UNIVERSITÀ DEGLI STUDI
DI MODENA E REGGIO EMILIA

Dottorato di ricerca in _Scienze Umanistiche

Ciclo XXXIV

Le fonti nelle società iperstoriche.

Analisi e critica dei processi di dematerializzazione, archiviazione e fruizione delle fonti digitali e digitalizzate per una diffusione sociale della storia.

Candidato

Scanagatta Manfredi

Relatore (Tutor):

Prof. Lorenzo Bertucelli

Coordinatore del Corso di Dottorato:

Prof.ssa Marina Bondi

Indice generale

Introduzione....2
1 - Una società Iperstorica....8
1.1 Tecnica, scientifica, digitale. Una rivoluzione?....8
1.2 Potere computazionale....21
1.3 Tecnologia. Essere tra....26
1.4 Iperstoria, un concetto cronosofico....34
1.5 Tecnologia, informazione, società....39
1.6 Big data una sfida quantitativa o qualitativa?....44
1.7 Digital divide e accessibilità....54
1.8 Da dato a informazione. Da evento a fonte....58
2 – Fonti e rivoluzione digitale....65
2.1 Da artefatti tangibili a oggetti digitali. La nuova forma del documento....65
2.2 Sovrapproduzione e metodo....68
2.3 Condivisione fisica e digitale....69
2.4 Condivisione e Infodemia....73
2.5 E le fonti materiche?....88
2.6 Presenza....89
2.7 Telepistemica e didattica....96
2.8 Archivi fisici e rivoluzione digitale....101
2.9 Archivi, una definizione?....105
2.10 Archivi in trasformazione....110
2.11 Chi ha paura dell'algocrazia?....115
2.12 Archivi correnti digitali....119
2.13 Medium-documento....126
3 ISAD (G) e ISAAR-CPF. Standard per una prospettiva internazionale degli archivi....137
3.1 ISAD(G) il padre degli standard....147
3.2 ISAAR (CPF), l'importanza dell'agente....151
3.3 EAD....156
3.4 EAC-CPF....166
4 Metadati....175
4.1 Standard....185
4.1.1 METS....189
4.1.2 MAG....191
4.1.3 PREMIS....194
4.1.4 Dublin Core Metadata Initiative....198
4.2 Omeka, una sperimentazione....206
5 Da OAIS al Web semantico....215
5.1 Interoperabilità....222
5.2 Il web semantico....226
RDF Schema OWL e ontologie....235
Bibliografia....249
Sitografia....260

Introduzione

Durante il seminario *Introduction to accademic writing,* tenuto dalla Professoressa Marina Bondi rimasi particolarmente colpito dall'accento che la Professoressa pose sulla necessità di individuare una nostra "tribù" di appartenenza, dunque quei ricercatori e professori che potessero essere interessati al nostro lavoro. La questione per quanto all'apparenza banale mi fece molto riflettere.
La prima domanda che mi posi era puramente funzionale, da chi posso trovare indicazioni e spunti che mi possano aiutare a sviluppare un impianto concettuale? Subito dopo però mi chiesi a chi sarebbe potuto interessare questo mio lavoro, in che modo avrei potuto provare a dare il mio contributo per approfondire un dibattito teorico e metodologico, che coinvolgesse una riflessione sulla digitalizzazione fonti?
La sensazione era che per trovare indicazioni utili a sviluppare il mio ragionamento avrei dovuto guardare al di fuori di quella che ritengo possa essere la mia tribù, non mi si fraintenda, storiche, storici e pubic historian stanno facendo lavori egregi nell'ambito delle Digital Humanities, nella costruzione e diffusone del contenuto storico a e con pubblici eterogenei, progetti sempre più strutturati e interessanti coinvolgono il lavoro di istituti e enti storici anche al di fuori dell'ambito accademico.
Ma osservando gli ambiti di ricerca e i progetti sviluppati da storici, storiche e public historian, salvo pochi casi, ho avuto la sensazione di trovarmi davanti alle fasi conclusive di un elaborazione teorica, concettuale e tecnica che aveva le proprie radici in ambiti differenti da quello della storia o della Public History.
La mia sensazione, in quanto tale assolutamente confutabile, era che in pochi, tra storiche e storici, si stessero interrogando sul valore reontologizzante dei processi di digitalizzazione e conseguentemente non ci si stesse chiedendo in che modo questa rivoluzione potesse coinvolgere il nostro lavoro, non solo da un punto di vista tecnico, ma epistemologico.
Cosa succede alle fonti quando da materiche subiscono un processo di dematerializzazione e di rimediazione, il contenuto informativo rimane immutato in seguito a questo processo? È corretto osservare in senso acritico la costruzione di archivi digitali e conseguentemente una nuova accessibilità alle fonti, o siamo chiamati a provare ad elaborare un impianto che ci aiuti a comprendere in che modo l'avvento delle società iperstoriche - quelle società in cui, secondo la definizione data da Luciano Floridi, è determinante la relazione con tecnologie ICT per la capacità

di produrre, scambiare e archiviare informazioni - possa ridefinire la nostra realtà se letta attraverso la lente dell'infosfera?

Per riuscire a dare un taglio di più ampio respiro al mio lavoro, mi sono dovuto chiedere cosa si intenda quando si parla di fonti nate digitali. Cosa sono, quali sono i processi oggi messi in atto per assicurarne la conservazione in modo che storiche e storici del futuro possano fare il loro lavoro?

E ancora, è vero che con la digitalizzazione dell'informazione si assiste ad un processo di disintermediazione o ci stiamo trovando davanti a pratiche che consentono una neointermediazione digitale, che avviene nelle fasi preliminari alla diffusione del contenuto informativo.

La storiografia sta già cambiando forma, ma i processi di narrazione e spiegazione che necessariamente ne costituiscono parte dell'ossatura stanno anch'essi mutando?

A queste domande ne seguì un'altra ben più fondamentale, a chi serve il lavoro che sto facendo, per chi sto scrivendo?

All'inizio di questo mio lavoro le domande erano sicuramente molte più delle risposte e invero poco è cambiato ora che sono giunto alla conclusione. Temo di essere riuscito a rispondere a pochi dei quesiti che mi sono posto, ma d'altra parte spero di aver elaborato un impianto di ragionamento che possa, in futuro, aiutarmi a trovare alcune risposte nella collaborazione con diverse discipline.

E' infatti nell'interdisciplinarità che ho trovato l'unica condizione possibile d'analisi.

Per comprendere il concetto di rivoluzione digitale mi sono affidato agli studi di filosofi dell'informazione e teorizzatori dell'ontologia sociale, solo così ho potuto comprendere che nel momento in cui mutano alcuni paradigmi di analisi della realtà, muta necessariamente anche l'ambiente in cui la realtà prende forma e che questo non è mai un passaggio indolore per chi si occupa di scienza.

Ma è altresì dovere dello scienziato riuscire ad elaborare dei concetti capaci di generare significazione sugli eventi, così da fornire degli strumenti a chi, da studente, osserva non senza perplessità le azioni di trasformazione della società, sapendo che i concetti e la società stanno in un continuo rapporto di tensione.

Questa rivoluzione va in prima istanza osservata a partire dagli atomi che la compongono, dai dati, elementi puri che esistono a prescindere dall'interpretazione e che si manifestano a noi come qualcosa la cui sola esistenza è espressione di una differenza. Il dato è la scintilla, l'elemento che consente la costruzione di una traccia, il dato è l'iscrizione che è alla base dell'oggetto sociale.

Senza i dati non possono esistere fonti e senza fonti non si può giungere alla conoscenza, fosse anche solo dell'evento.

Ma come possiamo fare, in una società che solo nel 2020 ha prodotto 64 zettabyte di dati, a estrapolare e a contestualizzare un'informazione comprensibile in senso qualitativo e non solo

quantitativo? E come possiamo immaginare di poter gestire l’archiviazione di tutti questi dati senza correre il rischio che si perdano nell’universo del Web?

E’ per provare a rispondere a queste domande che ho cercato l’aiuto di archivisti e informatici, che da anni lavorano insieme nella strutturazione di ciò che è definita archivistica informatica e che vede nell'elaborazione di standard internazionali, una strategia atta a costruire delle infrastrutture in prima istanza logiche, capaci di organizzare e veicolare la conoscenza che nello scorrere del tempo va a stratificarsi negli archivi.

La cosa sorprendente è che perché questo processo possa prendere forma è necessario avvalersi di uno strumento tecnico che è in prima istanza anch’esso un concetto, sono i metadati infatti ad essere l’elemento centrale del processo rivoluzionario che sta coinvolgendo l'organizzazione e la fruizione di informazioni digitali.

I metadati, semplicisticamente definibili come dati su altri dati, sono gli strumenti tramite i quali è possibile costruire dei processi di sovrascrittura atti a contestualizzare il dato e a rendere possibile una sua archiviazione ordinata e una sua fruizione. Senza metadati non esiste oggetto digitale.

Ogni fonte digitalizzata o nata digitale esiste nel Web come oggetto digitale, ciò implica che il portato informativo oltre ad aver vissuto una rimediazione, dunque una trascrittura, ha subito anche una sovrascrittura ed è in questo processo che noi in quanto storici e storiche dobbiamo cogliere l’elemento che maggiormente richiede una nuova analisi epistemologica delle fonti.

Grazie ai metadati le informazioni si stratificano e rendono possibile un processo di interoperabilità atto potenzialmente a costruire discorso storiografico. La rappresentazione dell'informazione che avviene all’interno di specifici processi di gestione dei dati, rende la fonte un oggetto informativo che grazie ai metadati descrittivi può essere individuato da un qualsiasi fruitore con l’utilizzo di semplici parole chiave.

La flessibilità di alcuni schemi di metadati consente di immaginare pratiche davvero innovative di archiviazione delle fonti, non solo perché consentono una catalogazione automatica, ma perché, se è vero come è vero che gli archivi sono strumenti sì di conservazione ma anche di comunicazione, la metadatazione descrittiva può divenire una nuova forma di spiegazione, che unita all’interoperabilità delle fonti può dare forma a complessi intrecci semantici in cui la rappresentazione simulacrale della fonte diviene essa stessa elemento narrativo del discorso storiografico.

Questa nuova forma di storiografia può apparire disintermediata, ma nella realtà dei fatti l’intermediazione avviene in una fase preliminare a quella della narrazione ma non di meno la determina.

Ed è in questa possibilità di neointermediazione nella costruzione dei processi informazionali, che si sviluppa il Web semantico come teoria atta a costruire una rete semantica, comprensibile dalle macchine, che colleghi tra di loro non più i file, ma i dati stessi.
I processi del Web semantico sono determinati dalla costruzione di triple composte da soggetto, predicato e oggetto. Tramite un'azione di incapsulamento il nucleo originario, il dato, mano a mano si stratifica di informazioni atte a renderlo autoesplicativo sia nei suoi contenuti, sia nella sua storia archivistica. Gli oggetti informativi divengono dati univoci e autoconsistenti.
I dati e le informazioni si liberano delle strutture degli standard, ma non ne perdono il portato logico, li rendono ancora più flessibili e interoperabili. Grazie a vocabolari comuni, che consentono una disambiguazione dei significati, e l'utilizzo delle triple, è possibile costruire ontologie digitali dove il predicato che unisce soggetto e oggetto è un arco che si manifesta come un concetto.
E' in questa nuova predominanza dei concetti all'interno del Web che immagino si possa inserire il lavoro di storiche e storici. Se i metadati possono essere la rappresentazione di parte del contenuto informativo della fonte, possiamo immaginare che dopo un'attenta analisi e critica delle fonti, storiche e storici attribuiscano ai dati, dei metadati che ho definito concettuali.
Ma in che modo questo potrebbe aiutare la diffusione sociale della storia?
Lavorare sui concetti consente di attuare un processo a mio avviso altamente funzionale per la comprensione della storia, ossia risemantizzare nell'oggi ciò che è evento nel passato e viceversa.
Mi trovo a scrivere queste pagine nei giorni in cui la Russia ha invaso l'Ucraina, difficile liberare la testa da questo pensiero. Ci stiamo trovando davanti ad un tipo di guerra che credevamo ormai passato, un conflitto che porta le vesti del secolo scorso, siamo nel cuore dell'Europa, ed è la Russia a muovere guerra per provare a riprendere spazio verso occidente. Rivendicazioni che fanno scorrere la memoria ad antichi pretesti ma che si mescolano con la rivoluzione energetica figlia dei nostri tempi.
L'invasione di terra di uno stato sovrano sorretta da una minaccia atomica rivolta verso il resto del mondo.
Nato e Unione Europea si trovano davanti ad una situazione complessa, la consapevolezza che un loro intervento militare in Ucraina non potrebbe che portare ad una escalation del conflitto, il timore di un allargamento del fronte della guerra. Impossibile non chiedersi cosa faranno Ungheria, Polonia, Bulgaria, Repubblica Ceca, si formerà un fronte unico contro la Russia o alcuni di questi paesi potranno essere tentati dalle idee di Putin. La Cina sfrutterà l'occasione per muoversi contro Taiwan? Quale la posizione dei paesi sul Baltico, perché Finlandia e Svezia si stanno preoccupando di dover difendere confini all'apparenza così distanti dal conflitto ad ora in atto? E ancora, quale sarà il ruolo della Turchia?

Queste domande per uno storico non possono che avere un valore speculativo, la nostra disciplina richiede la pazienza del tempo, la possibilità dello sguardo sul lungo periodo; il ruolo di storiche e storici non è quello di prevedere, fare ipotesi o giudicare, ma siamo chiamati e chiamate a descrivere gli eventi non tanto nel loro svolgersi, ma piuttosto in seguito alla loro conclusione.

Dobbiamo essere in grado di osservare i rapporti che si sviluppano tra causa ed effetto in senso retrospettivo e proiettivo, fornendo alla società delle chiavi di lettura per poter comprendere il passato così da essere cittadini e cittadine capaci di agire nel presente con la consapevolezza che le nostre azioni costruiranno il futuro.

Davanti a questi avvenimenti mi è inevitabile chiedermi quale sia il senso della mia ricerca, quale l'apporto che questo lavoro può avere in una prospettiva più ampia di quella squisitamente accademica?

Eppure, nonostante si parli di epistemologia, della differenza tra dati e metadati, di complessi standard archivistici e di ancor più complessi funzionamenti di un Web che aspira a diventare semantico, credo che tutto ciò possa avere un valore proprio in momenti così complessi, dove le informazioni si muovono in modo incontrollato e troppo spesso la storia viene piegata a pure necessità funzionali e dietrologiche.

I Sistemi Informativi come il SAN o il SIUSA, sono già oggi potenti strumenti dove poter trovare informazioni e contesti storiografici altamente attendibili e dove, in alcuni casi, è possibile accedere alla fonte primaria digitalizzata. Ma pensiamo alle possibilità di un Web dove le fonti storiche sono organizzate attraverso una contestualizzazione semantica comprensibile dalle macchine e dove i metadati consentono l’accesso alla fonte tramite l’utilizzo di poche parole chiave, quale il valore che i concetti possono assumere in questo processo.

In questi giorni mi è capitato di leggere un articolo di Paolo Valentino, nel pezzo, uscito per il Corriere della Sera il giornalista compie un’analisi sulla Russia di Putin, che tiene al centro il concetto di “culto della personalità”. Da storiche e storici sappiamo bene quanto questo concetto sia stato determinante nella costruzione delle grandi dittature del Novecento, a partire da Mussolini fino a Stalin passando per Mao e per Hitler, non vi fu dittatore che non costruì sulla sua figura un culto della personalità.

Pensiamo se a specifiche fonti, relative all’URSS, all’Italia fascista o alla Germania nazionalsocialista, fossero attribuiti dei metadati concettuali incentrati proprio sul culto della personalità; quanto sarebbe semplice per ogni fruitore, a prescindere dalle competenze pregresse, compiere una ricerca che fornirebbe come risultati le fonti primarie che storiche e storici hanno precedentemente metadatato con questo concetto. Certo, la narrazione non avverrebbe attraverso un

rapporto causale tra le fonti, tutto si svilupperebbe su un una linea tematica, ma ciò consentirebbe una comprensione davvero maggiore di ciò che oggi sta succedendo.

In questi giorni sono girate in rete molte immagini che mostrano l'espansione della NATO ad est come un'azione all'apparenza di conquista di territori un tempo afferenti all'URSS o al patto di Varsavia, queste fonti sono rappresentazione plastica della misinformation, processo che però può trovare un argine proprio grazie alla digitalizzazione e alla contestualizzazione delle fonti primarie. Se siamo nell'epoca della miscredenza delle competenze, far sì che non siano lo storico o la storica a parlare, ma che siano direttamente gli archivi potrebbe avere un impatto davvero rivoluzionario sui processi di diffusione sociale della storia, in prima istanza perché per la prima volta la potrebbe davvero rendere un bene comune.

Riportare l'attenzione alle fonti, unica vera rappresentazione che si ha della storia, ritengo che sia un grande passo avanti in società sempre più confuse e travolte da troppi punti di vista, dove *episteme* e *doxa* si confondono fino a creare un paradosso dove più nulla è vero perché tutto potenzialmente lo è.

«Un essere umano che non possieda né linguaggio, né abitudini, né memoria, che sia cioè privo di iscrizioni e di documenti, difficilmente potrebbe coltivare intenzioni, sentimenti e aspirazioni sociali.»

1 - Una società Iperstorica

1.1 Tecnica, scientifica, digitale. Una rivoluzione?

Macchina Automatica - no anima
Macchina Automatica - no anima
Ecco la Terra in Permanente Rivoluzione
Ridotta imbelle sterile igienica
Una Unità di Produzione
Unità di Produzione
Tecnica d'Acciaio
Scienza Armata Cemento
Tabula Rasa Elettrificata

G. L. Ferretti[1]

Non si faccia l'errore di descrivere, osservare e vivere la rivoluzione digitale come un processo che coinvolge solo o prevalentemente alcuni strumenti del nostro esistere, sarebbe come sostenere che la rivoluzione industriale sia stata tale, rivoluzionaria, solo perché ha cambiato gli strumenti e le modalità del lavorare, come se avere a che fare con sistemi meccanizzati di lavoro e di produzione non avesse modificato ogni fase della vita. I tempi del lavoro cambiano e di conseguenza cambiano i tempi della vita, la produzione modifica il concetto di consumo che a sua volta modifica l'idea stessa di profitto, come scrive Reinhart Koselleck:

La seconda rivoluzione industriale, si dice, non si limita a liberare il mondo del lavoro umano dallo sforzo fisico, ma affida già alcuni processi mentali a macchine automatiche. La cibernetica, la microfisica e la biochimica ricadono sotto il concetto di una seconda rivoluzione industriale che supera di gran lunga la

1 Testo tratto dal brano dei C.S.I. (Consorzio Suonatori Indipendenti), *Unità di produzione*, dall'album, *Tabula rasa elettrificata*, Polygram 1997

prima, in cui si trattava ancora di aumentare la produttività umana al di sopra dei bisogni tradizionali, con il capitale, la tecnica e la divisione del lavoro.[2]

Se davvero fossimo disposti a osservare le grandi rivoluzioni industriali e tecniche come processi che hanno modificato solo il modo di lavorare e di produrre beni, concentrandoci dunque sulla necessità di analizzare gli strumenti, le parole di Herbert Marcuse risuonerebbero nel vuoto.
«Una confortevole, levigata, ragionevole, democratica non-libertà prevale nella civiltà industriale avanzata, segno di progresso tecnico»[3].
Queste sono le righe con cui Herbert Marcuse apre *L'uomo a una dimensione. L'ideologia della società industriale avanzata*, e non lasciano scampo al lettore. Che si condivida o meno il pensiero dell'autore ciò che immediatamente ci si pone davanti è la correlazione tra civiltà industriale avanzata, progresso tecnico e assenza di libertà.
Non è mio interesse concentrarmi sul valore delle critiche espresse dal filosofo riguardo la società industriale, ma è mia volontà cogliere i livelli di profondità necessari a comprendere non tanto le cause, ma gli effetti di lungo periodo dell'industrializzazione della società.
Anche se in una prospettiva storica, cosa che Marcuse non smette mai di sottolineare, il pensiero dell'autore vuole da subito portare l'attenzione del lettore sulle implicazioni dell'industrializzazione della società, un processo avvenuto in seguito a una rivoluzione tecnica, resa possibile da una società dove l'iniziativa privata trovava sempre più margine d'azione e i capitali economici potevano essere investiti per dare forma a nuovi modi di creare e gestire il lavoro.
Marcuse porta delle critiche molto chiare agli effetti che l'industrializzazione ha avuto sulla società, per il pensatore siamo stati assoggettati, incapaci di esprime dissenso abbiamo accettato il progresso perché inevitabile, ci siamo privati della possibilità della critica, abbandonati all'alienazione non siamo stati capaci di reagire alla «Ragione» imposta dalla produzione, fino ad assistere inermi
«al tramonto delle forze storiche che, nella fase precedente della società industriale, parvero rappresentare la possibilità di nuove forme di esistenza»[4].
Le riflessioni del filosofo, intense e cariche di idealismo, ci descrivono come vittime di sistemi coercitivi che hanno annullato il nostro Io fino a sminuzzare lo spazio privato. La tecnologia come strumento di controllo, alienazione depotenziamento dell'essere umano.
Triste dunque il nostro passato tecnologico e, se non volessimo discostarci da questa visione assolutista, spaventoso il nostro futuro fatto di *device* tecnologici dai quali non riusciamo più a staccarci, noi eradicati dalla vita reale ci rifugiamo inconsapevoli in ambienti virtuali, dove il tutto è

2 R. Koselleck, *Futuro passato. Per una semantica dei tempi storici*, Clueb, Bologna 2007, p. 56
3 H. Marcuse, *L'uomo a una dimensione. l'ideologa della società industriale avanzata*, Einaudi, Torino 1970, p. 21
4 Ivi, p. 30

finzione, le relazioni fittizie, i sensi annullati, il controllo delle nostre vite demandato ad algoritmi che decidono anche chi può dire cosa e tutto è gestito e mediato attraverso gelidi schermi che ci rendono atomi soli.

Distopico e spaventoso questo modo di osservare il futuro, o meglio il presente.

Ciò che ci dobbiamo chiedere è se sia corretto definire l'innovazione tecnologica contemporanea con il concetto di rivoluzione digitale e mettendo in relazione l'analisi critica di Marcuse verso le rivoluzioni industriali, con il pensiero di alcuni filosofi contemporanei come Giorgio Agamben, sempre di più mi convinco del portato rivoluzionario della trasformazione digitale che stiamo vivendo.

Un elemento a favore di questa riflessione può essere il terrore con cui alcuni osservatori stanno descrivendo il nostro contemporaneo.

«Fa parte della barbarie tecnologica che stiamo vivendo la cancellazione dalla vita di ogni esperienza dei sensi e la perdita dello sguardo, durevolmente imprigionato in uno schermo spettrale.»[5]

E riferendosi all'università e alla didattica digitale

[…] Di ogni fenomeno sociale che muore si può affermare che in un certo senso meritava la sua fine ed è certo che le nostre università erano giunte a tal punto di corruzione e di ignoranza specialistica che non è possibile rimpiangerle e che la forma di vita degli studenti si era conseguentemente altrettanto immiserita. Due punti devono però restare fermi:

1. i professori che accettano – come stanno facendo in massa – di sottoporsi alla nuova dittatura telematica e di tenere i loro corsi solamente *on line* sono il perfetto equivalente dei docenti universitari che nel 1931 giurarono fedeltà al regime fascista. Come avvenne allora, è probabile che solo quindici su mille si rifiuteranno, ma certamente i loro nomi saranno ricordati accanto a quelli dei quindici docenti che non giurarono.[6]

In questi estratti di un articolo molto criticato del filosofo contemporaneo Giorgio Agamben ritroviamo molti termini simili a quelli utilizzati da Marcuse per descrivere l'infausta ricaduta sulla società dei processi di meccanizzazione e industrializzazione. L'assunto di base è che più il

5 G. Agamben, *Requiem per gli studenti*, originariamente pubblicato il 23/05/2020 Per l'Istituto Italiano per gli Studi Filosofici https://www.iisf.it/index.php/attivita/pubblicazioni-e-archivi/diario-della-crisi/giorgio-agamben-requiem-per-gli-studenti.html) ma adesso non più disponibile a questo link. E' possibile ritrovare l'articolo su Corpi e Politica https://www.corpiepolitica.it/2020/05/24/requiem-per-gli-studenti/ e sulla pagina FaceBook di IISF https://www.facebook.com/studifilosofici/posts/2184226538369148, ultima consultazione 27/04/2021. Quando l'articolo uscì suscitò molte polemiche anche perché seguiva una altro articolo riguardante il biopotere e terrore sanitario in relazione alle misure di contenimento varate a causa della pandemia di Covid-19 https://www.quodlibet.it/giorgio-agamben-biosicurezza ultima consultazione 27/04/2021

6 Ibidem

progresso avanza ed entra in relazione con le nostre vite, più ne saremo schiavi e dunque costretti ad assoggettarci alle volontà di un controllo politico ed economico che tenderà sempre e comunque a limitarci come esseri umani liberi. Agamben si spinge addirittura a paragonare l'accettazione della didattica digitale a distanza, dovuta a una pandemia, a chi accettò passivamente il fascismo. La barbarie tecnologica va combattuta, è necessario opporsi a una rivoluzione non voluta, imposta, fascista.

Ma se fosse più complesso di così, se l'innovazione tecnologica e digitale ci richiedesse lo sforzo di osservare gli avvenimenti e l'organizzazione dell'informazione in modo differente, se ci stesse ponendo in relazione al mondo in un modo assolutamente nuovo che ancora ci spaventa perché incapaci di comprenderlo nella sua potenziale infinita articolazione?

Per provare a cercare una spiegazione, può essere utile "rifugiarci" nelle parole di Thomas Samuel Kuhn, che con una dolcezza filosofica invidiabile si limita a porre l'attenzione sulle figure degli scienziati e degli studenti, cercando in poche e semplici parole di descrivere il contesto che si genera mentre una rivoluzione scientifica è in atto

> […] dopo un mutamento di paradigma, gli scienziati non possono non vedere in maniera diversa il mondo in cui sono impegnate le loro ricerche. Nei limiti in cui i loro rapporti con quel mondo hanno luogo attraverso ciò che essi vedono o fanno, possiamo dire che, dopo una rivoluzione, gli scienziati reagiscono ad un mondo differente […] Soltanto quando il suo modo di vedere le cose ha subito trasformazioni di questo genere, lo studente entra a far parte del mondo dello scienziato, riesce a vedere le cose che lui vede e a reagire come lui.[7]

Mentre il cambiamento è in atto è difficile scorgerlo e ancor più complesso comprenderlo, i pochi che hanno questa possibilità sono gli scienziati, ossia coloro che stanno direttamente agendo nel cambiamento, i quali però possono non avere risposte certe su quelli che saranno gli effetti della rivoluzione, conoscono ciò che sta cambiando ma possono solo immaginare le reali implicazioni. Fuori dal mondo degli scienziati, ma in relazione con essi, vi sono gli studenti, ossia tutti coloro che non sono attori del cambiamento ma che lo possono osservare nel suo verificarsi, anche se appunto per riuscirci devono essere predisposti a spostare il paradigma con cui solitamente osservano il loro contemporaneo, il nuovo non può essere guardato con le stesse lenti con cui si osservava il vecchio.

Non illudiamoci però che questo scarto sia semplice e indolore.

7 T. S. Kuhn, *La struttura delle rivoluzioni scientifiche,* Einuaudi, Torino 2009, pp. 139-140

Il mondo in cui lo studente entra in quel momento, però, non è un mondo stabilito una volta per tutte dalla natura dell'ambiente, da un lato, e dalla natura della scienza, dall'altro. Esso, piuttosto, è determinato unitamente dall'ambiente e dalla particolare tradizione di scienza normale che lo studente ha imparato a seguire. Perciò in periodi di rivoluzione, quando la tradizione della scienza normale muta, la percezione che lo scienziato ha del suo ambiente deve venire rieducata: in alcune situazioni che gli erano familiari deve imparare a vedere una nuova *Gestalt*.[8]

Dopo aver letto le riflessioni di Kuhn, sono da osservare in modo differente le critiche di Marcuse e Agamben ai mutamenti sociali dovuti a specifiche rivoluzioni tecnologiche e scientifiche, a partire dal fatto che Marcuse costruisce la sua critica ponendo una distanza temporale tra sé e i processi rivoluzionari, distanza che gli consente di osservare gli effetti negativi che tali rivoluzioni hanno avuto sulla società. È altresì vero che nella sua riflessione non tiene conto di quegli effetti inconfutabilmente positivi che seguirono i processi di rivoluzione. Basti pensare che nel 1800 la popolazione mondiale aveva raggiunto, dopo secoli di sviluppo, il primo miliardo. Nel 1900 si sono aggiunti altri seicentotrenta milioni di persone. Nel 1930, nonostante gli eventi della prima guerra mondiale, verrà raggiunto il secondo miliardo, il terzo appena trent'anni dopo. L'aspettativa di vita a metà ottocento era di poco più di quarant'anni nei paesi sviluppati e meno di trenta in quelli più arretrati. Nel 1900 supera i cinquant'anni nei paesi avanzati e rimane pressoché identica in quelli meno sviluppati.[9] In Italia oggi le aspettative di vita raggiungono circa gli ottanta anni.[10]
Marcuse, anche se non utilizza tutti i dati a sua disposizione, avanza una critica basandosi sull'osservazione e l'analisi degli effetti che l'industrializzazione della società ha avuto sui cittadini, il suo è dunque un pensiero che può non essere condiviso ma che ha preso forma dopo aver accettato il cambio di paradigma dato dalla rivoluzione, quando quindi, nonostante fosse uno studente è riuscito a osservare la realtà dello scienziato.
Agamben pone in atto un processo molto differente, la sua non è un'analisi, ma un'opinione, sicuramente rilevante dato la caratura culturale di chi la esprime, ma è opinione di uno studente ancora non in grado di osservare quello che sta succedendo. Con questo non voglio dire che la rivoluzione digitale ha e avrà solo aspetti positivi, in quanto studente non potrei farlo se non anche io esprimendo opinione, certo meno valida di quella di Agamben. Quello che intendo fare è provare a non utilizzare delle lenti categoriche per osservare il cambiamento, cercando dunque di pormi solo delle domande, sperando siano giuste, che mi aiutino a descrivere quello che sono in grado di osservare.

8 Ivi, p. 140
9 M. Flores, *Il secolo mondo*, il Mulino, 2002 p. 30
10 https://www.istat.it/it/archivio/255785

Dopo questa digressione, spero utile a comprendere come la paura del non conosciuto e del nuovo possa modificare l'analisi e il giudizio di eventi troppo complessi per essere vittime di classificazioni dicotomiche buono/cattivo, vorrei riprendere il mio ragionamento ripartendo dalla stessa domanda: si può parlare di rivoluzione digitale? E se sì perché?
Prima di procedere nel tentare di comprendere se quella digitale possa essere definita una rivoluzione, vorrei fermarmi brevemente sul concetto stesso di rivoluzione. Per Kosseleck tra le altre cose studioso di storia dei concetti, rivoluzione è un termine che indica sia il rovesciamento violento o la guerra civile, sia trasformazioni a lunga scadenza, dunque eventi e strutture che penetrano profondamente nella nostra vita quotidiana. Non solo, il pensatore tedesco ci suggerisce che nella stessa parola rivoluzione sia insita una forza così rivoluzionaria da permettere di espandersi continuamente e abbracciare tutto quanto esiste nel globo.

> La rivoluzione diventa un concetto metastorico, solo che si stacca interamente dalla sua origine naturale e ora tende a dare un ordine storico alle esperienze sovvertitrici. In altre parole la rivoluzione assume un aspetto trascendentale, diventa un principio regolativo sia per la conoscenza sia per l'azione di tutti gli uomini che ne sono coinvolti. Il processo rivoluzionario e la conoscenza della rivoluzione, che ne è influenzata e retroagisce su di esso, da questo momento in poi sono indissolubilmente congiunti. Tutti gli altri contrassegni del concetto moderno di rivoluzione si alimentano in questo significato misterioso che sta sul suo sfondo.[11]

Possiamo dunque dire che il termine rivoluzione può essere usato solo se il cambiamento in atto arriva a coinvolgere in modo sensibile tutta la società. Questo implica che la rivoluzione non può essere osservata nello strumento, ma è necessario ampliare lo sguardo, aumentare i livelli di astrazione e chiedersi se e come la tecnica possa modificare il vivere quotidiano di ognuno di noi.
Quello che è semplice sostenere è che la tecnica interviene reinventando alcuni processi della società, impattando inizialmente solo su azioni specifiche compiute dall'essere umano, ma che con il tempo inizia a modificare abitudini e pensiero dando nuova forma alla realtà. Questo ci aiuta a comprendere come l'innovazione tecnologica, che materialmente agisce nell'invenzione o nel rimodellamento di specifici strumenti, in realtà arrivi a modificare pensieri e azioni delle società nelle quali avvengono questi mutamenti, coinvolgendo in modo più o meno marcato la vita di tutti gli esseri umani che abitano queste società[12].
Tra il 1730 e il 1760 John Kay realizza la "navetta volante", un marchingegno in grado di spostare meccanicamente e non più a mano il filato di cotone sul telaio e quindi di produrre stoffe di

11 R. Koselleck, *Futuro passato. Per una semantica dei tempi storici*, Clueb, Bologna 2007, p. 63
12 L'utilizzo al plurale del sostantivo "società" facendo dunque presumere risultati differenti in differenti società tra di loro contemporanee è necessario per poi comprendere come Floridi divida le società tra preistoriche, storiche e iperstoriche non attraverso una lente temporale

ampiezza maggiore rispetto a quella delle braccia del tessitore, ciò comportò un cospicuo aumento della produttività e come è facile da immaginare, o meglio come ormai ben sappiamo, diede una spinta innovativa tale che nel giro di pochi anni si vedranno realizzati i telai meccanici a energia manuale, ma soprattutto i telai idraulici.

Questo schema a «botta e risposta», che coinvolge e stravolge il settore tessile inglese, implica che ad ogni processo innovativo segua una domanda che richiede un nuovo progresso tecnologico. Ciò arriva ad avere degli effetti anche sull'industria siderurgica attraverso innovazioni che si impongono con maggiore lentezza, ma che pongono le basi per una radicale trasformazione nell'assetto industriale urbanistico e sociale inglese.

I processi sul trattamento della ghisa che ottiene Henry Cort tra il 1783 e il 1784 permettono di produrre una ghisa più malleabile e di sottoporla a una torchiatura così da poterla trasformare in barre di ferro. E' così che sarà possibile utilizzare il ferro per realizzare strutture complesse di notevoli dimensioni, come i nuovi telai. Questo portò a un miglioramento degli impianti e la meccanizzazione dei processi industriali venne nuovamente stimolata.

La necessità sempre maggiore di carbone portò la scienza e la tecnica a interrogasi su come poter più facilmente pompare in superficie l'acqua che risiedeva in fondo a pozzi di carbone sempre più profondi. La macchina inventata da Savery nel 1680 disperdeva troppa energia, lo stesso problema si riscontrò anche con la pompa ideata da Newcomen, più sicura e potente, ma ancora lenta nel funzionamento e ad alta dispersione energetica. Sarà nel 1775, grazie all'unione dell'ingegno di James Watt alle capacità finanziarie di Mattew Bolton, che vedrà la luce la prima macchina a vapore.[13]

L'impiego del vapore come principale forma di energia e le adozioni delle innovazioni tecniche nelle produzione tessile cambiarono il paesaggio e la società inglese, la necessità di concentrare le macchine e i lavoratori nelle fabbriche sconvolse la geografia e i costumi di vita, modificando anche le vie di comunicazione terrestri e fluviali. Nel 1804 la prima locomotiva a vapore trainò 10 tonnellate di minerali e 70 passeggeri a 8 chilometri orari. Tra il 1800 e il 1820 in Inghilterra vennero costruiti 320 chilometri di ferrovia, i trasporti ormai non riguardavano più solo i minerali o le lavorazioni industriali, molti furono i passeggeri a usufruire di questo nuovo modo per spostarsi.

Nel 1807, in seguito alla realizzazione di una fitta rete di canali artificiali e navigabili, venne creato il primo battello a vapore.

13 D.S. Landes, *Prometeo liberato. Trasformazioni tecnologiche e viluppo industriale nell'Europa occidentale dal 1750 ai giorni nostri*, Einaudi, Torino 2017

La realizzazione di uno strumento il cui utilizzo era legato unicamente a poter spostare meccanicamente del cotone, in circa 50 anni porta a una profonda trasformazione dei sistemi di produzione, trasporto e comunicazione in tutta l'Inghilterra.

Ma questi non sono gli unici cambiamenti significativi, la svolta tecnologica infatti non impatta solo l'ambito dell'economia, della produzione e dei trasporti, ma coinvolge l'insieme delle gerarchie dei valori e delle relazioni sociali. I lavoratori impiegati nelle nuove fabbriche rappresentano rispetto alle epoche precedenti un'incredibile novità, in relazione ai ritmi produttivi cambiano le abitudini, le mentalità e i modi di vita degli strati di popolazione chiamati a far funzionare il nuovo *moloch*[14] dell'industria. Ciò portò alla creazione di nuove forme di associazionismo – luddismo - e di resistenza che nei primi anni dell'ottocento presero la forma di atti di violenza e di sommosse atte a distruggere i macchinari e le fabbriche, creando nuovi sistemi di conflitto sociale fino a sperimentare le *Trade Unions*, unioni di mestiere a metà strada tra le associazioni di mutuo soccorso e i sindacati modernamente intesi[15].

Nel 1866 a Ginevra, il Congresso Internazionale dei Lavoratori – la prima Internazionale, nata nel 1864 – propose otto ore come limite legale dell'attività lavorativa, il primo maggio del 1867 in Illinois venne approvata una legge che istituisce le otto ore lavorative. Nel 1884, la Federation of Organize Trades and Labor Unions, sindacato confederato americano, chiese ai movimenti operai di unirsi per rivendicare la riduzione delle ore di lavoro.[16]

Nel 1882, quasi due miliardi di persone viaggiavano annualmente per ferrovia, nelle regioni sviluppate dell'occidente dovevano esserci a quella data ben poche persone che una volta o l'altra non fossero entrate in contatto con la rete ferroviaria. Probabilmente solo la rete di linee telegrafiche sorrette da file interminabili di pali di legno lunga tre o quattro volte il sistema ferroviario mondiale era nota più universalmente.

Ma anche l'immaginario paesaggistico e romantico del telegrafo con le sue lunghe linee di pali e cavi è destinato a scomparire. Era il 1896, dopo una anno di esperimenti per inviare onde lunghe alla distanza di un chilometro, quando Marconi si rivolse al British Post Office per vendere il sistema telegrafico senza fili che il Ministero delle Poste e Telecomunicazioni italiano si rifiutò di prendere in considerazione.

---

14 Utilizzo qui il termine Moloch riprendendo la metafora espressa da Allen Ginsberg in *Howl*, poema pubblicato nel 1956

15 F. Benigno, *L'età moderna. Dalla scoperta dell'America alla restaurazione,* Editori Laterza, Roma Bari, 2006 p. 350

16 L. Pezzica, *Giorni di Storia. Festa dei lavoratori.* prodotto multimediale realizzato da DHLab http://memoria.lazio.it/iniziative/giorni-di-storia/festa-dei-lavoratori/ ultima consulktazione 02/05/2021

Nel 1897 Marconi fondò la Wireless Telegraph and Signal Company. Nel 1901 ricevette il primo messaggio transatlantico senza fili inviato dalla Cornovaglia[17].

Negli ultimi anni del 1800 sono state inventate e sviluppate le turbine e motori a combustione interna, il telefono, il grammofono, la lampadina elettrica a incandescenza, le automobili rese per tutti da Daimler e Benz, per non parlare della cinematografia e dell'aeronautica[18].

Il processo di rivoluzione tecnologica e scientifica ha una chiara ricaduta anche nell'ambito sociale, o come scrive Hobsbawm nelle «statistiche morali»[19].

L'alfabetismo era palesemente in aumento, alcuni esempi sono che al momento delle guerre napoleoniche le lettere spedite in Inghilterra erano forse due l'anno per cittadino, mentre tra il 1880 e il 1885 erano circa 42, nel 1880 negli Stati Uniti uscivano ogni mese 186 milioni di copie di giornali e periodici, mentre nel 1788 erano solo 330 mila, ma ancor più interessante è che le persone con interessi scientifici aderenti alle società di cultura britannica erano 44 mila, probabilmente quindici volte di più di cinquant'anni prima[20].

Creighton, storico e vescovo anglicano, come riportato dalla Cambridge Modern History, scrisse «Non possiamo non presupporre, come ipotesi scientifica in base alla quale è stata scritta la storia, un progresso nelle cose umane»[21].

Massimo Livi Bacci attraverso un'analisi di natura demografica ci aiuta a comprendere gli effetti delle rivoluzioni tecnologiche e scientifiche dell'ottocento spostando il nostro sguardo anche sulle campagne, dove la produttività della terra era poco dinamica e l'introduzione del capitale e della meccanizzazione accrebbero la produttività del lavoro, implicando un forte aumento di eccedenti di manodopera, spesso dislocata da attività e modi di vita tradizionali, improvvisamente messi in crisi[22].

E' sempre a cavallo tra la fine dell'ottocento e i primi anni del novecento che si moltiplicano i modi per trattare e prevenire le malattie batteriche, con sieri, antitossine e vaccini. Nel 1901 venne isolato il virus della febbre gialla, le terapie basate su teorie cellulari diedero impulso ai rimedi sintetici e allo sviluppo dell'industria farmaceutica. Nascono la Bayer, Ciba, Hoechst, Sandoz, e questo portò a un aumento della produzione di pillole e medicamenti e la loro distribuzione di massa, l'aspirina comparve nel 1899.

Nel 1901 viene introdotto l'elettrocardiografico, si diffonde negli ospedali l'utilizzo dei raggi X, individuati nel 1895 da Wilhelm Rontgen, scoperta a breve distanza seguita da quella della

17 M. Flores, *Il secolo mondo,* il Mulino, 2002 p. 33
18 E.J. Hobsbawm, *L'età degliimperi. 1875 1914,* Editori Laterza, Roma Bari, 2000
19 Ivi, p. 35
20 ibidem
21 *Cambridge Modern History,* Cambridge 1902, I, p. 4
22 M.L. Bacci, *Storia minima della popolazione mondiale,* Il Mulino, Bologna 2005, p. 173

radioattività da parte di Henri Becquerel e Pierre e Marie Curie che probabilmente si attestò come la più grande scoperta avvenuta a cavallo tra l'ottocento e il novecento e che verrà compresa in modo sensibile come una scoperta di utilità collettiva[23].

In circa centocinquanta anni tutto il mondo vive dei cambiamenti impressionanti, che avranno ricadute importanti anche nell'ambito bellico, dove tecnologia vuol dire potere, anche se sarà solo nei primi anni del novecento e in particolar modo con la Prima Guerra mondiale che sarà possibile osservare i nefasti frutti dell'innovazione tecnica e scientifica su un campo di battaglia.

In queste pagine si è sintetizzato alcune delle fasi che maggiormente hanno determinato ciò che è stato definito dagli storici "rivoluzioni industriali"[24] - dando quindi una definizione cronosofica ad una periodizzazione - la prima, che ha coinvolto prevalentemente l'Inghilterra e che prese forma nella seconda metà del settecento coinvolgendo in prima istanza l'industria tessile e sviluppando i primi sistemi di motorizzazione attraverso il carbone. La seconda, con una ricaduta intercontinentale anche se coinvolse prevalentemente le società nordoccidentali e che vide nell'evoluzione della chimica e dell'utilizzo dei carburanti fossili il centro di un cambiamento sociale ed economico senza precedenti.

In questa breve disamina non sono state affrontate le innovazioni tecnologiche e scientifiche arrivate con la prima e la seconda guerra mondiale.

In tempi di guerra le economie spostano i loro assi di interesse, vengono nazionalizzate e lo sforzo è tutto incentrato verso il conflitto attraverso la difesa dei propri confini e l'annullamento dell'avversario. In tempi di pace invece l'innovazione è spinta da volontà differenti, si muove più liberamente, non si concentra solo su alcuni settori e come abbiamo potuto vedere piccole innovazioni possono avere una ricaduta sensazionale in campi molto diversi.

Le innovazioni date dalla guerra hanno bisogno che la guerra finisca perché possano essere apprezzate dalla popolazione, mentre ciò che avviene in tempi di pace ha un'immediata ricaduta sulla società, una società che non è stressata dal conflitto e che dunque può interagire in modo molto più costruttivo con il cambiamento tecnologico.

Il processo rivoluzionario non può che essere compreso sul lungo periodo, ma è indubbio che per essere così definito sia necessario che coinvolga diversi aspetti della società e che abbia un impatto diretto sui cittadini.

La rivoluzione, se è rivoluzione, deve riguardare tutti.

---

23 M. Flores, *Il secolo mondo,* il Mulino, 2002 pp. 29-30

24 Nel terzo capitolo ci soffermeremo su come l'attribuzione del termine "rivoluzione" sia funzionale alla storiografia e non alla storia, e di come sia esemplificativa di quanto il linguaggio determini i processi di spiegazione storiografica.

Prima di descrivere in modo più dettagliato gli ambiti in cui la rivoluzione digitale ha portato e sta portando cambiamenti molto profondi, vorrei concentrarmi su un altro dato che a mio avviso lega la rivoluzione contemporanea con quelle del passato.

E' grazie alla reazione a catena, che prima abbiamo chiamato *botta e risposta* - riprendendo la definizione data da Hobsbawm – che può prendere forma un processo innovativo capace di modificare aree molto differenti dell'esistere. L'invenzione di un macchinario per l'industria tessile inglese, attraverso una reazione a catena, porta alla costruzione di una macchina a vapore, che ci conduce fino alla rivoluzione dell'industria siderurgica europea, che arriva a condizionare gli esiti di una guerra mondiale.

In quanto storici siamo abituati a osservare gli avvenimenti con lo sguardo di lungo periodo, così come siamo consapevoli dei rapporti di causalità che determinano ogni evento/processo.

Ciò che però non sempre consideriamo è come e quanto ricerche e innovazioni in settori specifici, anche molto specialistici, possano apportare cambiamenti decisivi in settori che nulla hanno a che fare con il campo di appartenenza in cui la ricerca si origina.

Per meglio spiegare ciò che intendo proverò a riassumere brevemente la vita e l'apporto dato allo sviluppo di ARPANET da parte di uno scienziato del quale, nonostante un ruolo decisivo nello sviluppare l'idea della rete, non avevo mai sentito parlare prima di confrontarmi con l'ottimo lavoro fatto da Detti e Lauricella in *Le origini di internet*[25].

Joseph Carl Robnett Licklider nel 1942 ottenne un Phd in psicologia alla University of Rochester con una tesi dal titolo *An Electrical Study of frequency Localization in the Auditory Cortex of the Cat.* Con il suo lavoro aveva misurato la grandezza, la frequenza e la localizzazione dell'attività elettro-neurale corticale nel cervello degli animali in risposta ai toni pervenuti all'orecchio[26].

Se è vero che il suo lavoro non continuò in questa direzione, è altresì vero che continuò a lavorare nell'ambito della percezione e dell'udito. Nel 1948 studiò le prestazioni del sistema uditivo attraverso auricolari quando l'intelligibilità della parola è disturbata dal rumore (il problema del cocktail party). Al tempo stesso però i suoi interessi si estesero dalle reti esterne di comunicazione tra esseri umani stabilite mediante l'emissione e la ricezione di segnali sonori a quella sulla rete interna che collega le orecchie al sistema nervoso centrale[27].

In seguito ai suoi studi riguardanti due reti di comunicazione interfacciate dalle orecchie: una formata da più esseri umani connessi tra loro mediante la voce e l'udito, l'altra costituita dal sistema

25 T. Detti, G. Lauricella, *Le origini di internet*, Mondadori, Milano 2013

26 R.M. Fano, *Joseph Carl Robnett Licklider,* in National Research Council Biographical memories V.75, https://www.nap.edu/read/9649/chapter/12#191 ultima consultazione 29/04/2021

27 T. Detti, G. Lauricella, *Le origini di internet*, Mondadori, Milano 2013 p. 68

nervoso, nel 1950 sintetizzò il senso del lavoro svolto dall'Acustic Laboratory e dal Research Laboratory of Electronic del MIT usando queste parole:

Il problema su cui questo gruppo sta lavorando è più o meno il seguente. C'è un certo numero di esseri umani connessi da link di comunicazione in modo da formare una rete. Hanno accesso a determinate informazioni e un compito da svolgere, che richiede un intervento basato sull'informazione. La domanda è: in quale modo la loro performance nello svolgimento del compito dipende dalla configurazione della rete?[28]

Persone connesse l'una all'altra attraverso link, l'informazione e la capacità di elaborarla come oggetto di valutazione del funzionamento e infine la modalità di organizzazione della rete come elemento decisivo per la capacità di scambiarsi e elaborare informazioni.
Se prima di leggere queste righe non avessimo conosciuto l'oggetto degli studi di Licklider, sarebbe stato davvero difficile non cadere nell'errore di credere che queste sue parole fossero rivolte a commentare il funzionamento di una rete per come oggi intendiamo internet.
Teorie dell'informazione e della comunicazione iniziano a intrecciarsi con la ricerca di natura psicologico cognitiva e trovano un loro punto di incontro nelle innovazioni date dalle apparecchiature elettroniche e dalla conseguente ricerca sperimentale riguardo la cibernetica.
Secondo gli studi di Detti e Lauricella anche in questo caso le ricerche teoriche hanno come volano l'industria bellica.

A porre la macchina al centro della psicologia di quegli anni e a far crescere al suo interno una «psicologia cibernetica» era innanzi tutto il suo oggetto, costituito dal fattore umano di sofisticati sistemi uomo-macchina come gli aerei, i sommergibili, i carri armati e i servomeccanismi che dovevano gestirli o neutralizzarli.[29]

Nella mia ricerca affronterò il tema di quanto e di come la digitalizzazione delle fonti storiche materiche e la costruzione di archivi digitali ci consenta già oggi di sviluppare ricerca storica coadiuvati dall'interoperabilità delle fonti, possibile anche all'interno di confederazione di archivi e di come questo processo metta al centro il tema dell'accessibilità e dell'informazione. Anche se studenti ancora forse non del tutto in grado di cogliere il cambio di paradigma e dunque di osservare il mondo per come lo riescono a vedere gli scienziati, osserveremo anche le implicazioni dello sviluppo del Web Semantico, come una tecnologia che tende ad assegnare significati concettuali alle rete che collega le informazioni. Questo non può che portarci a riflettere riguardo il valore e l'importanza della *spiegazione* nell'elaborazione di contenuti storiografici, in particolar modo

28 Ivi, p. 69
29 Ivi, p. 70

quando gli oggetti digitali – le fonti – sono archiviati mediante l'utilizzo di metadati descrittivi, che uniti a specifici linguaggi rendono possibile la creazione di ontologie digitali all'interno degli archivi.
Il tema è oggi al centro dell'interesse scientifico di diverse discipline, dall'informatica, alla filosofia dell'informazione, passando per la linguistica fino alle digital humanities e appunto la storia.
Ma se oggi siamo capaci di elaborare sistemi di astrazione analitica così complessi in relazione a procedure tecniche informatiche altrettanto complicate è anche grazie alle intuizioni di Licklider, uno psicologo che studiava la relazione tra suoni e percezioni che iniziò a immaginare esseri umani connessi in una rete di link atta allo scambio e all'analisi di informazioni e che alla fine degli anni cinquanta arrivò a teorizzare una rete di *thinking centers*

> […] una delle possibilità interessanti aperte dal concetto di una rete di *thinking centers* è quella di una discussione tecnica *all'interno* del sistema. Se diversi centri sono connessi tra loro da canali di comunicazione ad alta capacità, posiamo immaginare una conversazione telefonica notevolmente potenziata dall'accesso, alle due estremità della linea, alle strutture del sistema. Oppure possiamo immaginare una conferenza facilitata dal fatto che sarebbero immediatamente disponibili biblioteche di informazione tecnica e anche i consigli di consulenti specializzati[30].

Secondo le intuizioni di Licklider la conoscenza poteva essere messa e organizzata in rete. Ora, se questo è senza dubbio uno dei principi base su cui si è sviluppato internet per come lo conosciamo oggi, non possiamo non farci affascinare dall'idea dei *thinking service*, se proviamo a osservare i nodi di comunicazione di questa rete non tanto come produttori di informazione, ma come archivi di informazioni che se uniti consentono l'accesso da un unico portale a una conoscenza incredibilmente vasta e eterogenea.
Nei prossimi capitoli proveremo a descrivere come possa essere possibile dare istruzioni, che per noi hanno natura semantica, in un linguaggio comprensibile dalle macchine e che consente una rielaborazione meta-sintattica dei contenuti, così da rendere possibile un'interazione uomo macchina incredibilmente funzionale anche nell'analisi di contenuti umanistici.
Ad ora abbiamo potuto osservare tre elementi che secondo me possono essere distintivi di ciò che chiamiamo rivoluzione industriale, scientifica o digitale.
Il ruolo degli intellettuali, indicatori precisi di quanto un processo di innovazione tecnica possa essere o già sia anche un processo di innovazione sociale. La reazione dell'intellettuale alle innovazioni può essere un metro interessante per comprendere se ciò che si sta mostrando come

30 Ivi, p. 78

novità determinerà o meno un cambiamento nella società. Più forte sarà il grido di allarme, maggiormente saranno definiti gli schieramenti tra favorevoli e contrari, più complesso sarà il ragionamento che tenta di descrivere ciò che sta cambiando e più probabilità ci sono che effettivamente la rivoluzione non riguardi solo lo strumento, ma l'esistenza dell'essere umano.
Per comprendere il valore del concetto di rivoluzione abbiamo volto lo sguardo verso la storia, verso la sua complessità data dai lunghi periodi e da costanti relazioni causali. Abbiamo osservato come le grandi rivoluzioni industriali tecniche e scientifiche del passato si siano verificate a partire da una nuova e all'apparenza quasi insignificante invenzione, che ha determinato però cambiamenti così profondi da modificare i paesaggi del continente europeo e nord americano, cambiando per sempre il nostro rapporto con il tempo e con lo spazio. Infine ci siamo confrontati con l'ibridazione, ibridazione di idee, tecniche, competenze e obiettivi. l'idea che solo nella transdisciplinarità sia possibile trovare soluzioni tanto innovative da arrivare a modificare l'esistenza delle persone, che a loro volta, vivranno l'innovazione come una ibridazione esistenziale, prima tra uomo e meccanica, ora tra uomo e digitale.
Nulla è fisso, tutto è in divenire, esattamente come la storia.

1.2 Potere computazionale

Per arbitrate una partita di tennis servono sei giudici di linea, tre per parte del campo, e un giudice di sedia, un tempo vi erano anche i giudici di rete che tenendo la mano appoggiata sul bordo bianco superiore della rete chiamavano la ripetizione del servizio "let!" se la palla aveva sfiorato il nastro. Il giudice di rete, è stato sostituito ormai da molti anni da un sensore che invia un segnale se la palla sfiora il nastro. Questo ha permesso di rilevare anche il più leggero contatto tra la palla e la rete e ha tolto un essere umano da una posizione in cui ricevere una "pallata" non era poi possibilità così remota.
Durante i quarti di finale degli US Open 2004, una tecnologia utilizzata solo a fini televisivi, per dare allo spettatore una moviola digitale in tempo reale di dove cadeva la palla, divenne un elemento determinate per la carriera della giudice di sedia Mariana Alves. Nel corso di quella partita Serena Williams contestò una serie di chiamate del giudice e i successivi replay televisivi dimostrarono che aveva ragione la giocatrice. Alla Alves fu impedito di arbitrare altri match durante

il torneo newyorkese e i suoi errori spinsero a valutare l'introduzione in campo della tecnologia finora utilizzata solo per le moviole televisive.[31]

Nel finire della stagione 2005 l'International Tennis Federation ha testato il sistema dell'*Hawk-Eye* a New York City e dopo il controllo di ottanta tipi di tiro misurati dall'ITF high speed camera, è stato approvato l'utilizzo a livello professionistico.

Gli US Open nel 2006 annunciarono che ogni giocatore avrebbe avuto diritto a due controlli (o *challenge*) per set.

Nel 2007 anche Wimbledon, il più importante e "conservatore" torneo del circuito ATP, ha introdotto l'Hawk-Eye sul Centre Court e il Campo 1, concedendo ad ogni tennista tre controlli per set. In caso di tie-break viene assegnato un challenge extra, inoltre nell'ultimo set dove non esiste il tie-break vengono assegnati altri tre controlli ogni 12 game.

Fino al marzo 2008 l'ITF, ATP, WTA e i tornei del Grande Slam avevano introdotto delle regole diverse sull'utilizzo dell'Hawk-Eye, quali il numero di challenge permessi per set.

Il 19 marzo 2008 le varie federazioni hanno uniformato il sistema con le regole attuali: tre challenge per set con uno extra in caso di tie-break. Il primo torneo ad aver adottato le nuove regole è stato il Sony Ericsson Open 2008.[32]

Gli US Open 2020 sono stati il primo torneo del Grande Slam a utilizzare l'Hawk-Eye senza la presenza dei giudici di linea, con il sistema che analizzava istantaneamente il gioco ed era collegato all'altoparlante che subito chiamava out la palla uscita dal campo. Fu usato però solo nei campi secondari del torneo, mentre nei due campi principali si continuò ad avere i giudici di linea. All'Australian Open 2021 questa soluzione è stata utilizzata per la prima volta in tutti i campi.

Le partite di tennis di uno dei tornei più importanti del mondo sono state arbitrate da dalle macchine. Niente più giudici di linea, presente solo il giudice di sedia il cui ruolo però è quasi solo di presenza, sono le macchine infatti che osservano dove cada la palla, valutano se sia dentro o fuori e attribuiscono il punto.

Se ci dovessimo chiedere se una telecamera ha una accuratezza maggiore dell'occhio umano nel registrare informazioni, la risposta sarebbe no, perché l'occhio vede meglio di una telecamera, ma per un essere umano è impossibile non avere mai neanche un decimo di secondo di distrazione, mentre per le macchine questo non è contemplato.

Hawk-Eye funziona tramite sei, a volte sette telecamere ad alta prestazione, normalmente posizionate sul lato inferiore del tetto dello stadio, che seguono la palla da diverse angolazioni.

31 https://www.popularmechanics.com/adventure/sports/a5772/cameras-fouls-and-referees/#sidepanel ultima consultazione 17/05/2021

32 http://news.bbc.co.uk/sport2/hi/tennis/7305404.stm ultima consultazione 17/05/2021

Il video delle sei telecamere viene quindi triangolato e combinato per creare una rappresentazione tridimensionale della traiettoria della palla. Hawk-Eye non è infallibile, ma dovrebbe avere una scarto di errore entro 3,6 millimetri.[33]

Ciò che avviene è dunque che delle macchine riprendono la palla da diverse angolazioni, per poi triangolare le informazioni che hanno ricevuto, trasformare queste informazioni in una rappresentazione digitale e tridimensionale del movimento della palla per poi restituire queste immagini digitali ai giocatori su un grande schermo presente nello stadio, contemporaneamente la stessa informazione viene condivisa con chi sta guardando la partita da casa che sia in televisione o tramite internet grazie a canali di streaming digitale.

Per rendere possibile questo processo e per trovare il modo di far interagire Hawk-Eye con la partita in tempo reale, le più importanti federazioni internazionali di tennis hanno modificato, o meglio, aggiunto dei regolamenti andando *de facto* a cambiare alcuni funzionamenti delle regole del gioco e contemporaneamente gli stadi sono stati modificati in modo da rendere l'ambiente adatto al funzionamento delle macchine.

Massimo Durante sintetizza così quella che definisce la tesi di base del suo testo *Potere computazionale:* «Il potere computazionale si esercita adattando o pretendendo di adattare a poco a poco, non solo il mondo ma anche la rappresentazione della realtà al modo di funzionamento delle tecnologie digitali dell'informazione e della comunicazione (ICT) che esso stesso alimenta»[34]

Il potere computazionale è dato dalla capacità di processare dati e trasformarli in risultati, questo processo è di ogni soggetto capace di interagire con l'ambiente. L'essere umano ha un grande potere computazionale semantico, siamo incredibilmente capaci di dotare di senso il mondo e le nostre visioni.

Le tecnologie digitali non hanno capacità semantiche, ma sono ottimi motori sintattici con una straordinaria capacità di calcolo, siamo noi esseri umani a dotare di significato i calcoli e i processi sintattici che sempre noi chiediamo alle macchine di svolgere.

Quello che stiamo chiedendo alla tecnologia è dunque un nuovo modo di poter definire il mondo e di essere in grado di osservarlo, per farlo inevitabilmente mutiamo la nostra relazione con la realtà, modifichiamo l'ambiente e le macchine stesse, al fine di ottenere nuovi risultati, nuovi output, modificando la nostra capacità di interagire e collezionare gli input.

La pallina da tennis che rimbalza un centimetro fuori dalla riga di fondo campo è l'input che il nostro occhio riesce a leggere e che nel contesto di una partita, dopo l'elaborazione computazionale

33 https://www.hawkeyeinnovations.com/sports/tennis ultima consultazione 17/05/2021
34 M. Durate, *Potere computazionale. l'impatto delle ICT su diritto società, sapere,* Meltemi, Milano, 2019 p. 11

del giudice di linea fa sì che questo chiami un “out!” trasformando dunque l’input in un output che porta all'assegnazione del punto al giocatore che non ha sbagliato.
Attraverso Hawk-Eye l’input rimane lo stesso, ma si modifica la capacità computazionale che oltre a portare come output l’assegnazione del punto al tennista, rende a noi esseri umani una rappresentazione digitale dell’evento.

L’esercizio del potere computazionale implementato in una molteplicità di sistemi, non si traduce soltanto in un progressivo adattamento dell’ambiente, con una conseguente trasformazione del mondo, che investe settori sempre più diffusi e rilevanti delle nostre società, ma si traduce anche e soprattutto in una vera e propria produzione del sapere. Questa produzione del sapere concerne, da una parte, la formazione della base di conoscenza di cui i sistemi computazionali hanno bisogno per assolvere i loro compiti, (la loro peculiare rappresentazione della realtà) e, dall’altra la raffigurazione che tali modelli computazionali offrono di noi stessi (la loro peculiare rappresentazione di noi).[35]

Nel suo libro *La quarta rivoluzione*, un testo capace di descrivere in modo profondo e puntuale le trasformazioni che stanno avvolgendo la nostra società, Luciano Floridi si chiede se ci sia la possibilità di individuare una prospettiva attraverso la quale interpretare i fenomeni maggiormente rilevanti di questa trasformazione – le nanotecnologie, l’internet delle cose, il Web 2.0, il Web semantico, il cloud computing ecc. - come aspetti di un unico trend macroscopico.
Una delle domande che si pone Floridi è dunque molto simile alla domanda portante di queste prime pagine: si può parlare di rivoluzione digitale? E se sì, perché?

Parte della difficoltà con cui rispondiamo a questa domanda risiede nel fatto che siamo abituati a considerare le ICT come strumenti mediante i quali interagiamo con il mondo e tra noi. In realtà, tali tecnologie sono divenute forze ambientali, antropologiche, sociali e interpretative. Esse creano e forgiano la nostra realtà fisica e intellettuale, modificano la nostra autocomprensione, cambiando il modo in cui ci relazioniamo con gli altri e con noi stessi, aggiornano la nostra interpretazione del mondo e fanno tutto ciò in maniera pervasiva, profonda e incessante.[36]

Nel suo lavoro Floridi intitola i primi tre capitoli rispettivamente “Tempo”, “Spazio” e “Identità”[37].

35 ibidem
36 L. Floridi, L*a quarta rivoluzione. Come l’infosfera sta trasformando il mondo,* Raffaello Cortina editorie, Milano, 2017 p. IX
37 I titoli completi sono “Tempo: l’iperstoria”; “Spazio: l’infosfera”; “Identità: l’onlife”

Con questa triade così definita è chiaro che il filosofo ci stia già facendo capire come e quanto la rivoluzione in atto sia pervasiva nelle nostre esistenze e di come possa essere in grado di modificare la nostra realtà, a partire da quella individuale fino ad arrivare a quella sociale e politica.
Da studiosi della storia non si può che essere attratti dall'idea stessa di "Tempo" che tanto ha a che fare con la nostra disciplina ma che è così complessa da afferrare. In quanto storici tendiamo ad analizzare gli eventi in senso diacronico, dandogli dunque una lettura legata alla relazione tra spazio e tempo, dove *prima* e *dopo* sono avverbi che assumono un valore di significato determinante all'interno della costruzione del discorso storico. In occidente e per riconoscimento dell'ONU il tempo del mondo e dell'umanità viene definito essenzialmente in prima di Cristo e dopo Cristo.
Curiosa la nostra difficoltà a definire il tempo, considerando che è l'astrazione con cui abbiamo deciso di descrivere forse il primo significato della nostra esistenza.

«Finché l'uomo è desto la sua vita in sviluppo in continua realizzazione è rappresentata nella sua coscienza dall'elemento del divenire (questo dato si chiama presente) e come divenire essa ha la misteriosa caratteristica della direzione che l'uomo in tutte le lingue più evolute ha cercato di dominare intellettualmente mediante la parola tempo e i problemi ad essa connessi»[38]

Il tempo è dunque per noi qualcosa di misurabile in senso lineare, per riuscire a procedere in questa lettura da sinistra verso destra - in una rappresentazione grafica il futuro è sempre a destra - ci aiutiamo con le periodizzazioni, mettendo dunque su questa linea retta una serie di date che arrivano a caratterizzare in senso generico delle epoche, all'interno delle quali noi troviamo e comprendiamo l'evoluzione del pensiero, dell'arte, della politica.
Jacques Le Goff nelle sue riflessioni sul tempo continuo della storia ci porta a riflettere sul valore e la necessità delle periodizzazioni.
«Molto prima di aver ottenuto diritto di cittadinanza nella storiografia e nella ricerca storica, il concetto di periodo era stato utilizzato per organizzare il passato»[39].
Il "Tempo" di Floridi, che come vedremo viene definito per alcune società contemporanee come *iperstorico*, mentre per altre *storico e preistorico*, esiste attraverso una periodizzazione che non è temporale, ma tecnica.
Per comprendere l'iperstoria, credo che prima sia necessario confrontarsi con lo "Spazio", comprendere cioè in che modo la tecnica modifica l'ambiente delle nostra vite, ambiente che il filosofo definisce come *infosfera.* Dove l'identità viene definita *onlife.*

38 O. Spengler, *Il tramonto dell'occidente*, Longanesi, Milano, 2012, p. 92
39 J. Le Goff, *Il tempo continuo della storia*, La Terza, Bari-Roma 2022, p. 9

È sul concetto di tecnologia che ci dobbiamo soffermare se vogliamo comprendere come tempo, spazio e identità possano modificarsi all'interno del processo rivoluzionario che stiamo vivendo.
Come scrive Durante «si sono, in effetti, susseguite così tante e profonde concezioni della tecnologia dal novecento a oggi che non è neppure pensabile renderne conto»[40], pretendere dunque di dare un resoconto complessivo di come filosofia e tecnologia si siano confrontati nel corso del novecento è veramente difficile e non è neanche il reale interesse di questa ricerca, per questo andremo a lavorare su quelle che oggi sono considerate le riflessioni maggiormente interessanti e innovative sul pensare la tecnologia in diretta relazione con i cambiamenti tecnologici di natura digitale.

1.3 Tecnologia. Essere tra

Peculiarità della tecnologia è il suo *esere-tra* un suggeritore e un utente. Floridi trova un modo semplice per spiegare questa relazione e fa l'esempio del sole e del cappello, il cappello è la tecnologia che viene creata dall'uomo per inserirsi tra l'uomo stesso e il sole, dove il sole è il suggeritore, ossia l'elemento che ci fa porre la domanda e l'uomo è l'utente che ha bisogno della tecnologia per coprirsi[41].
Per capire quale sia il rapporto tra tecnologia e ambiente è necessario rendere più complessa la relazione tra suggeritore e utente andando a individuare una possibile tripartizione delle tecnologie che vengono definite come tecnologie di primo, secondo e terzo ordine.
Questa tripartizione non coincide sempre con una progressiva evoluzione delle tecnologie in una rappresentazione cronologica[42].
Le tecnologie di primo, secondo e terzo ordine possono coesistere nello stesso tempo, così come in una visione globalizzata individuando dunque come spazio il mondo intero, possono coesistere nello stesso tempo le società preistoriche, storiche e iperstoriche.
La differenza che esiste tra tecnologie di primo, secondo e terzo ordine è data dall'identità del suggeritore e dell'utente.
Quando i suggeritori sono naturali e gli utenti sono esseri umani si è di fronte a tecnologie di primo ordine.

40 M. Durate, *Potere computazionale. l'impatto delle ICT su diritto società, sapere,* Meltemi, Milano, 2019 p. 47
41 Luciano Floridi, *la quarta rivoluzione. Come l'infosfera sta trasformando il mondo,* Raffaello Cortina editorie, Milano, 2017 p. 27
42 M. Durate, *Potere computazionale. l'impatto delle ICT su diritto società, sapere,* Meltemi, Milano, 2019 p. 53

Un semplice esempio per comprendere cosa sia una tecnologia di primo ordine è dato dalle scarpe, dunque da un artefatto tecnologico utile all'uomo per poter camminare su tutte le superfici senza ferirsi, un altro esempio che porta Floridi è l'aratro, così come l'arco.
Le tecnologie di primo ordine non sono necessariamente tecnologie più basiche rispetto a quelle di secondo e terzo ordine, un fucile d'assalto è sempre una tecnologia di primo ordine perché si inserisce tra l'essere umano e un altro essere umano, mentre una tecnologia di secondo ordine è un cacciavite, tecnologicamente meno sviluppata rispetto a un fucile d'assalto, ma che che interviene tra l'essere umano e un'altra tecnologia, in questo caso la vite.
Le tecnologie di secondo ordine dunque si inseriscono tra l'essere umano e un'altra tecnologia, dove l'essere umano è l'utente e la tecnologia il suggeritore. Le tecnologie di secondo ordine mediano tra l'umanità e la tecnologia stessa.
Oltre all'esempio del cacciavite che media il rapporto tra l'essere umano e una vite, è facile comprendere il significato delle tecnologie di secondo ordine pensando alle chiavi, che mediano la relazione tra l'essere umano e il blocchetto della serratura.
La rivoluzione industriale di cui abbiamo provato a tracciare alcuni riferimenti nel primo paragrafo è contraddistinta dallo sviluppo di tecnologie di secondo ordine.

Il motore, inteso come ogni tecnologia che fornisce energia a un'altra tecnologia, è probabilmente la più importante tecnologia di secondo ordine. Mulini ad acqua e a vento hanno trasformato l'energia in movimento per millenni, ma è stato soltanto a partire dal momento in cui il motore a vapore, a combustione interna e quello elettrico sono diventati fornitori di energia "portabile", collocata tra le altre tecnologie e gli utenti ovunque questi ne avessero bisogno, che la rivoluzione industriale ha potuto tradursi in una realtà diffusa[43].

L'era digitale nella quale ci stiamo addentrando ha come elemento decisivo della sua propria peculiarità la possibilità di individuare un terzo grado tecnologico, dove utente e suggeritore sono entrambi tecnologie.
Le tecnologie di secondo ordine sono legate all'emergere di forme più complesse di civilizzazione e socializzazione, dal momento che presuppongono la capacità di costruire tecnologie a partire da altre tecnologie e sono un tratto distintivo della specie animale uomo.
A partire dalla fine del settecento come abbiamo potuto vedere è la tecnologia a svolgere il principale ruolo di suggeritore, stimolando in noi bisogni e sollecitazioni. Ciò che Hobsbwam definiva come uno schema a botta e risposta.

43 Luciano Floridi, *la quarta rivoluzione. Come l'infosfera sta trasformando il mondo,* Raffaello Cortina editorie, Milano, 2017 p. 30

Alcune tecnologie di primo ordine sono inutili in mancanza delle corrispondenti tecnologie di secondo ordine alle quali son associate, Floridi ci fa capire bene questo punto facendoci riflettere sul fatto che le strade come tecnologie di primo ordine, non hanno bisogno delle macchine per essere utili, mentre una vite ha bisogno del cacciavite.

Anche le tecnologie di secondo ordine implicano un certo grado di reciproca dipendenza con quelle di primo ordine, e questo denota l'avanzamento organizzativo e la capacità di specializzarsi di una società. «O abbiamo dadi e viti o nessuno dei due.»[44]

Ad ora abbiamo visto che nelle tecnologie di primo ordine la relazione si sviluppa tra

natura – tecnologia – umanità, dove quindi le tecnologie mediano il rapporto tra umanità e natura.

Nelle tecnologie di secondo ordine la relazione si trasforma in

umanità – tecnologia – tecnologia, dove quindi le tecnologie mediano il rapporto tra umanità e tecnologia.

Il nuovo scarto rivoluzionario avviene nel momento in cui è la tecnologia a mediare tra una tecnologia come suggeritore e una tecnologia come utente.

Lo schema delle tecnologie di terzo ordine si descrive quindi con un rapporto

tecnologia – tecnologia – tecnologia.

L'essere umano dunque non ha più un ruolo all'interno del processo, ma sul processo.

Ad oggi la maggior parte delle ICT sono tecnologie di terzo ordine e l'umanità diventa parte integrante di un ambiente che è ora definito in misura crescente in termini tecnologici. Le tecnologie di terzo ordine concorrono a definire in modo sempre più decisivo e pervasivo l'ambiente in cui gli esseri umani vivono e interagiscono.

Vi è dunque una terza articolazione del rapporto tra tecnologia e ambiente: in quest'ultimo caso la tecnologia non è più ciò che ci consente di adattarci all'ambiente né ciò che ci consente di estrarre energia dall'ambiente da destinare ad uno scopo ulteriore e diverso. È ciò che concorre a conformare l'ambiente stesso.[45]

Per fare un esempio semplice da comprendere basti pensare alle tecnologie domotiche, che trasformano le case in ambienti intelligenti, capaci per esempio di controllare e regolare in modo sempre più preciso il riscaldamento e la fornitura di acqua calda nelle nostre case. Oppure se vogliamo possiamo tornare sul campo da tennis dell' Austalian Open e vedere come telecamere e sensori riescano ad interagire tra di loro senza la presenza umana, modificando a tal punto l'ambiente, da arbitrare una partita di tennis.

44 Luciano Floridi, *la quarta rivoluzione. Come l'infosfera sta trasformando il mondo,* Raffaello Cortina editorie, Milano, 2017 p. 30

45 M. Durate, *Potere computazionale. l'impatto delle ICT su diritto società, sapere,* Meltemi, Milano, 2019 p. 58

Le tecnologie di terzo ordine funzionano grazie ad un continuo scambio di dati che permette di generare informazioni. Quello che sta succedendo oggi è che le ICT sono in grado di scambiarsi automaticamente dati tra di loro e dunque di generare informazione che ha una conseguenza diretta sulle nostre esistenze. L'informazione è alla base di ogni processo epistemico e da sempre l'essere umano genera e scambia informazione con altri esseri umani e animali non umani.
Gli stessi telai meccanici analogici funzionavano in base a delle informazioni che venivano scambiate tra l'essere umano e le macchine, oggi l'essere umano non è più al centro di questo scambio di informazioni nel senso che non c'è più bisogno della sua presenza, non è più un elemento decisivo nella produzione di informazioni, come nota il filosofo francese Michel Serres, la rivoluzione digitale sta mutando il rapporto tra soggetto e oggetto.
Il soggetto è sempre stato concepito come un esser umano pensante e senziente capace di determinare il modo in cui la realtà oggettiva è sperimentata e conosciuta, mentre l'oggetto è sempre stato concepito come qualcosa che viene determinato dal soggetto e da questo può essere manipolato.

> L'essere umano è stato costantemente recepito come l'unica fonte di informazioni rilevanti dal punto di vista politico, giuridico, economico, morale e sociale […] gli oggetti del mondo meritano di regola attenzione e rispetto, nella misura in cui il loro valore è il riflesso di un interesse umano. L'epistemologia moderna è, pertanto, alla base di una rappresentazione asimmetrica tra soggetto e oggetto.[46]

Questa asimmetria si disvela in modo inequivocabile difronte ai progressi dell'esplorazione spaziale. Il 19 aprile 2021, all'interno della missione spaziale della Nasa su Marte, il dispositivo di volo Ingenuity Mars Helicopter, si è staccato dal suolo marziano e ha effettuato il primo di sei voli. Ciò che rende questo evento di particolare interesse per il nostro ragionamento è che data la distanza che vi è tra la Terra e Marte, non è possibile comandare a distanza il piccolo velivolo poiché le informazioni inviate dalla base sulla terra tramite onde radio non possono essere comunicate in diretta al dispositivo su Marte, il quale dunque è stato capace di procedere nella sua missione senza essere governato da un essere umano.
Ingenuity in costante comunicazione con il Rover Perseverance - robot anch'esso su Marte che ha trasportato il dispositivo di volo in una zone del pianeta con caratteristiche che potessero facilitare il decollo - seguendo una serie di informazioni ricevute in anticipo, è riuscito a decollare in assoluta autonomia, senza che nessuno lo potesse controllare, tutto il sistema informativo è stato gestito da delle tecnologie digitali capaci di prendere delle decisioni e di svolgere compiti autonomamente.[47]

46 Ivi, p. 36
47 https://www.jpl.nasa.gov/news/press_kits/mars_2020/launch/ ultima consultazione 29/05/2021

Gli stessi ingegneri della NASA presenti nella base di Cape Canaveral non erano nient'altro che spettatori.[48] Tecnologia – tecnologia – tecnologia, ecco il terzo ordine nella sua espressione massima.

Difficile vedere un collegamento tra un drone che spicca automaticamente il volo mentre si trova su Marte e i processi di digitalizzazione, ordinamento e diffusione delle fonti storiche, eppure, come vedremo nell'ultimo capitolo, il Web semantico consente la possibilità, tramite una serie di indicazioni date anticipatamente alle macchine e collegate direttamente all'oggetto digitale, di fornire al fruitore una conoscenza maggiore rispetto alle richieste espresse. Questo perché le macchine grazie ad un processo di contestualizzazione, "capiranno" ciò che gli stiamo chiedendo e ci forniranno tutte le informazioni in loro possesso con la capacità di disambiguare la nostra richiesta e dunque la loro risposta.

La rivoluzione digitale non può essere osservata in senso strumentale, dunque come un processo che ha generato tecnologie utili a svolgere specifici compiti, perché nella realtà dei fatti ha modificato e sta modificando il nostro modo di fare esperienza del mondo e della conoscenza.

Le nostre vite sono quotidianamente in relazione con strumenti digitali che producono e scambiano informazione attraverso la produzione e lo scambio di dati che possiamo fruire grazie a software che a loro volta hanno la stessa forma dei dati che veicolano, sono bit.

Ciò significa che siamo parte di un ambiente che è sempre più tecnologicamente conformato, siamo parte di un tutto in costante interazione con gli agenti e i dispositivi che popolano questo ambiente, rispetto al quale viene meno l'idea di entrata e di uscita, come si trattasse di un ritaglio dello spazio: mondo analogico, mondo digitale, mondo virtuale sono espressioni che non fanno più riferimento a ragioni dello spazio, ma a modi dell'esperienza[49].

La rivoluzione digitale ha coinvolto in modo così determinante lo spazio e l'ambiente che Floridi ha sentito l'esigenza di dover coniare un termine che riuscisse a definire la costante relazione informazionale che vi è tra agenti umani e tecnologici con il termine *infosfera*.

Scarpe Nike e iPod comunicano tra di loro già dal 2006, grazie alle calzature *Nike+* connesse all'iPod nano attraverso lo *Sport Kit Nike+iPod*, le informazioni su tempo, distanza, calorie bruciate e ritmo vengono archiviate sull'iPod e visualizzate sul display; un feedback ascoltabile in tempo reale viene inoltre fornito attraverso le cuffie[50]. Oggi sul mercato è possibile trovare L'Apple Watch Nike, un "orologio" capace di collezionare e condividere con l'utente tutti i dati sulla corsa in

48 https://mars.nasa.gov/technology/helicopter/#Milestones ultima consultazione 29/05/2021
49 Ivi, p. 40
50 https://www.apple.com/it/newsroom/2006/05/23Nike-and-Apple-Team-Up-to-Launch-Nike-iPod/ ultima consultazione 29/05/2021

tempo reale e contemporaneamente sostituire tutte le funzionalità dello smartphone e chiaramente essere sempre connesso ad un sistema di tracciamento GPS.[51]

Apple ha appena commercializzato un nuovo prodotto che si chiama *AirTag*[52], dei dischetti bianchi lucidi da 31,9 millimetri di diametro, con uno spessore di 8 millimetri, che appartengono alla categoria dei tracker, consentendo dunque il tracciamento di oggetti che non vogliamo perdere. Si possono mettere nel portafoglio, oppure con degli accessori venduti da Apple possono essere allacciati a uno zaino, a un mazzo di chiavi, a un trolley.

Per ritrovarli, basta aprire l'app "Dov'è", che fino ad ora si poteva utilizzare per localizzare i propri device Apple, ma solo se localizzati alla rete. Con gli *AirTag* il funzionamento è molto più preciso e affidabile. Questo perché oltre al bluetooth, la localizzazione funziona grazie a "Posizione precisa", che ci guida verso il punto esatto con indicazioni tipo bussola sul display dell'iPhone sfruttando giroscopio e accelerometro.

L'indicazione è così precisa grazie alla tecnologia Ultra Wideband, che sfrutta impulsi di energia a radiofrequenza. La bussola digitale ci dà informazioni molto puntuali, come direzione e distanza, e se proprio non riusciamo a finire la caccia al tesoro si può attivare un piccolo suono. E se siamo fuori dal raggio di azione del bluetooth, o abbiamo dimenticato le chiavi in ufficio entra in gioco l'aspetto più interessante. Apple sfrutta tutta la sua base di device iOS, arrivata a un miliardo, e fa sì che se un iPhone passa vicino al nostro AirTag, lo riconosca. L'informazione viene mandata sul cloud di Apple e nella nostra app "Dov'è" la posizione verrà aggiornata.

Se perdiamo l'AirTag, sempre da "Dov'è" si può impostare la "modalità smarrito". In questo modo si riceve una notifica quando viene ritrovato e a scelta dell'utente si può impostare un numero di telefono e un messaggio da far apparire sui device Apple che passeranno nelle vicinanze, in modo che ci possano avvertire.

La nostra costante relazione informazionale con strumenti digitali ci "trasforma" in ciò che Floridi definisce *inforganismi*, una rete composta da miliardi di esseri umani e miliardi di strumenti digitali costantemente interconnessi uno all'altro.

L'AirTag utilizza la tecnologia RFID, identificazione a frequenza radio[53], che può immagazzinare e recuperare da remoto dati da un oggetto, assegnando a esso un'identità unica, come un codice a barre, ciò che impareremo a conoscere come URI e che come vedremo è possibile utilizzare in un programma di archiviazione digitale, dove i documenti digitali, gli oggetti digitali, potranno essere messi in relazione con altre fonti se collegate attraverso un URI.

51 https://www.nike.com/it/help/a/apple-watch-nike ultima consultazione 29/05/2021
52 https://www.apple.com/it/airtag/ ultima consultazione 21/05/2021
53 https://rfid.it/apple-punta-sulla-ultra-wideband-con-i-nuovi-airtag/ ultima consultazione 21/05/2021

Acronimo inglese di Radio Frequency Identification, l'RFID è la tecnologia di identificazione automatica basata sulla propagazione nell'aria di onde elettro-magnetiche, consentendo la rilevazione univoca, automatica (hand free), massiva e a distanza di oggetti, animali e persone sia statici che in movimento.

L'RFID interviene a monte della filiera del dato, così da diventare l'abilitatore dell'infrastruttura IoT (Internet of Things), la scintilla per l'interconnessione di persone, oggetti, servizi e processi[54].

È sufficiente incorporare questo minuscolo microchip in qualsiasi cosa, ivi compreso un essere umano o animale, per ottenere un Itente[55].

Il dato che ci dovrebbe far riflettere è che la produzione mondiale di RFID è aumentata di 25 volte tra il 2005 e il 2010, raggiungendo i 33 miliardi di tag, nel 2012 il valore globale di RFID ammontava a 7,67 miliardi di dollari, si stima che nel 2022 arriverà a quasi 22 miliardi[56].

Tramite questi minuscoli tag è possibile dare ad ogni oggetto fisico una identità digitale; uno dei campi in cui gli RFID sono molto utilizzati è nell'ambito della gestione dei documenti cartacei[57], una volta applicato il tag ad un documento questo può iniziare ad avere una doppia vita, una fisica e una digitale, sarà dunque possibile creare dei database in cui riportare i tag dell'RFID e da un punto di vista gestionale sarà impossibile perdere il documento così da rendere l'archiviazione e la fruizione molto più sicura e semplice.

Questo processo di identificazione digitale di documenti cartacei può apparire avveniristico nel campo dell'archiviazione e gestione di documenti storici, ma in realtà è ben poca cosa rispetto al cambiamento che sta coinvolgendo anche il mondo degli archivi e delle fonti.

RFID infatti consente di rintracciare e organizzare i documenti che comunque rimangono materici dunque non cambiano nulla nel loro statuto epistemico e ontologico, intenzione di questa ricerca è di andare ad osservare cosa succede quando il documento di archivio subisce una rimediazione, muta il suo stato, dematerializzandosi e consentendo un modo completamente nuovo di accesso e fruizione dei documenti.

Su questo tema ci confronteremo più avanti e andremo ad osservare come la costruzione di standard di archiviazione, l'utilizzo di metadati e la trasformazione della fonte in un oggetto informativo, porti a processi assolutamente innovativi di ordinamento e fruizione delle informazioni. Certo è che già oggi le fonti digitalizzate o digitali e gli archivi digitali, sono a tutti gli effetti elementi propri dell'infosfera.

---

54 https://www.rfidglobal.it/tecnologia-rfid/ ultima consultazione 21/05/2021

55 Luciano Floridi, *la quarta rivoluzione. Come l'infosfera sta trasformando il mondo,* Raffaello Cortina editorie, Milano, 2017 p. 49

56 ibidem

57 https://www.rfidglobal.it/case-history-gse/ ultima consultazione 21/05/2021

La sfida sarà comprendere come lo storico possa divenire un inforganismo capace di sviluppare nuove e interessanti modalità di relazione con l'archivio e le fonti stesse. Questo processo ci porterà a dover riflettere su termini quali storia e storiografia, reimmaginando un apparato teorico che possa condurre il lavoro dello storico all'interno della rivoluzione digitale con una nuova prospettiva teoretica, dove le informazioni che per il lavoro dello storico sono definite fonti, non sono più prodotte e ordinate solo da esseri umani, ma direttamente da delle tecnologie.

Le fonti per lo storico del futuro - e in parte anche per la storica del presente - che siano fonti digitalizzate o fonti digital born, saranno dati. Ogni file di testo che quotidianamente studiamo non è altro che un insieme di dati, intelleggibili per l'essere umano perché letti da delle macchine alle quali abbiamo detto di elaborare i dati in un contesto di significato che ci rende possibile la comprensione.

I processi di digitalizzazione della realtà, che come abbiamo potuto vedere stanno coinvolgendo in modo complesso tutto l'ambiente che ci circonda e l'esperienza che facciamo delle cose, consegnano la nostra era ad un nuovo tempo, ciò che Floridi definisce appunto iperstoria.

Sul concetto di iperstoria, che in queste pagine osservo come una periodizzazione cronosofica, non si è ancora sviluppato un dibattito accademico e scientifico tale da poter avvalorare o confutare le astrazioni di Luciano Floridi. Userò dunque questo concetto in senso funzionale, un grimaldello con il quale provare a portare il dibattito legato alla rivoluzione digitale all'interno della comunità scientifica delle storiche e degli storici.

Secondo Floridi, l'essere umano fa il suo ingresso all'interno della storia grazie alle tecnologie, nel senso che è grazie a diversi strumenti tecnologici, che l'uomo nel suo evolversi è riuscito a produrre e lasciare tracce che storiche e storici oggi possono analizzare. È indubbio che la pervasività delle tecnologie digitali unita alla quantità dell'informazione prodotta, scambiata e archiviata ci conduce verso un tempo nuovo, del quale forse ancora dobbiamo comprendere la vera essenza, ma che senza ombra di dubbio ci sta mostrando le possibilità di un radicale cambiamento. Non di meno è necessario chiedersi quale sia il valore epistemologico che in quanto storici possiamo assegnare al concetto di dato. Le società iperstoriche, per come definite da Floridi, divengo tali anche nel momento in cui le tecnologie di terzo ordine, quelle in cui è una tecnologia a mediare tra altre tecnologie, assumono un ruolo preponderante nella produzione e condivisione di dati. Saranno quei dati di interesse per lo storico del futuro? E se sì, come potrà utilizzarli all'interno del proprio percorso di ricerca?

È importante che la storia come disciplina continui ad interrogarsi in modo complesso non solo sugli strumenti ma sui cambiamenti di natura epistemologica che questo passaggio epocale ci consegna, altrimenti c'è il rischio che la storia, intesa come le tracce oggi presenti negli archivi in

forma di documenti materici, non riesca ad entrare in questo nuovo tempo o che lo faccia in un modo che in futuro non saremo in grado di controllare.
Prima però di approfondire in che modo è possibile far sì che la storia non rimanga fuori dall'iperstoria è bene provare a capire cosa si intenda con questo termine, in prima istanza chiedendosi se l'iperstoria abbia a che fare con il tempo e se si in che termini.
Indagare l'idea, il concetto e il significato di tempo va sicuramente al di là delle mie competenze, innegabile però la fascinazione verso un qualcosa del quale ancora non si ha una precisa definizione, di cui riconosciamo solo la relatività, e forse proprio per questo continuiamo a cercare modi per osservarlo, renderlo contabile e comprensibile, col sospetto che non potremmo mai afferrarlo.

1.4 Iperstoria, un concetto cronosofico

Da storico credo che possa essere interessante chiedersi quali modi ci si sia dati per provare non tanto a definire il concetto di tempo, ma per calcolarlo anche in una prospettiva storica.
Ricoeur, Le Goff, Koselleck e Pomian ci aiutano in questa riflessione facendoci presente che sono stati molti e diversi i modelli concettualizzati dall'antico testamento in poi per provare a comprendere il tempo.
Nella tradizione ebraico cristiana sono stati sviluppati due modelli di periodizzazione, ognuno dei quali utilizza due cifre simboliche: il 4 riferito alle stagioni dell'anno e il 6 relativo alle età della vita. I due modelli sono proposti da Daniele nell'antico testamento e da Sant'Agostino nel IX libro della *Città di Dio.*[58]
Nella loro suddivisone del tempo, sia Daniele sia Agostino si ispirano ai cicli della natura. I quattro regni del Profeta corrispondono alle quattro stagioni, mentre i sei periodi descritti da Agostino rinviano da una parte ai sei giorni della creazione, dall'altra alle sei età della vita: infanzia, fanciullezza, adolescenza, gioventù, maturità e vecchiaia. Entrambi però attribuiscono alle proprie periodizzazioni un significato simbolico.
Per Le Goff «nella concezione del lontano passato, i periodi non possono essere sequenze neutre, bensì esprimono sentimenti diversi riguardo al tempo e a quella che, in una plurisecolare elaborazione, si chiamerà la «storia»»[59].
La teoria agostiniana delle sei età consegna l'idea di avanzamento del tempo all'inevitabile senescenza, dunque ad un progressivo decadimento analogo a quello della vita umana, una visione

58 J. Le Goff, *Il tempo continuo della storia,* La Terza Bari Milano, 2022, pp. 10,11
59 Ivi, p. 12

pessimistica dell'avanzare del tempo che ha contribuito in modo sensibile all'impedimento fino al XVIII secolo dello svilupparsi di un'idea di progresso, seguendo la quale, più si avanza nel tempo e più le società, dunque la vita, progrediscono in senso positivo.
Sarà con la prima lezione tenuta da Michelet al College de France, che per la prima volta in modo netto si definisce un'epoca come migliore di quella precedente. L'uscita da ciò che ormai a gli occhi dello storico francese è un medioevo cupo, ci consegna alla libertà e al progresso delle cose umane date dal Rinascimento.
In questo processo la periodizzazione assume particolare valore per un'analisi di tipo qualitativo, si individua un periodo che riteniamo sia migliore di quello precedente e per questo ne riconosciamo dei valori culturali, estetici e politici che lo caratterizzano in modo così deciso da renderlo unico e dunque periodizzante, una pietra miliare nella strada del tempo alla quale attribuire valori specifici così da definire un prima e un dopo.
L'immagine del Rinascimento brilla particolarmente perché narrativamente contrapposta al periodo a lui precedente, un Medioevo che però fino al 1839, anno della morte della moglie di Michelet, lo stesso storico francese descriveva come un periodo festoso, luminoso, vitale, esuberante.
Le periodizzazioni sono frutto di una volontà narrativa dello storico che le studia, un utile strumento per provare a contare e dividere il tempo del passato, ma ancora una volta frutto di un processo che non si può definire scientifico né empirico.
Ed è esattamente in questo ambito che ho intenzione di inserire il concetto di Iperstoria, come possibile periodizzazione del tempo che stiamo vivendo.

La periodizzazione della storia non è mai un atto neutro o innocente: lo testimonia l'evoluzione dell'immagine del medioevo in epoca moderna e contemporanea. Attraverso di essa si esprime un apprezzamento delle sequenze di epoche così definite, un giudizio di valore, per quanto collettivo. D'altronde l'immagine di un periodo storico può modificarsi con il tempo.
Ogni periodizzazione opera dell'uomo, è insieme artificiale e provvisoria, e si evolve insieme alla storia stessa. Da questo punto di vista essa riveste una duplice utilità, in quanto permette sì di padroneggiare al meglio il tempo passato, ma evidenzia anche la fragilità di quel particolare strumento del sapere umano che è la storia.[60]

Pomian, in *Lordre du Temps*[61], prova a descrivere i processi attraverso i quali siamo arrivati a descrivere il tempo, partendo da costruzioni analitiche spinte da una volontà empirica fino ad

60 Ivi, p. 23
61 K. Pomian, *L'ordre du temps*, Gallimard, Paris 1984

arrivare a processi di identificazione del tempo dati da una significazione dei fatti sviluppata dallo storico stesso.

Pomian come scrive Ricoeur parla di un tempo cronologico, composto da cicli brevi o lunghi che ritornano, come i giorni, le settimane, i mesi e gli anni e un tempo cronometrico che designa il tempo lineare dei periodi lunghi, come i secoli o i millenni la cui scansione è punteggiata da eventi fondamentali e fondatori. Questa è la rappresentazione del tempo che secondo Ricoeur più si avvicina al modo storico. Si ordinano gli avvenimenti in funzione di una serie di date e di nomi, mettendo in sequenza le ere e le loro suddivisioni.[62]

Vi è poi la cronografia, un sistema di notazione del tempo che può fare a meno del calendario, dove gli episodi registrati vengono definiti dalla loro posizione rispetto agli altri: successione di eventi unici.

E infine la cronosofia, che secondo Ricoeur consente di «elaborare una "storia della storia" da cui gli storici di mestiere non riescono forse mai a liberarsi, dal momento che si tratta di assegnare una significazione ai fatti: continuità vs discontinuità, ciclo vs linearità, distinzione in periodi o in ere»[63].

Le riflessioni sulla cronosofia torneranno di grande utilità, insieme alla storia dei concetti, quando ci interrogheremo sulla possibilità di utilizzare i metadati in senso concettuale per consentire una accessibilità e un'interoperabilità delle fonti, trasformate in oggetti informativi, utilizzando concetti trasversali che aiutino i fruitori a trovare una connessione tra il passato e il presente, così da poter meglio comprendere la storia.

Il pensiero di Koselleck rispetto alle temporalità storiche può essere utile a comprendere come poter osservare il concetto di iperstoria, partendo dal presupposto che una società e i suoi concetti stanno in un continuo rapporto di tensione, dove specifici processi sociali, politici e culturali vengono tematizzati dai concetti ancora prima di poter essere esperiti e compresi dalla società che li genera.

Il concetto di iperstoria appartenendo al nostro tempo, va analizzato in modo sincronico, non di meno però, per poterlo comprendere all'interno del concetto di tempo, così dove lo posiziona lo stesso Floridi, dobbiamo porci delle domande che normalmente trovano una soluzione all'interno di una analisi diacronica, dal momento che nella storia di un concetto vengono misurati lo spazio di esperienza e l'orizzonte di aspettativa del concetto e del tempo in questione.

Per provare a decifrare l'idea di iperstoria è indubbio che ci si debba confrontare con la questione cronologica legata ad un'idea di progresso, ma anche con una definizione cronosofica, in cui una periodizzazione può assumere un significato valoriale e qualitativo.

62 P. Ricoeur, *La memoria, la storia, l'oblio*, Raffaello Cortina Editore, 2003 pp. 218, 219
63 Ivi, p. 219

Noi dobbiamo in primo luogo articolare la differenza tra le categorie temporali naturali e le categorie storiche. Ci sono intervalli temporali per esempio che durano finché non è decisa una battaglia, ossia periodi intersogettivi di attività in corso, durante i quali il tempo naturale è restato per così dire escluso. Ovviamente eventi e situazioni restano riferibili alla cronologia naturale, nella quale è persino insito un presupposto minimale della loro interpretazione. Il tempo naturale e il suo decorso (comunque esperito) appartengono alle condizioni dei tempi storici; ma questi ultimi non si risolvono mai nel primo. I tempi storici hanno serie temporali diverse dai ritmi temporali imposti dalla natura. […] Per noi oggi sembra ovvio che lo spazio dell'azione politica e sociale si sia fortemente staccato dalla natura attraverso la coazione sistematica della tecnica.[64]

Per Koselleck i tempi storici possono dunque essere definiti anche in un ordine temporale non naturale, ma definito dalla tecnica, dunque da un'azione diretta dell'uomo che modifica il proprio ambiente e la propria società in modo così determinante da creare un nuovo tempo incastonato tra un prima e un dopo, dove il prima si denota con un senso di mancanza, mentre il dopo si ammanta di un valore evolutivo. Evoluzione e rivoluzione sono due termini che è possibile mettere in relazione, secondo Koselleck vi può essere la libertà di scambiare i due concetti se osservati non tanto dagli avvenimenti che li compongono nei casi specifici, ma nella portata sociale e collettiva dei propri effetti, dove è possibile osservare un processo generale di emancipazione sociale che si diffonde e viene promosso con l'industrializzazione e nel nostro caso specifico con la digitalizzazione.

Osservare la storia attraverso i concetti ci pone davanti al tempo sempre in modo sia diacronico che sincronico il che porta passato presente e futuro a trovare una compenetrazione nel nostro pensare. Il passato si è manifestato in se stesso per poi essere colto grazie alle tracce del suo passaggio, tracce che carichiamo di significato in base ad un metodo analitico di comparazione e critica, ma che inevitabilmente finiamo per concettualizzare tendendo nello stesso piano del discorso il significato dei concetti per come si sono espressi nel passato e il significato dei concetti per come possiamo comprenderli oggi, e attraverso questo tendiamo a costruire dei significati generali che ci aiutano ad orientarci temporalmente nel passato e che contemporaneamente ci spingono verso il futuro con una prospettiva evolutiva.

E' possibile definire l'iperstoria come concetto in relazione ai tempi storici solo se ne accettiamo il valore cronosofico e se decidiamo di osservare il passato in base a concetti che solo nel nostro contemporaneo siamo in grado di caricare di un significato così rilevante. Dobbiamo però essere

64 R. Koselleck, *Futuro passato. Per una semantica dei tempi storici*, Clueb, Bologna 2007, p. 113

consapevoli che le nostre astrazioni semantiche se possono essere caricate di verità nello svolgersi della nostra società, sono solo ipotesi se proiettate nel futuro.

In presenza di un'identità cronologica naturale si possono avere differenti classificazioni di corsi storici. In questa rifrazione temporale sono contenuti strati temporali diversi, che hanno una durata differente a seconda dei soggetti dell'azione o delle situazioni esaminate, e che dovrebbero essere misurati gli uni rispetto a gli altri. Nel concetto della contemporaneità del non-contemporaneo sono anche contenute diverse estensioni temporali. Rinviano alla struttura prognostica del tempo storico, poiché ogni prognosi anticipa eventi che sono si impliciti nel presente, e in questo senso esistono già, ma che non sono ancora avvenuti.[65]

Tempo e significazione si intrecciano e vengono determinati da natura e tecnica. Questo incrocio di concetti credo ci aiuti a comprendere come il tempo sia relativo non solo in termini fisici, ma anche semantici e per quanto ci possa sembrare ovvio collegare il progresso ad un concetto lineare di sviluppo Pomian, così come Koselleck, ci mette in guardia su questa semplificazione.

Lo sviluppo delle conoscenze preistoriche e archeologiche tende a disporre nello spazio forme di civiltà che eravamo propensi a immaginare come successive nel tempo. Il che significa due cose: anzitutto che il progresso (se questo termine è ancora adatto a designare una realtà diversissima da quella a cui era stato in un primo tempo applicato) non è necessario ne continuo; procede a salti, a balzi, o, come direbbero i biologi, per mutazioni. Tali salti e tali balzi non consistono nell'andare sempre più lontano nella stessa direzione; si accompagnano a mutamenti di orientamento, un po' alla maniera del cavallo negli scacchi che ha sempre a sua disposizione svariate progressioni ma mai nello stesso senso. L'umanità in progresso non assomiglia certo a un personaggio che sale una scala, che raggiunge con ogni suo movimento un nuovo gradino a tutti quelli già conquistati; evoca semmai il giocatore la cui fortuna è suddivisa su parecchi dadi e che ogni volta che li getta, vede sparpagliarsi sul tappeto, dando luogo via via a computi diversi. Quel che si guadagna sull'uno, si è sempre disposti a perderlo sull'altro, e solo di tanto in tanto la storia è cumulativa, cioè i computi si addizionano in modo da formare una combinazione favorevole[66]-

L'iperstoria se connessa all'idea di tempo è una visione dell'oggi nel futuro, comprensibile se comparata al passato, mentre assume un significato nel contemporaneo solo se il concetto è messo in relazione con l' esperienza. Un'esperienza che però si verifica solo per quegli esseri umani il cui ambiente è modificato dalla tecnologia, una tecnologia che consente un costante scambio informazionale tra persone e sistemi ICT o come abbiamo visto anche solo tra tecnologie ICT.

65 Ivi, p. 112
66 K. Pomian, *L'ordine del tempo*, Einaudi, Torino 1992 p. 162

L’iperstoria è un tempo che si estende globalmente, ma che non coinvolge in egual modo tutta la popolazione mondiale, accogliendo al suo interno società che non sono descritte in ambiti geografici, ma tecnologici. Ciò che appunto Floridi ha definito infosfera.
«A un livello massimo l’infosfera è un concetto che può essere utilizzato come sinonimo di realtà, laddove interpretiamo quest’ultima in termini informazionali. In tal caso l’idea è che ciò che è reale è informazionale e ciò che è informazionale è reale.»[67]

1.5 Tecnologia, informazione, società

Nell'idea di iperstoria l’ambiente è determinante, per come si sono evolute le tecnologie all’interno della civiltà umana, non vi può essere iperstoria al di fuori e senza le tecnologie digitali, la rete e le ICT.
Il tempo e lo spazio ancora una volta esistono in un rapporto simbiotico, anche se in questo caso il concetto di spazio non ha un valore misurabile in base ad una estensione geometrica. Lo spazio qui è definito dall’ambiente che a sua volta è determinato dalla società, una società, quella iperstorica che non ha confini specifici, si apre in senso globale ma ad oggi non coinvolge tutto il mondo, anche se è molto probabile che ogni essere umano a breve potrà vivere all’interno di una società iperstorica. Per ora vi sono delle società ancora non in grado di sviluppare delle relazioni informazionali uomo digitale così evolute da aumentare il grado di produzione e archiviazione delle informazioni.
Ancora una volta ci troviamo difronte alla questione dell’informazione, che nell’idea di tempo che esprime Floridi assume un valore determinante. E’ infatti l’informazione secondo il filosofo a determinare l’ingresso di una società all’interno della storia.

> Solo quando sono diventati disponibili sistemi per registrare eventi e, in tal modo, accumulare e trasmettere le informazioni per un futuro consumo, le lezioni apprese dalle generazioni passate hanno iniziato ad evolvere esponenzialmente, in maniera debole o lamarkiana e così l’umanità ha fatto ingresso nella storia.[68]

Secondo il filosofo le società si dividono in preistoriche, storiche e iperstoriche in base alla loro capacità di produrre, scambiare e archiviare informazioni, dove l’interazione con le tecnologie ICT è determinante. Ad oggi la maggior parte delle società sono ancora storiche, dove le ICT non hanno ancora preso il sopravento su le altre tecnologie, ma vi sono parti del mondo dove ancora esistono

67 Luciano Floridi, *la quarta rivoluzione. Come l’infosfera sta trasformando il mondo,* Raffaello Cortina editorie, Milano, 2017 p. 47
68 Ivi, p. 1

società preistoriche, dove ancora si vive senza registrare informazioni, ne son un esempio società isolate nella regione amazzonica.
Contemporaneamente a società preistoriche ne esistono di iperstoriche, dove le ICT e le loro capacità di processare dati sono condizioni essenziali per assicurare e promuovere il benessere sociale, la crescita individuale e lo sviluppo generale. Ne son un esempio tutti i membri del G7.

In ciascuno di questi paesi almeno il 70% del PIL dipende da beni intangibili, fondati sull'uso di informazioni piuttosto che da beni meteriali prodotto di processi agricoli o manufatturieri. Le loro economie riposano in larga misura su risorse basate sull'informazione (economia della conoscenza), servizi ad alta intensità di informazione (in particolare nell'ambito di servizi commerciali, proprietà, comunicazione, finanza, assicurazione e intrattenimento) e settori pubblici orientati all'informazione (in particolare l'educazione, la pubblica amministrazione e la sanità)[69]

Se poniamo l'informazione al centro dei processi storici e evolutivi della società umana ci troviamo davanti a tre stadi, la preistoria in cui non ci sono ICT; la storia, in cui ci sono ICT che registrano e trasmettono informazioni ma le società umane dipendono principalmente da altre tecnologie che riguardano le risorse primarie e l'energia e l'iperstoria, in cui ci sono ICT che registrano, trasmettono e processano informazioni in modo sempre più autonomo in cui le società umane dipendono in modo cruciale dalle ICT e dall'informazione in quanto risorsa essenziale per la loro crescita.
All'interno di questo sistema così complesso, è necessario chiedersi se e come la storia intesa come disciplina e la storiografia, subiscano un processo di mutazione metodologica e epistemica.
I prossimi capitoli saranno interamente incentrati sul provare a proporre una risposta a questa domanda, ma prima di addentrarci in queste riflessioni penso sia necessario chiedersi cosa voglia dire Floridi quando afferma che: «storia è sinonimo di informazione.»[70]
La storia come disciplina esiste a prescindere dalle tecnologie siano queste di primo, secondo o terzo ordine? La risposta è no.
Come scriveva Bloch la storia è la conoscenza dell'uomo in società[71]. Dunque lo storico è chiamato a ricercare le tracce che gli uomini hanno lasciato nel loro esistere all'interno di una società, così da poter comprendere e raccontare avvenimenti, pensieri e azioni che in modo più o meno marcato hanno definito il nostro passato, dunque il nostro presente.

69 Ivi, p. 4
70 Ivi, p. 3
71 M. Bloch, A*pologia della storia. O mestiere di storico,* Giulio Einaudi Editore, Torino 1998 p. XLV.

Le fonti infatti altro non sono che informazioni che lo storico è in grado di trovare attraverso uno scambio informazionale tra lui e un archivio e tra lui e un'altra persona.

Come avremo modo di approfondire nei prossimi capitoli quando parlo di archivi alludo non solo ad archivi fisici o istituzionali, ma anche ad archivi digitali pubblici e privati, archivi inventati, archivi di persona, di audiovisivi, musei, pinacoteche, emeroteche, fonoteche, YouTube, Google Art and Culture, ma anche Facebook, Instagram, Twitter, tutti "luoghi" di deposito e conservazione di fonti, dove queste possono essere documenti ufficiali, diari, autobiografie, musica, arte, tweet, post, video, canzoni, quadri, ecc. tutto ciò che può essere definito come fonte primaria, ossia una fonte che si è formata all'interno dello stesso tempo che lo storico vuole indagare per la propria ricerca.

Tutte queste tipologie di fonti, unite alle fonti orali, sono l'unico modo che si ha per provare ad avere una rappresentazione abbastanza eterogenea delle azioni dell'uomo nel passato e sono l'unico strumento a disposizione dello storico per procedere in modo scientifico nella sua ricerca. Vedremo come non vi è storico in vero che utilizza solo le fonti per produrre discorso storico, infatti ogni azione di rappresentazione storiografica passa anche attraverso un processo di spiegazione e narrazione. Non solo, secondo Topolski le fonti dirette possono parlare solo a se stesse, poiché le si osserva in un tempo diverso da quello in cui son state create[72].

Ognuna delle tipologie di fonti citate, tranne le fonti orali, e molte sono rimaste fuori da questo breve elenco, esiste grazie allo sviluppo di una specifica tecnologia, a partire dalle tavolette sumeriche dove grazie a tecnologie di primo e secondo ordine è stato possibile incidere con un cuneo su una superficie di argilla.

La produzione di fonti è dunque direttamente legata alle tecnologie che di fatto rendono possibile la realizzazione e la conservazione delle fonti.

Con i processi di digitalizzazione della società ci troveremo davanti a tre tipologie di fonti, le fonti materiche, le fonti digitalizzate e le fonti digital born. Ognuna di queste ha una sua peculiarità epistemica sia nella produzione che nella conservazione, ma solo le fonti digitalizzate e le fonti digital born possono far parte di società iperstoriche dove l'informazione è composta da dati e governata da sistemi ICT.

In informatica «i dati sono delle rappresentazioni originarie, cioè non interpretate, di un fenomeno, evento, o fatto, effettuate attraverso simboli o combinazioni di simboli, o di qualsiasi altra forma espressiva legate a un qualsiasi supporto »[73].

72 J. Topolski, *Narrare la storia. Nuovi principi di metodologia storica*, Bruno Mondadori, Milano, 1997 p. 54

73 G.M. Di Nunzio, *Sistemi di elaborazione delle informazioni*, Università degli studi di Padova, 2014, http://www.dei.unipd.it/~dinunzio/fdi-2014-2015/04_dati_informazioni.pdf ultima consultazione 30/05/2020

Ogni oggetto digitale è composto di dati, enumerabili in bit e si trasformano in informazione nel momento in cui al sistema in utilizzo per conservare o rappresentare i dati, vengono fornite indicazioni tali da generare un contesto capace di produrre informazione.

Nelle prossime pagine cercheremo di approfondire la questione dell'informazione e di come questa sia in realtà secondo la filosofia di Ferraris, legata al processo dell'iscrizione, dunque alla scrittura.

Vedremo anche come e quanti dati vengono generati all'interno dell'era digitale e di come sia complesso gestirli e archiviarli. Il tema della memoria, intesa come conservazione di medio e lungo periodo delle informazioni, sarà un tema fondamentale da comprendere, perché se è vero che i dati sono più semplici da archiviare rispetto a fonti materiche è altresì vero che ne generiamo talmente tanti che già oggi pare impossibile poterli conservare tutti e anche ci riuscissimo, i sistemi tecnologici evolvono così rapidamente che l'obsolescenza tecnologica è un problema tangibile che ci porta ad interrogarci su come poter assicurare la permanenza nel tempo di un dato archiviato.

Ora sappiamo di essere all'interno di una rivoluzione che coinvolge l'ambiente e l'informazione, una rivoluzione dove la tecnologia pervade ogni esperienza e l'essere umano non è più al centro dei processi informativi.

Floridi la definisce la quarta rivoluzione.

La prima è data dalle scoperte di Copernico, che nel 1543, con *Sulle rivoluzioni dei corpi celesti* estromette l'uomo dalla sua posizione di centro dell'universo.

La seconda rivoluzione arriva tramite un altro libro, è con *L'origine della specie* che Darwin pone il tema dell'evoluzione estromettendo l'essere umano dal centro del regno biologico, costringendoci ancora una volta a dover riosservare la nostra posizione in relazione con l'esterno e con noi stessi.

Con Des Cartes e con il suo celebre "penso dunque sono" l'essere umano trovava essenza del suo esistere nella sua mente, dove la nostra centralità risiedeva nella coscienza, avevamo pieni poteri sulla nostra mente, convinti che tramite l'introspezione bastasse guardarci dentro per conoscere il contenuto e averne pieno controllo. Con gli studi sulla psicanalisi di Freud e oggi con le neuroscienze si da forma alla terza rivoluzione. L'uomo scopre l'inconscio, comprende dunque che sotto la coscienza esiste altro e che non ci è possibile accedere a tutti i contenuti della nostra mente.

«Siamo stati spodestati dal centro del reame della coscienza pura e trasparente, e riconosciamo di essere opachi a noi stessi»[74].

Dopo queste tre rivoluzioni l'essere umano di sé sa che non è più immobile al centro dell'universo, non è un'anomalia all'interno del mondo animale e che la comprensione di sé non è interamente trasparente.

74 L. Floridi, *la quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Raffaello Cortina editorie, Milano, 2017 p. 102

Oggi per Floridi stiamo vivendo una quarta rivoluzione, dove a breve scomparirà la divisione tra online e offline, per proiettarci all'interno dell'onlife dove non vi è più separazione tra essere o non essere connessi alla rete poiché le due cose non saranno più divisibili, tranne che attraverso un processo di eremitaggio e di rifiuto estremo delle tecnologie, ognuno di noi vivrà all'interno dell'infosfera, ma non al centro.

Stiamo lentamente accettando l'idea che si fa strada a partire da Turing, per cui non siamo agenti newtoniani isolati e unici, come una sorta di Robinson Crosue su un'isola. Piuttosto siamo organismi informazionali (inforg) reciprocamente connessi e parte di un ambiente informazionale (l'infosfera), che condividiamo con altri agenti informazionali naturali e artificiali, che processano informazione in modo logico e autonomo.

L'orco di Bloch, per cercare carne umana deve dotarsi di nuovi sistemi computazionali.

1.6 Big data una sfida quantitativa o qualitativa?

Information is not knowledge,
Knowledge is not wisdom,
Wisdom is not truth,
Truth is not beauty,
Beauty is not love,
Love is not music.
Music is the best[75]

L'innovazione metodologica che ha permesso la rivoluzione digitale nella storiografia si manifesta sopratutto nella ricerca documentaria, nel nuovo rapporto con le fonti primarie digitali, la loro configurazione e morfologia, nelle strategie di reperimento delle informazioni e nelle modalità di accesso alle fonti primarie e secondarie e in fine negli strumenti che possediamo per analizzarle e per conservarle. […] È la dimensione euristica che delinea lo spartiacque tra chi usa alcuni strumenti del digitale e chi, invece, si proietta nell'analisi computazionale della documentazione digitale.[76]

Queste parole di Serge Noiret ci aiutano a comprendere e ridefinire il terreno su cui si sta provando a costruire un dibattito di natura metodologica, che aiuti gli storici e le storiche ad individuare i giusti percorsi per confrontarsi con la rivoluzione digitale che sta attraversando e modificando la nostra società.
Quello che abbiamo visto anche nelle pagine precedenti è quanto i processi di digitalizzazione abbiano modificato il nostro modo di produrre fonti, fonti che noi iniziamo a chiamare dati. Sempre Noiret quando parla di storia digitale parla di una storia «dominata dalla presenza di dati digitali, *data driven*, termine che ha soppiantato l'utilizzo del sostantivo fonte.»[77]
Il nostro essere inforganismi in una costante relazione informazionale con sistemi ICT modifica in senso ontologico le modalità che oggi si hanno di produrre e lasciare tracce delle nostre azioni e dei nostri pensieri, ciò che muta è sia la forma che la sostanza, e questo è indubbiamente vero quando ci si riferisce alle fonti digital born, prodotte dunque attraverso strumenti digitali, ma non di meno

75 Testo tratto dal brano di Frank Zappa, *Packard goose*, dall'album *Joe's Garage,* Zappa Records RykoDisk, 1979
76 S. Noiret, *Storia contemporanea digitale*, in R. Minuti, a cura di, AA.VV, *Il web e gli studi storici. Guida critica all'uso della rete*, Carocci Editori, 2017 Roma p. 273
77 Ivi, p. 271

interessa anche le fonti materiche che una volta dematerializzate e digitalizzate subiscono un processo di rimediazione[78].

Nel prossimo capitolo ci concentreremo in modo particolare sulla rimediazione di fonti materiche, adesso però credo sia necessario soffermarci ad analizzare criticamente le fonti digital born, consapevoli però che ancora non possiamo avere gli strumenti adatti per comporre un'analisi critica da un punto di vista storiografico di questa tipologia di fonti.

Sono ancora troppe le domande a cui non è possibile per uno storico dare una risposta a riguardo, ciò che si può fare è provare ad avanzare delle ipotesi, consapevoli però che non ci è possibile dare definizioni strutturate e immutabili.

Riguardo alle fonti digital born, i principali problemi con i quali filosofi, sociologi, informatici, digital humanists e scienziati della comunicazione si stanno confrontando riguardano il tema dei big data, la proprietà dei dati e la conservazione di questi su piattaforme proprietarie. In poche parole noi ad oggi non abbiamo ancora risposte definitive e univoche su come sia possibile organizzare i big data, a chi appartengano i dati e se una piattaforma come Facebook, solo per fare un esempio, tra dieci anni consentirà ancora l'accesso ai dati prodotti oggi dai sui utenti e che risiedono all'interno dei suoi server.

Un ragionamento a parte va fatto per i Linked Open Data[79], ossia dataset aperti di entità pubbliche e governative[80]. In Italia, il Codice dell'Amministrazione Digitale[81] ha introdotto i principi della disponibilità dei dati pubblici, è dunque possibile accedere ai dati senza restrizione da parte di soggetti pubblici o privati. Il CAD è un testo unico che riunisce e organizza le norme riguardanti l'informatizzazione della Pubblica Amministrazione nei rapporti con i cittadini e le imprese.

I dati della PA saranno fonti fondamentali per lo storico del futuro e nel prossimo capitolo affronteremo in modo più approfondito il tema del Linked Open Data, per ora però non possiamo che porci il problema di come poter gestire in senso diacronico tutte le altre fonti nate digitali. Dobbiamo essere consapevoli dell'indeterminatezza che oggi viviamo e che consente appunto di produrre solo ipotesi e proposte, interessanti per sviluppare un dibattito ma troppo deboli per poter essere inserite in un ragionamento di metodo.

Ciò non di meno è assolutamente importante comprendere ciò di cui stiamo parlando, dal momento che lo storico del futuro, per osservare il nostro presente, non potrà che confrontarsi con fonti prodotte in ambito digitale.

---

78 J.D Bolter-R. Grusin, *Rimediation. Competizione e integrazione tra media vecchi e nuovi*, Guerini, Milano 2002 p73

79 http://dati.camera.it/it/linked-data/

80 T. Di Noia, R. De Virgilio, E. Sciascio, F.M. Donini, *Semantic Web. Tra ontologia e open data*, Maggioli Editori, Santarcangelo di Romagna 2018 p., p. 83

81 https://www.agid.gov.it/it/agenzia/strategia-quadro-normativo/codice-amministrazione-digitale/

Per provare a comprendere il fenomeno credo possa essere utile, per quanto non semplice, riportare alcuni dati[82] di ordine quantitativo, che ci diano una rappresentazione il più possibile veritiera della quantità di dati che oggi produciamo.

Per avere un'idea della complessità del cambiamento che stiamo provando ad osservare, credo possa essere utile partire da un articolo di «Science» dell'aprile del 2011, dunque non un articolo recente e per questo portatore di un dato assolutamente non più corrispondente al reale attuale, ma utile per darsi un metro di paragone che ci possa far comprendere come ciò che oggi osserviamo potrà decuplicarsi nei prossimi anni.

L'articolo di Martin Hilbert e Priscila Lopez, dal titolo *The World's Technological Capacity to Store, Communicate, and Compute Information*, sostiene che alla fine degli anni ottanta, meno dell'1% dell'informazione mondiale tecnologicamente archiviata fosse in un formato digitale, mentre lo era in una percentuale del 94% nel 2007[83], bisogna considerare che il 2002 è l'anno in cui l'umanità fu capace di immagazzinare una maggiore quantità di informazione in una forma digitale[84], piuttosto che analogica, è da qui che alcuni datano l'inizio dell'Era Digitale.

Dal 1989 al 2015 la capacità mondiale di scambiare informazioni attraverso le reti di telecomunicazione è quasi raddoppiata ogni 40 mesi, il termine internet 2.0, quando internet dunque diviene il luogo dell'autorialità diffusa, viene ufficializzato per la prima volta nel 2004 in occasione del *The web 2.0 conference* all'hotel Nikko di San francisco[85].

Nel 2015 si è stimato che la capacità mondiale di immagazzinare dati è aumentata da 2,6 exabytes nel 1986, a circa 5.000 exabites[86] nel 2014. Un testo di circa 300 pagine con all'interno delle immagini, esportato in pdf senza una compressione specifica, pesa circa 10 MB, un mega sono 1.048.576 byte, un exabites sono 1.152.921.504.606.846.976 byte.

I ricercatori della School of Information di Berkeley nel 2004 stimarono che l'umanità avesse accumulato in una traduzione digitale, approssimativamente 12 exabyte di dati nel corso della storia, fino alla diffusione dei computer, e a partire da questa, 180 exabyte già entro il 2006. Secondo uno studio del 2011 il totale è cresciuto fino a più di 1600 exabyte tra il 2006 e il 2011, oltrepassando in tal modo la soglia dello zettabyte (1 zettabyte sono 1.000 exabyte). Questo numero

82 Il termine dato, come vedremo, può creare della confusione semantica. In questo specifico caso mi riferisco ad analisi quantitative di dati che a loro volta generano dati, ma che a differenza di quelli analizzati, i dati che citerò sono già inseriti in un contesto di significato, presentandosi dunque come informazioni. Sul rapporto tra dato, informazione e conoscenza torneremo più avanti in questo capitolo.

83 *Science* 01 Apr 2011:Vol. 332, Issue 6025, pp. 60-65 DOI: 10.1126/science.1200970

84 M. Hilbert, *Video animation on The World's Technological Capacity to Store, Communicate, and Compute Information from 1986 to 2010*

85 https://web.archive.org/web/20080913174125/http://www.web2con.com/web2con/. É interessante notare che l'URL del sito si trovi all'interno del portale *Web.Archive*, un arhivio che raccoglie vecchi siti.

86 Exybites è una unità di misura informatica, un exabytes è la sesta potenza di mille e vale un milione di terabite, un miliardo di miliardi di bite. 1000000000000000000 byte = 1000^6 = 10^{18} byte = 1 trilione di byte o un milione di terabyte

tende a crescere di quattro volte circa ogni tre anni, cosicché si sono raggiunti gli 8 zettabyte di dati entro il 2015.

«Ogni giorno viene generato un numero sufficiente di dati da riempire tutte le biblioteche americane più di otto volte.»[87]

Con lo sviluppo dell'infosfera e delle tecnologie ICT, la mole di dati prodotta è molto superiore a quanto gli analisti si fossero immaginati, oltre a ciò con la pandemia Covid -19 e il sistema di lockdown globale che ha di fatto chiuso in casa miliardi di persone contemporaneamente, si è creata una accelerazione spaventosa di dati digitali prodotti, consumati e archiviati.

L'editoria accademica e scientifica purtroppo è lenta nel pubblicare i risultati della ricerca, una lentezza che mal si concilia con la velocità evolutiva di cui è capace l'era digitale, volendo trovare informazioni attendibili sulla quantità di dati prodotti tra il 2019 e il 2020, tenendo anche conto delle implicazioni dovute alla pandemia di Covid-19, ho dovuto portare avanti la ricerca comparando tra di loro oltre che articoli di riviste di settore, i dati pubblicati da società di gestione dati, produttrici di software, società di statistica, di data analisi finanziarie e di marketing digitale. I dati digitali sono ad oggi uno degli asset finanziari maggiormente in crescita, saper raccogliere, osservare e clusterizzare dati al meglio vuol dire essere in grado di comprendere trend economici, sociali e culturali, acquisire così conoscenza da poter utilizzare o rivendere ad altre società.

Il potere computazionale di società che si occupano di gestione dati le rende *defacto* capaci di gestire quella che oggi è considerata tra le risorse più importanti e economicamente determinanti del nostro contemporaneo.

Secondo Arne Holst[88], esperto di tecnologia e telecomunicazioni, la quantità totale di dati creati, archiviati, copiati e consumati a livello globale ha già raggiunto i 64,2 zettabyte nel 2020 ed è destinata ad aumentare rapidamente. Credo sia utile ricordarsi che nel 2011 avevamo superato il primo zettabyte. Fino al 2025, si prevede che la creazione globale di dati crescerà fino a più di 180 zettabyte. Nel 2020, la quantità di dati creati e condivisi ha raggiunto un nuovo massimo, la crescita è stata più alta di quanto previsto in precedenza a causa dall'aumento della domanda in seguito alla pandemia COVID-19, lo smart working, le video conferenze, così come la didattica a distanza hanno apportato un aumento veramente notevole di dati prodotti e fruiti, così come la necessità di un intrattenimento domestico ha reso le piattaforme di streaming molto più diffuse.

È importante notare però che solo una piccola percentuale di questi nuovi dati creati viene conservata, solo il due per cento dei dati prodotti e consumati nel 2020 è stato salvato e conservato nel 2021.

87 L. Floridi, *la quarta rivoluzione. Come l'infosfera sta trasformando il mondo,* Raffaello Cortina editorie, Milano, 2017 p. 13

88 https://www.statista.com/statistics/871513/worldwide-data-created/ ultima consultazione 15/06/2021

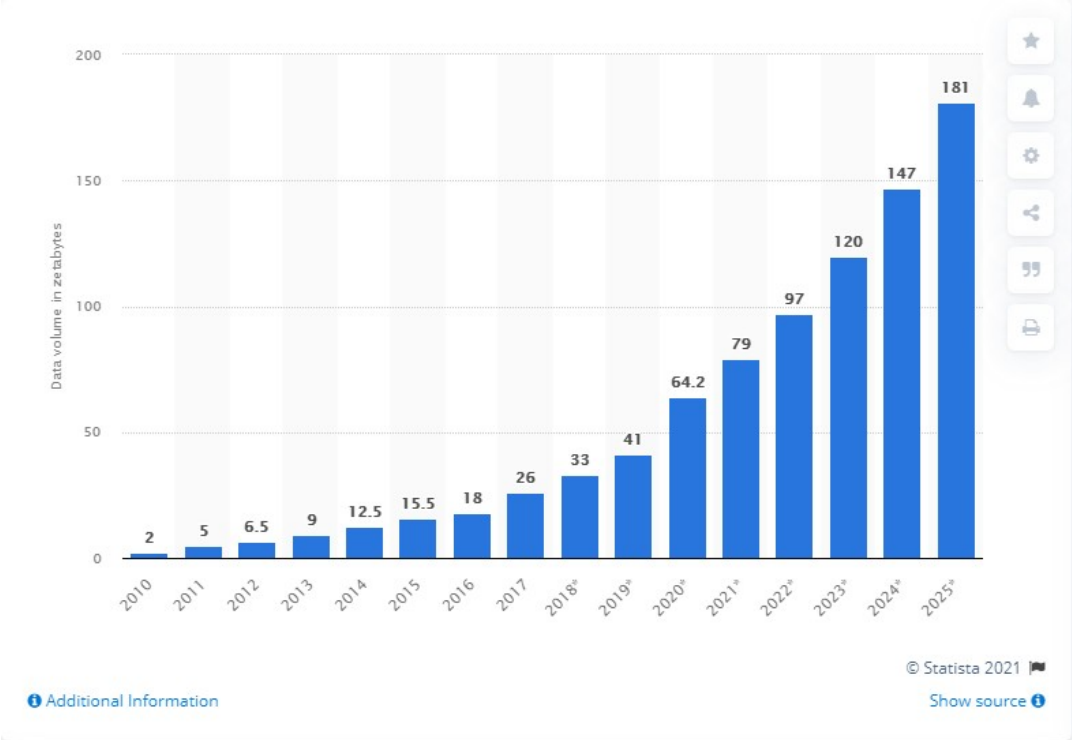
In linea con la forte crescita del volume di dati, si prevede che la capacità di archiviazione aumenterà, crescendo ad un tasso annuale del 19,2% nel periodo di previsione dal 2020 al 2025. Nel 2020, la capacità di archiviazione ha già raggiunto i 6,7 zettabyte.

Nella tabella che riporto di seguito è possibile osservare un'infografica che riproduce la curva di crescita della quantità di dati prodotti dal 2010, con una previsione fino al 2025, scontato ma interessante notare l'incredibile aumento che c'è stato tra il 2019 e il 2021[89].

I dati sono stati presi da Statista, secondo l'Università di Pisa una delle più grandi piattaforme online di dati statistici e di mercato al mondo, con accesso a più di 1,5 milioni di statistiche, previsioni, dossier, report ed infografiche su 80 mila argomenti provenienti da più di 22 mila fonti da tutto il mondo con un incremento di 300-500 nuovi dati al giorno. La piattaforma combina indicatori economici, dati di consumo, sondaggi di opinione e trend demografici[90].

Il grafico ci mostra, in zettabyte, una crescita costante nella produzione dei dati, che come abbiamo visto non è corrispettiva ad una crescita altrettanto esponenziale per l'archiviazione di questi dati.

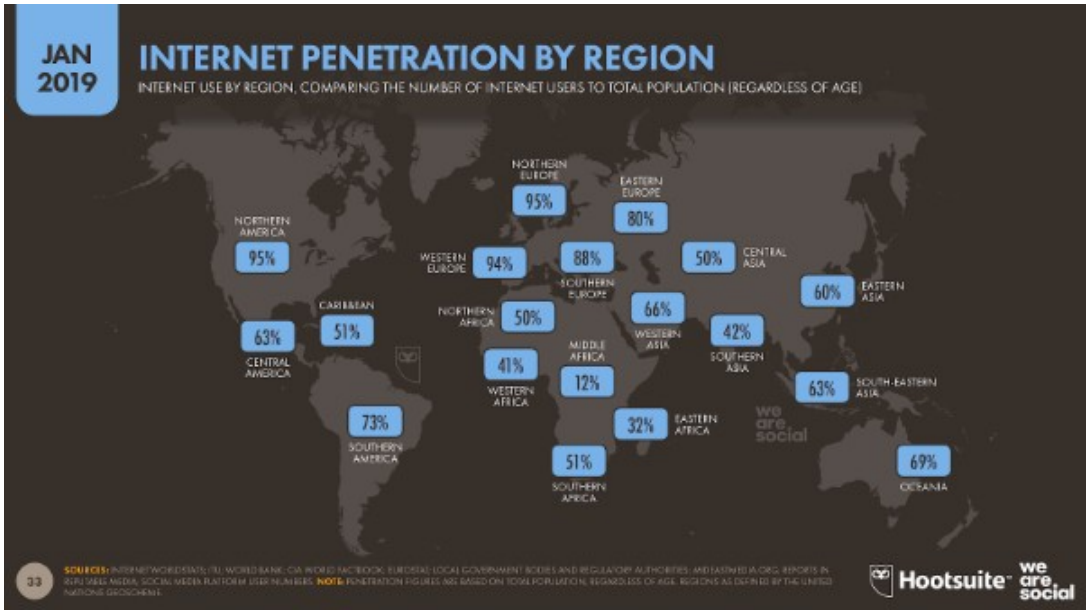
Tabella 1.3.1



Per comprendere al meglio come sia possibile raggiungere una mole tale di dati prodotti nel mondo ci è utile andare ad osservare altre infografiche prodotte da Wearesocial.com[91].

89 https://www.statista.com/statistics/871513/worldwide-data-created/ a questo link è possibile trovate l'infografica dinamica ultima consultazione 15/06/2021.
90 https://www.sba.unipi.it/it/risorse/banche-dati/statista ultima consultazione 1/07/2021
91 https://wearesocial.com/blog/2019/01/digital-2019-global-internet-use-accelerates ultima consultazione 1/07/2021

tabella 1.3.2



In questa infografica è possibile vedere in percentuale quanti sono gli utilizzatori di internet nel mondo, suddivisi per aree geografiche. Ciò che notiamo è che al 2019, data a cui si riferisce la tabella, internet è utilizzato dal 95% degli abitanti del nord America, dal 94% degli abitanti dell'Europa occidentale e dal 95% degli abitanti del nord Europa. Le percentuali di utilizzo rimangono tutte al disopra del 50% in gran parte del mondo tranne che nell'Asia centrale, dove si fermano al 50%, nel subcontinente indiano dove si arriva solo al 42% della popolazione, così come in Africa occidentale con un 41%, fino a toccare la percentuale più bassa nell'Africa centrale con un totale di 12%.[92]

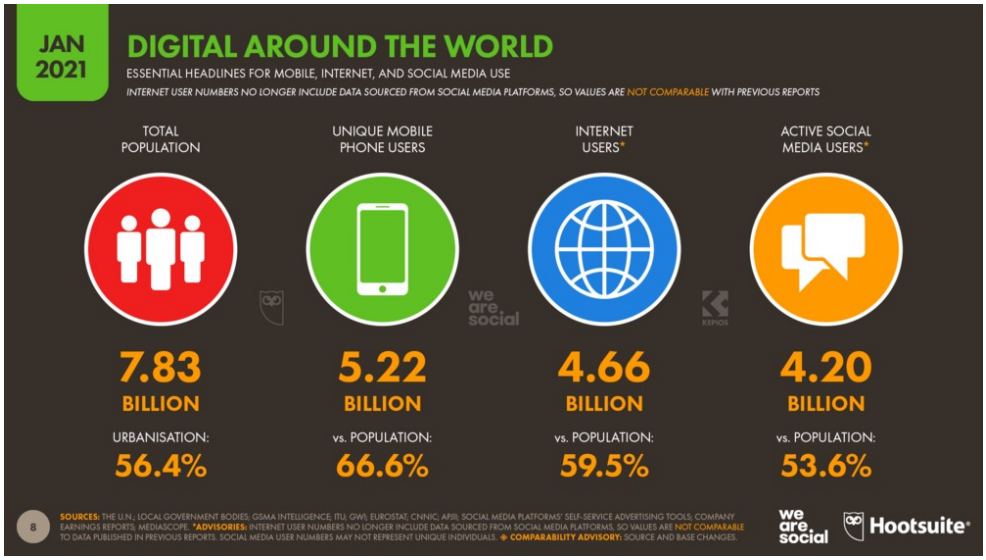
Questi dati non possono che farci riflettere su quanto scritto nel paragrafo precedente, dove riprendendo le teorie di Floridi, avevamo sottolineato come solo alcune zone del mondo vivano all'interno di società iperstoriche, mentre altre si trovano ancora in società storiche o preistoriche.

Dall'inizio del 2020 le Nazioni Unite riferiscono che la popolazione mondiale è cresciuta di 80 milioni di persone con un incremento annuo dell'1%, per un totale di 7,83 miliardi di persone presenti sul globo. Di queste 5,22 miliardi usano un telefono cellulare, pari al 66,6% della popolazione mondiale. Gli utenti unici di telefonia mobile sono cresciuti dell'1,8% (93 milioni) da gennaio 2020, mentre il numero totale di connessioni mobili è aumentato di 72 milioni (0,9%) per raggiungere un totale di 8,02 miliardi all'inizio del 2021. A gennaio 2021, 4,66 miliardi di persone in tutto il mondo usano internet, in aumento di 316 milioni, 7,3% rispetto a gennaio 2020. La penetrazione globale di internet è ora al 59,5%. Sui social media sono iscritti 4,20 miliardi di utenti nel mondo. Questa cifra è cresciuta di 490 milioni negli ultimi 12 mesi, con una crescita annuale di

92 https://wearesocial-net.s3.amazonaws.com/wp-content/uploads/2019/01/Screen-Shot-2019-01-30-at-12.02.22.png ultima consultazione 01/07/2021

oltre il 13%. Il numero di utenti dei social media è ora equivalente a più del 53% della popolazione mondiale.[93]

tabella 1.3.3



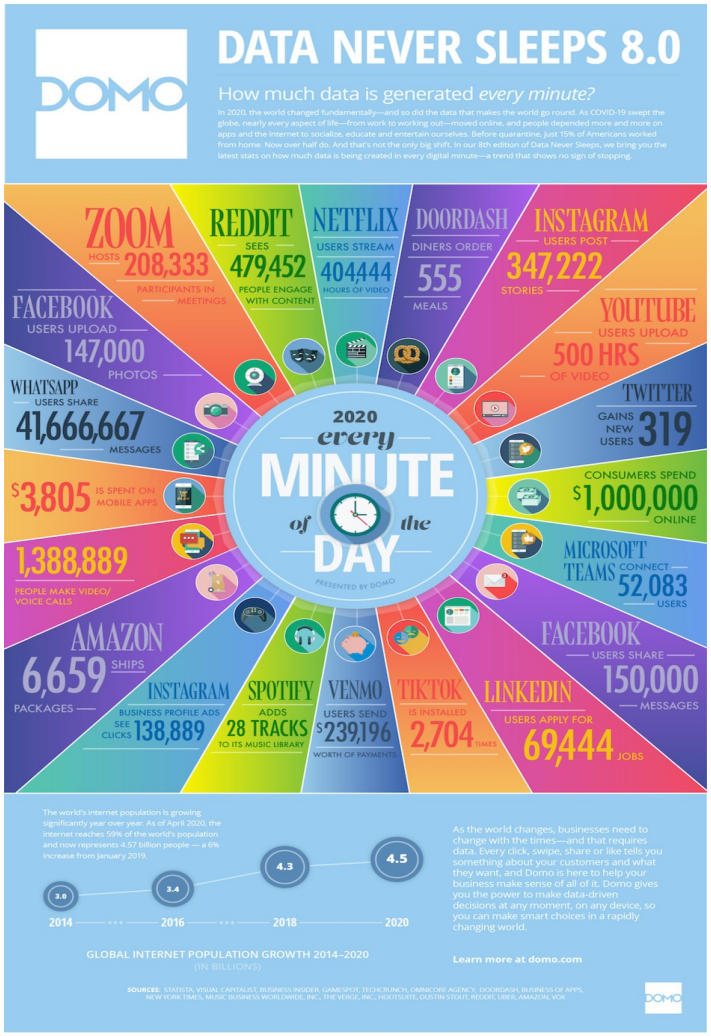
I dati che stiamo osservando ci sono necessari per dare una forma concreta a qualcosa che finché proviamo a comprendere tramite zettabyte può sembrarci eccessivamente astratta, queste infografiche rendono un quadro molto preciso di cosa sia oggi la nostra società e probabilmente ci restituiscono anche un' immagine più chiara di cosa siamo noi come cittadini e inforganismi.

Per l'ottavo anno DOMO[94] ha rilasciato l'infografica *Data never sleeps 8.0*[95] dove in modo semplice riunisce graficamente tutti i dati divisi per tipologia di utenza, che vengono generati ogni minuto di ogni giorno nel mondo.

In verità possiamo notare, sopratutto comparando con le infografiche degli anni passati, che DOMO non riporta tutti i cluster possibili dei dati prodotti, ma ne selezioni dei set andando a creare una media sui dati totali prodotti ogni minuto, cosa ancora più importante per noi in quanto storici, nelle grafiche di DOMO non sono presenti i dati generati da e con le pubbliche amministrazioni.

Tabella 1.3.4

93 https://datareportal.com/reports/digital-2021-global-overview-report ultima consultazione 01/07/2021
94 https://www.domo.com/ ultima consultazione 01/07/2021
95 https://web-assets.domo.com/blog/wp-content/uploads/2020/08/20-data-never-sleeps-8-final-01-Resize.jpg ultima consultazione 01/07/2021



Per quanto interessante, come abbiamo anticipato, questa infografica non può che essere una astrazione riportando una media di tutti i diversi dati che vengono generati ogni minuto, per questo mancano delle categorie di dati che invece potrebbero essere molto importanti nel lavoro dello storico del futuro, a parte da i dati sulle pubbliche amministrazioni, che non emergono mai nei lavori di DOMO, alcune informazioni le possiamo incorporare dall'analisi fatta sempre da DOMO nel 2019, particolarmente interessante è sapere che in un minuto nel mondo vengono spedite 188 milioni di email[96].

Possiamo ora dire di aver osservato anche se solo a livello superficiale, ciò che oggi si prova a gestire con i big data, ossia una mole enorme di dati che vengono generati e che rappresentano le nostre interazioni mutuate attraverso tecnologie digitali o, visto lo sviluppo delle ICT e dell'internet

96 https://web-assets.domo.com/blog/wp-content/uploads/2019/07/data-never-sleeps-7-896kb.jpg ultima consultazione 01/07/2021

delle cose, i dati generati e scambiati direttamente tra tecnologie, ciò che nel paragrafo precedente abbiamo imparato ad osservare come tecnologie di terzo ordine.
Comprendere i big data è complesso anche solo concettualmente perché richiede alla mente umana di doversi confrontare con delle quantità che non riusciamo a razionalizzare, è come provare a creare un immagine mentale del vuoto, per il nostro cervello è impossibile, ha bisogno di inserire nello spazio dei limiti o degli oggetti o dei segni. Il nostro è un limite computazionale.
Ma è giusto affermare che sia nel limite computazionale che paradossalmente risiede il significato dell'aggettivo *big* relativo ai dati? Ad oggi facciamo fatica a capire come gestire e come archiviare i big data perché le tecnologie a nostra disposizione non hanno sviluppato un potere computazionale tale da assicurarci di svolgere senza problemi questo lavoro, altresì è vero, come ci fa riflettere Floridi, che questo è solo un falso problema e che soprattutto quando svilupperemo sistemi con un potere computazionale sempre all'altezza di elaborare i dati che generiamo, questi stessi sistemi produrranno dati che andranno gestiti. Voler osservare la questione dei big data attraverso un punto di vista computazionale rischia dunque di essere un cortocircuito, una coazione a ripetere.
Ma in che senso questi dati sono grandi?
È grande lo spazio che serve per archiviarli? Sicuramente si, i server occupano uno spazio fisico anche se spesso non ci pensiamo, ogni mail che mandiamo, ogni messaggio che riceviamo, ogni post che pubblichiamo non esistono in un'astrazione di uno spazio immateriale, ma ne occupano appunto uno fisico e per esistere richiedono un grande consumo di energia. Ma chiaramente il problema non risiede nella grandezza fisica e neanche nelle capacità tecnologiche.
Ciò non dimeno lo stesso Floridi ammette che non è semplice definire il concetto big data.

A dispetto dell'importanza del fenomeno non è chiaro cosa esattamente significhi l'espressione big data. In questi casi si ha la tentazione di adottare l'approccio del Giudice della Corte Suprema degli Stati Uniti, Stewart Potter, allorché gli fu chiesto di descrivere la pornografia, rispose: "difficile da definire, ma la riconosco quando la vedo"[97]

E per rimanere in ambito sessuale è interessante la definizione di big data che diede Dan Ariely, docente di psicologia e behavioral economics alla Duke University: «Big data is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it… »[98]

97 L. Floridi, *la quarta rivoluzione. Come l'infosfera sta trasformando il mondo,* Raffaello Cortina editorie, Milano, 2017 p. 14
98 https://advanced-hindsight.com/blog/new-big-data/ ultima consultazione 01/07/2021

Negli Stati Uniti, la National Science Foundation (NSF) e il National Istitute of Healt (NIH) hanno identificato i big data come l'obbiettivo di un programma.

«L'espressione big data in esame fa riferimento ad ampi, diversi, complessi, longitudinali e/o distribuiti set di dati generati da strumenti, sensori, transazioni internet, email, video, numero di click e/o altre fonti digitali disponili oggi e nel futuro.»[99]

Secondo Floridi, che ci fa notare quanto sia «oscura e vaga»[100] la definizione appena data, per comprendere i big data dobbiamo concentrarci su un problema di potenza intellettiva e non computazionale, ciò che osserva è che la nostra difficoltà risieda nel non riuscire a realizzare pattern di piccola scala capaci di dare significato alla mole di dati prodotti e archiviati. Per farci capire cosa intende fa l'esempio di un enorme buffet del quale non riusciamo a mangiare tutto, ma nessuno ci costringe a masticare ogni singolo bit, siamo liberi di scegliere e in base alla nostra scelta provare a estrapolare soddisfazione per le nostre esigenze.

Questo passaggio credo che sia fondamentale per quello che sarà il lavoro dello storico del futuro, quali dati ci sono utili? In che modo possiamo pensare di organizzarli? Come sarà possibile inserirli in un contesto di significato che li trasformi da dati a informazioni?

«I big data sono destinati a crescere. l'unico modo per fronteggiare tale evento è sapere chi siamo e cosa stiamo cercando. Non si fa scienza semplicemente accumulando dati»[101]

A riguardo è interessante leggere ciò che scrive Francesco Maccelli nel suo saggio *Gli storici, le tecnologie e il lavoro*:

> La storia economica dipende intrinsecamente dall'acquisizione, dalla manipolazione e dalla pulizia delle scarse informazioni che provengo dal passato. […] Uno degli aspetti centrali appare l'estrazione dei pattern dal caos apparente delle fonti (eterogenee) e dai dati utilizzati dagli storici digitali. Infatti, gli studiosi sono abituati a problemi derivanti dal confronto tra quantità diverse per un lungo periodo di tempo. In questa direzione, i big data, forse più di altri, hanno la necessità di una metrica chiara.[102]

Come possiamo notare anche Macelli, da storico, pone la questione dell'analisi dei big data da un punto di vista non solo quantitativo, quando sostiene che probabilmente è necessario capire come estrapolare dei pattern per riuscire ad organizzare l'incredibile mole di informazioni, ci porta nuovamente a riflettere sulla possibile importanza dei concetti per l'analisi della storia. In un

99 https://www.nsf.gov/cise/bigdata/ ultima consultazione 5/07/2021

100 L. Floridi, *la quarta rivoluzione. Come l'infosfera sta trasformando il mondo,* Raffaello Cortina editorie, Milano, 2017 p. 15

101 Ivi, p.17

102 F. Maccelli, *Gli storici, le tecnologie e il lavoro,* D. Paci, a cura di, AA.VV., *Storia in Digitale. Teorie e metodologie*, Unicopli, Milano 2019, p. 179

sistema sociale che produce quantità enormi di dati, si stano sviluppando delle tecnologie che consentono di immaginare una catalogazione automatica del documento, queste tecnologie, come vedremo, hanno già una storia di applicazione abbastanza strutturata grazie alle potenzialità offerte da standard archivistici internazionali, che grazie ai metadati trovano una trasposizione nell'ambito informatico. Ma è nell'autoconsistenza del dato, che prende forma nel Web semantico, che potenzialmente si può assistere ad un processo di descrizione, sempre attraverso un lavoro di metadatazione, tale da consentire un processo preliminare di clusterizzazione potenzialmente utile a coadiuvare il lavoro dello storico che si troverà a doversi confrontare con i big data.

1.7 Digital divide e accessibilità

Le tecnologie digitali sono dotate di un grande potere computazionale sintattico, meta sintattico e quantitativo, ma non hanno capacità semantiche, per quanto in grado di apprendere non riconoscono i significati, come scrive anche Benasayag il cervello umano a differenza di un cervello elettronico è situato in un corpo che è capace di fare esperienza del mondo attraverso diversi livelli, le macchine invece non sono situate, non hanno percezione del loro io in uno spazio e in un tempo.
Per noi esseri umani i limiti del nostro corpo e delle nostre capacità cerebrali non sono un problema, ma la condizione necessaria per sviluppare conoscenza e pensiero[103].
In altre parole per riuscire a risolvere problemi complessi l'essere umano ha bisogno di fare esperienza del limite, così da poter generare un nuovo pensiero che in relazione con l'ambiente proponga delle soluzioni, le macchine ancora sono strumenti che noi stiamo addestrando ad essere sempre più funzionali, ma che sono solo in grado di leggere il mondo non di comprenderlo, per questo Floridi insiste sulla necessità di tenere l'essere umano al centro dell'analisi dei dati e di sviluppare le Human Data Science, certo è che se parte della popolazione non è in grado di fruire delle tecnologie digitali non potrà neanche essere inclusa all'interno di un processo di elaborazione dei dati.

Con Human Data Science si intende un modello multidisciplinare che vuole offrire una nuova chiave di lettura dei big data, mantenendo le persone al centro per capire i loro bisogni e verificare come soddisfarli. Ogni giorno le persone producono una quantità di dati infinita: si parla di 2,5 x1018 Terabyte, e il trend è destinato ad aumentare esponenzialmente. Come interpretarli e valorizzarli? Un aiuto arriva dalla tecnologia attraverso i big data e l'Intelligenza artificiale, ma per gestire questo prezioso patrimonio nel modo più

103 M. Benasayag, *Il cervello aumentato l'uomo diminuito*, Erickson, Trento 2015, p. 60

opportuno stanno entrando in gioco nuovi modelli. Modelli che, appunto, devono riuscire a comprendere i bisogni delle persone tramite la modellazione dei dati utilizzando un approccio umano-centrico[104].

Se vogliamo riuscire a tenere l'essere umano al centro delle analisi sui dati è necessario però che l'essere umano abbia compiutezza nell'utilizzo e del funzionamento dei sistemi tecnologici.
Serge Noiret, nel suo saggio *Storia contemporanea digitale*[105], porta la nostra attenzione sul concetto di datificazione che l'autore ha tradotto da Viktor Mayer-Schonberger e Kennet Cukier, non solo per riaffermare come e quanto statistiche, documenti e fonti siano oggi prodotti prevalentemente in formato digitale e che per essere esperiti richiedano dunque competenze tecniche e tecnologiche specifiche, ma ci invita a riflettere anche su un tema centrale all'interno dell'infosfera, ossia il digital divide.
Per digital divide si intende letteralmente divario digitale, ciò è inteso come mancanza di accesso e di fruizione alle nuove tecnologie di comunicazione e informatiche. Da qualche anno ormai si parla di questo argomento, che con il passare del tempo riguarda aspetti sempre diversi delle nuove tecnologie e non solo, molti sono gli aspetti anche sociali della questione.
Storicamente, i primi che parlarono di digital divide furono Al Gore e Bill Clinton, quando, all'inizio degli anni novanta, intrapresero una politica di forte sviluppo e potenziamento dell'infrastruttura di internet negli Stati Uniti. Il concetto di divario digitale era riferito alla difficoltà di accesso a internet in determinate zone del paese, una difficoltà intesa e questo è un passaggio centrale, anche sotto l'aspetto economico. In quegli anni internet esplose come fenomeno di massa e divenne sempre di più un mezzo di lavoro, non essere connessi alla rete o non avere gli strumenti cognitivi per farlo, significava quindi essere relegati ai margini della società[106].
Questo è valido ancora oggi dove il digital divide incide sulla società, anche italiana, su tanti differenti livelli; il più semplice da comprendere è quello dato da una appartenenza anagrafica, secondo dati Istat, tra le famiglie resta un forte divario digitale da ricondurre soprattutto a fattori generazionali e culturali. La quasi totalità delle famiglie con almeno un minorenne dispone di un collegamento a banda larga 95,1%, mentre tra le famiglie composte esclusivamente da persone ultra sessantacinquenni tale quota scende al 34,0%[107].
Non avere un accesso ad internet può diventare sempre di più un limite tangibile nella gestione della propria vita, pensiamo a come ormai anche se faticosamente, le pubbliche amministrazioni italiane abbiano digitalizzato i loro processi, ciò significa che sempre di più è possibile svolgere

104 https://www.insuranceup.it/it/call4ideas/luciano-floridi-per-avere-successo-con-i-dati-ci-vuole-etica/ ultima consultazione 3/07/2020
105 R. Minuti, a cura di, AA.VV, *Il web e gli studi storici*, Carocci Editori, 2017 Roma p. 275
106 http://freedom.dicea.unifi.it/pub/pluto/files/journal/pj0207/Digital_Divide.html ultima consultazione 2/07/2021
107 https://www.istat.it/it/archivio/236920 ultima consultazione 2/07/2021

adempimenti tecnici e burocratici tramite portali online, le banche stanno chiudendo molte delle loro filiali perché l'home banking sta diventando uno strumento sempre più diffuso e utilizzato, l'autodichiarazione che per alcuni mesi durante la pandemia è stata obbligatoria se si voleva uscire di casa, la si poteva ottenere solo scaricandola da internet, o se fermati dalle forze dell'ordine era possibile compilarla in loco, ma nessun ministero ha mandato a casa dei cittadini anche solo una copia cartacea da poter poi fotocopiare.

A giugno 2021 le identità digitali SPID erogate sono 22.195.216[108], e il Piano Nazionale di Ripresa e Resilienza (PNRR) ha, tra i suoi obiettivi, quello di diffondere l'identità digitale assicurando che venga utilizzata entro il 2026 dal 70% della popolazione.[109]

L'abbattimento del digital divide rientra quindi nei piani di ripresa del Governo italiano, ma come accennavamo prima, il digital divide non è legato solo a questioni di ordine anagrafico o culturale, vi è un livello molto più complesso da analizzare.

La maggior parte delle famiglie senza accesso a internet da casa indica come principale motivo la mancanza di capacità 56,4% e il 25,5% non considera Internet uno strumento utile e interessante. Seguono motivazioni di ordine economico legate all'alto costo dei collegamenti o degli strumenti necessari 13,8%, mentre il 9,2% non naviga in rete da casa perché almeno un componente della famiglia accede a Internet da un altro luogo.[110]

Il dato su cui voglio portare l'attenzione è quel 13,8% di famiglie che non hanno internet perché troppo costosa la connessione o gli strumenti per poterla sfruttare. Questo ci fa capire come la datificazione, per riprendere il ragionamento di Noiret, si presenti anche come un problema di natura economico, questo è un tema di natura sociale in termini generali, dove porzioni di popolazioni sono escluse dall'infosfera perché è troppo costoso farne parte. La questione economica si riverbera anche nell'ambito culturale e scientifico nel momento in cui non tutte le infrastrutture pubbliche si possono permettere di attuare un processo di digitalizzazione dei loro sistemi di archiviazione, così da rimanere *de facto* fuori dalla possibilità di partecipare con il loro patrimonio ad una società iperstorica.

Infatti, le biblioteche digitali accessibili con abbonamenti commerciali non sono alla portata di un click per permettere a chiunque di effettuare operazioni di tex mining per fare ricerca su migliaia di documenti scansionati in contesti scientifici. Una parte rilevante di questi documenti datificati, benchè disponibili ad essere analizzati grazie alle tecnologie, non è accessibile in quanto si tratta di dati commerciali. A differenza delle biblioteche digitali pubbliche […] le biblioteche commerciali accrescono il digital divide tra chi accede

108 https://avanzamentodigitale.italia.it/it/progetto/spid ultima consultazione 2/07/2021
109 https://innovazione.gov.it/progetti/identita-digitale-spid-cie/ ultima consultazione 2/07/2021
110 https://www.istat.it/it/archivio/236920 ultima consultazione 2/07/2021

e chi non se lo può permettere, tra paesi e biblioteche che hanno i mezzi per finanziare abbonamenti costosi e chi, a casa nostra come in molte parti del mondo non se lo può permettere. […] In questi casi il digitale perde un suo connotato essenziale, ossia il facile accesso alla cultura mondiale anche da parte di paesi che non possiedono infrastrutture paragonabili alle grandi biblioteche europee e americane e, in Europa e America, tra chi si può o meno permettere oggi di abbonarsi ai servizi digitali.[111]

La questione dell'accessibilità ai dati sta prendendo sempre più importanza anche all'interno del dibattito storico e storiografico, se pensiamo infatti all'appello della storia come bene comune lanciato da Andrea Giardina, Liliana Segre e Andrea Camilleri[112], come abbiamo scritto con Deborah Paci, Serge Noiret e Marcello Ravveduto[113], la digitalizzazione dovrebbe rendere la storia sia nell'ambito delle fonti che della storiografia, più accessibile per chiunque abbia una connessione ad internet.

Come abbiamo potuto vedere però il digital divide investe sia parte della popolazione che non ha la possibilità di avere una connessione ad internet o la capacità di usare in modo adeguato gli strumenti digitali, sia chi ha strumenti e competenze ma è comunque precluso per motivi commerciali e di proprietà dall'accesso diretto ai dati.

Il tema dell'accessibilità ai dati è un tema centrale e che ci porta a dover ragionare sulla relazione tra dato e informazione. Prima di provare a dare un contesto epistemologico a questa riflessione credo che possa essere utile partire da alcuni esempi pratici.

Poche pagine sopra è possibile vedere una serie di infografiche che sintetizzano in termini quantitativi quali e quanti dati produciamo nel mondo, prendiamo per esempio la tabella 1.3.1, quella relativa alla quantità di zettabyte prodotti nell'anno, l'ente che l'ha realizzata è Statista, società di statistica tedesca di cui abbiamo già parlato, ciò che ora ci interessa notare è che se tramite la loro pagina web è possibile vedere e fruire liberamente dell'infografica, se si vuole accedere ai dati che Statista ha utilizzato per realizzare quel grafico, dobbiamo sottoscrivere un abbonamento mensile a pagamento. In realtà, dopo le prime visualizzazioni, che avvengono senza nessun tipo di problema, l'interfaccia del sito muta, andando a coprire l'infografica con una finestra di dialogo che ti chiede di sottoscrivere l'abbonamento, anche per scaricare il grafico è richiesto di abbonarsi, fortunatamente questo problema l'ho risolto prima che si presentasse, andando a realizzare uno screenshot della pagina ritagliando solo la parte dell'infografica, così da

111 S. Noiret, *Storia contemporanea digitale,* in R. Minuti, a cura di, AA.VV, *Il web e gli studi storici. Guida critica all'uso della rete*, Carocci Editori, 2017 Roma p. 276

112 http://www.icviacutigliano.it/sito/news_2018-2019/Manifesto_Storia ultima consultazione 5/07/2021

113 S. Noiret, M. Scanagatta, D. Paci, M. Ravveduto, *La storia come bene comune, le nuove frontiere della public history digitale*, «Passato e Presente», F. 113, Franco Angeli, Milano 2021, p. 119-134

“appropriarmi” per fini puramente di ricerca, di una immagine digitale che ho potuto condividere in queste pagine[114].

Quello che però ci deve far riflettere è che se anche pur limitatamente l’informazione mi è stata possibile fruirla gratuitamente, per accedere ai dati, invece, mi è stato richiesto da subito di pagare. Il tema è delicato e controverso, se nella realizzazione dell’infografica è inevitabile riconoscere la proprietà anche intellettuale di chi l’ha realizzata, non possiamo non soffermarci sul fatto che i dati attraverso cui è stato possibile realizzare quel prodotto informativo li abbiamo creati noi utenti, quindi in qualche modo ci appartengono, appartengono all’umanità intera.

In questo caso credo che sia inevitabile riconoscere un ruolo centrale al potere computazionale come espressione di potere economico, non sono tanto i dati grezzi ad aver un valore, ma piuttosto gli strumenti per leggerli, comprenderli e organizzarli. Questi strumenti è possibile osservarli come delle infrastrutture che consentono l’utilizzo dei dati. Spesso si compara i dati al petrolio come risorsa di grande ricchezza, ma sia Durante che Floridi, condividono l’idea che questa comparazione sia sbagliata, il potere economico risiede nella capacità di estrapolare informazione da questi dati non tanto nei dati stessi, così come il potere economico non risiede tanto nel petrolio che giace sottoterra, ma nelle infrastrutture per estrarlo spostarlo e raffinarlo.

Ma quindi qual è la differenza tra dato e informazione?

1.8 Da dato a informazione. Da evento a fonte

I concetti di dato e informazione sono dalla metà degli anni cinquanta del novecento, al centro di un complesso dibattito teorico e filosofico che vede confrontarsi diverse discipline.

Il mio scetticismo riguardo a una analisi definitiva dell'informazione coinvolge la famigerata versatilità dell'informazione. La nozione di informazione è stata utilizzata per caratterizzare una misura dell'organizzazione fisica (ossia la riduzione dell'entropia), una configurazione [*pattern*] nella comunicazione fra emittente e ricevente, una forma di controllo e feedback, la probabilità che un messaggio venga trasmesso attraverso un canale di comunicazione, il contenuto di uno stato cognitivo, il significato di una forma linguistica o la riduzione dell'incertezza. Questi concetti di informazione vengono definiti in varie discipline come la fisica, la termodinamica, la teoria della comunicazione, la cibernetica, la teoria statistica dell'informazione, la psicologia, la logica induttiva, ecc. Sembra quindi non esserci un'unica idea di

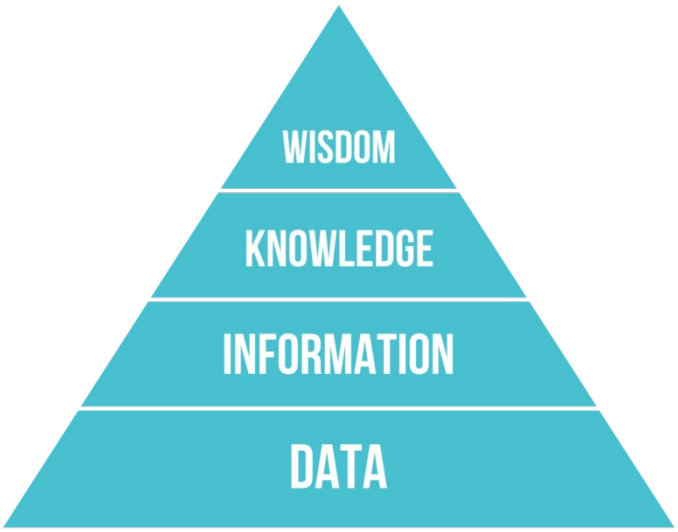
114 Il tema dell'appropriazione delle fonti digitali e digitalizzate ritengo che sia di grande interesse su almeno due livelli, il primo è di natura telepistemica, il secondo riguarda direttamente le pratiche possibili di Public History e la possibilità di mettere in scena le fonti. Entrambi questi argomenti li affronteremo nei capitoli successivi.

informazione su cui questi vari concetti convergano e quindi neppure una specifica teoria dell'informazione[115].

Per alcuni filosofi il concetto è così ambiguo che dovrebbe essere bandito dalla scienza[116], per altri invece assume un valore metafisico, come per Buckland che scrisse «non possiamo dire con certezza di niente che non sia informazione».[117]

Ciò che però a quanto pare non è messo in discussione da nessuna teoria sull'informazione è che alla base di questa vi sono i dati. Quando scrivo alla base non è solo una locuzione metaforica, infatti nelle scienze dell'informazione molto del ragionamento si sviluppa di conseguenza alla realizzazione della piramide DIKW, che vede messe in scala dal basso verso l'altro: i dati, l'informazione, la conoscenza e la saggezza.

In realtà altre piramidi sono state teorizzate in conseguenza a questa, sempre aggiungendo degli strati, tra cui l'esperienza, ma in questo momento noi ci concentreremo solo sui primi due livelli, i dati e l'informazione.



Una prima definizione di dato, in ambito informatico, l'ho già data nel paragrafo precedente, ma credo che possa essere utile andarla a riprendere così da potercisi soffermare.

115 R.J. Bogdan, *Grounds for cognition: how goal-guided behavior shapes the mind*. Hillsdale: Erlbaum, 1994, p. 53

116 R. Thom, *Dalla morfogenesi alla struttura*. In: *Teoria dell'informazione*, atti del Seminario tenutosi a Venezia nella sede della Fondazione Giorgio Cini dal 28 maggio al 1 giugno 1973, a cura di Jacques Roger. Bologna: Il Mulino, 1974, p. 73-86: p. 83. Cfr, R. Ridi, *la piramide dell'informazione. Una introduzione*, https://aibstudi.aib.it/article/view/11903/11481#5 ultima consultazione 3/07/2021

117 Michael K. Buckland, *Information and information systems*. New York: Praeger, 1991, p. 50

In informatica «i dati sono delle rappresentazioni originarie, cioè non interpretate, di un fenomeno, evento, o fatto, effettuate attraverso simboli o combinazioni di simboli, o di qualsiasi altra forma espressiva legate a un qualsiasi supporto»[118].

I dati sono rappresentazioni, dunque di per sé l'espressione di qualcosa che esiste in se stessa, sono originari e non interpretati, una manifestazione pura, non soggettiva.

Ricardio Ridi li definisce come informazioni più semplici, grezzi, non integrati, quando invece l'informazione è più sofisticata, ma non è chiaro se nella sua complessità l'informazione sia oggettiva o soggettiva[119]. È per questa ambiguità che Ridi nel suo lavoro *La piramide dell'informazione*, torna sulla differenza tra dato e informazione sostenendo che sarebbe sbagliato omologarli o comunque definire le informazioni solo dei dati maggiormente contestualizzati, questo perché il dato ha di per sé un valore, sarebbe altrimenti superfluo destinare a informazioni e dati due livelli della piramide[120].

Foskett ha provato a descrivere il rapporto tra dati, informazione e conoscenza con un'altra rappresentazione grafica che ha chiamato "piramide della saggezza" che prevede anche un quinto livello, preliminare ai quattro citati, ossia l'universo dei fenomeni. Da esso gli esseri umani estraggono, grazie all'investigazione, i dati, che diventano informazione dopo essere stati organizzati e conoscenza dopo essere stati assimilati. Infine, grazie all'esperienza sociale, la conoscenza diventa saggezza, che «si fonda su un'ampia conoscenza e una profonda comprensione di tutti gli aspetti e gli eventi della vita in una comunità, ossia sulle nostre relazioni con le persone viventi, incluse le relazioni indirette che instauriamo attraverso i libri» [121].

Leggendo queste parole, si può trovare una convergenza con il pensiero di Bloch quando scrive che

l'oggetto della storia è per natura l'uomo. O meglio: gli uomini. Più che il singolare, favorevole all'astrazione, il plurale, che è grammaticale della relatività, conviene a una scienza del diverso. Dietro i tratti concreti del paesaggio, dietro gli utensili o le macchine, dietro gli scritti che sembrano più freddi e le istituzioni in apparenza più totalmente distaccate da coloro che le hanno fondate, sono gli uomini che la storia vuole afferrare[122].

118 G.M. Di Nunzio, *Sistemi di elaborazione delle informazioni*, Università degli studi di Padova, 2014, http://www.dei.unipd.it/~dinunzio/fdi-2014-2015/04_dati_informazioni.pdf ultima consultazione 30/05/2020

119 R. Ridi, *Piramidi delle informazioni e KO,* 10° incontro ISKO Italia, Università di Pavia, 14/05/2021, a questo link è possibile sentire il suo intervento http://www.riccardoridi.it/video/isko2021.mp4 ultima consultazione 3/07/2021

120 R. Ridi, *La piramide dell'informazione, una introduzione, https://aibstudi.aib.it/article/view/11903/11481#5* ultima consultazione 4/07/2021

121 D. J. Foskett, *Libraries and information systems: a fruitful partnership*. In: *Proceedings of the Annual Conference of CAIS*. Canadian Association for Information Science, 1995, p. 13 https://journals.library.ualberta.ca/ojs.cais-acsi.ca/index.php/cais-asci/article/view/405/353 ultima consultazione 3/07/2021

122 M. Bloch, *Apologia della storia o mestiere di storico,* Einaudi Editore, Torino 1998, p. 22

Aggiungendo in nota, citando Monod, che cita Michelet al corso all'Ecole normale del 1829 «noi ci occupiamo sia dello studio dell'uomo singolo, e sarà filosofia, sia dello studio dell'uomo sociale, e sarà storia»[123]

I dati sono dunque alla base dell'esperienza cognitiva umana, esperienza che quando organizzata si manifesta nella strutturazione della società che a sua volta è manifestazione dell'agire umano, ossia di ciò che secondo Bloch deve o dovrebbe interessare lo storico e la storica.

Per tornare sulla definizione di dato, Durante sintetizza la complessità mutuando il pensiero di Floridi e scrivendo che un dato può definirsi come una mancanza di uniformità nel mondo, un punto di discontinuità, un suono che interrompe il silenzio, un punto nero su una pagina bianca, una discontinuità che ci fa percepire una differenza.

E cos'è un evento se non una discontinuità? Un fatto lo si afferma attraverso l'azione ed è in quell'agire che l'uomo si manifesta con la sua esistenza nella società, si tratta poi, da un punto di vista storico, di comprendere cosa può rimanere di quel fatto nel tempo.

Se camminando inciampo e cado, è un dato di fatto che sia successo qualcosa, se quella caduta però non è osservata da nessuno, se io stesso non ne farò mai parola, esisterà solo in se stessa, se invece quell'evento è stato osservato e mettiamo caso registrato perché ho deciso di andare in ospedale, vi saranno delle informazioni su quell'evento, delle tracce che se organizzate potranno dare conoscenza.

Senza voler andare a toccare la parte più alta della piramide, non interessandoci dunque della saggezza che può diventare azione positiva, troviamo una correlazione con quanto afferma Cooley «i dati opportunamente organizzati e trattati possono diventare informazione, e l'informazione che viene assorbita, compresa e applicata dalle persone può diventare conoscenza.»[124]

Riccardo Ridi ci può essere ulteriormente di aiuto[125].

Definire i dati come «una differenza che genera una differenza»[126] è una mossa sicuramente azzeccata, perché da una parte libera il concetto di informazione dall'impegnativa caratteristica della sussistenza anche

123 Ibidem, nota n° 4

124 M. Cooley, *Architect or bee: the human price of technology*. London: Hogarth, 1987, p. 11

125 Per completezza di informazione nella citazione riporterò anche le note presenti originariamente nel testo.

126 La prima formulazione di Bateson (successivamente ripetuta, con piccole varianti, in numerose sue opere che vengono di volta in volta citate come fonte) risale probabilmente a una relazione sul *Doppio vincolo* tenuta nell'agosto 1969 a un convegno dell'American Psychological Association, il cui testo è stato tradotto in italiano, con tale titolo, in Gregory Bateson, *Verso un'ecologia della mente*, traduzione di Giuseppe Longo. Milano: Adelphi, 1976, p. 293-302. Benché in tale edizione il testo originale «any difference which makes a difference» sia reso, a p. 294, come «una differenza che genera una differenza», va segnalato che in altri capitoli dello stesso volume e in altre traduzioni di varie opere di Bateson ricorrono anche l'espressione «una differenza che produce una differenza» e le ancora più criptiche varianti «una differenza che fa differenza» e (di gran lunga la più diffusa, secondo Google e Google scholar) «una differenza che fa la differenza».

in assenza di un soggetto percipiente (come può esserci informazione senza qualcuno che venga informato? [127]), e dall'altra svincola il concetto di dato dall'altrettanto delicato legame con la verità (un'informazione può essere vera o falsa, ma non ha senso porsi dubbi sulla veridicità di un dato definito diaforicamente perché esso non afferma alcunché, ma semplicemente esiste[128]). Inoltre definire i dati come mere discontinuità oggettive rispetto a uno sfondo neutro, prive di un proprio significato autonomo e univoco in assenza di una interpretazione esterna, costituisce un'ottima premessa per definire poi le informazioni come dati contestualizzati e interpretati, e quindi (solo ora) significativi per un soggetto (diventando «una differenza che genera una differenza per qualcuno o per qualcosa o da un punto di vista[129]») e passibili di essere considerati veri o falsi. [130]

Questa relazione tra dati e informazione, così descritta da Riccardo Ridi, può portarci a riflettere sulla relazione tra fatto/dato e informazione/fonte. È ovvio che non si possa affermare che ogni dato rappresenti un evento, ma credo che sia interessante sottolineare come la definizione di dato si possa effettivamente avvicinare a quella di evento e di come in questo processo emerga il complesso tema della verità, declinato anche nell'accezione di oggettività e soggettività.
Il dibattito metastorico sul valore epistemologico delle fonti ci chiede di doverci confrontare tra gli altri sul lavoro di Topolski che in una visione costruzionista, ma non antirealista, osserva le fonti come complessità narrative atte a descrivere l'avvenire di un evento.
È possibile dunque osservare le fonti come elementi informativi, complessità composte da una eterogeneità di dati.
Secondo Floridi la visione diaforica del dato, dunque l'affermarsi di qualcosa che esiste nel manifestarsi come differenza, va osservata attraverso una triplice declinazione[131]:

- I dati come *diaphora de re*: si tratta della differenza che è percepibile nel mondo reale, come ad esempio un suono che interrompe il silenzio o una luce rossa su uno sfondo nero.
- I dati come *diaphora de signo*: si tratta della differenza che è percepibile tra almeno due stati fisici di un sistema o segnali, come ad esempio il voltaggio più alto o più basso di una carica elettrica o il punto e la linea dell'alfabeto Morse.

127 Per una sintesi della discussione intercorsa fra Marcia J. Bates e Birger Hjørland, fra il 2005 e il 2011, sulla natura oggettiva o soggettiva dell'informazione cfr. Riccardo Ridi, *Livelli di irrealtà: oggettività e soggettività nell'organizzazione della conoscenza*, «Bibliotime», 18 (2015), n. 2, http://www.aib.it/aib/sezioni/emr/bibtime/num-xviii-2/ridi.htm, par. 2

128 «La veridicità non è una proprietà intrinseca dei dati stessi, ma è piuttosto determinata dalla situazione in cui i dati sono generati. [...] I dati non sono, di per sé, una forma di conoscenza. Piuttosto, i dati hanno bisogno di essere interpretati per produrre conoscenza» (S. Leonelli, *The philosophy of data* cit., p. 199)

129 Birger Hjørland, *Information: objective or subjective/situational?*, «Journal of the american society for information science and technology», 58 (2007), n. 10, p. 1448-1456: p. 1449

130 R. Ridi, *La piramide dell'informazione, una introduzione*, *https://aibstudi.aib.it/article/view/11903/11481#5* ultima consultazione 4/07/2021

131 L. Floridi, The Philosophy of information, Oxford University Press, Oxford 2011, p. 85

- I Dati come *diaphora de dicto*: si tratta della differenza che è percepibile tra due simboli come ad esempio tra due lettere dell'alfabeto

I dati di cui abbiamo parlato nelle pagine precedenti, possono essere letti come *diaphora de dicto*, sono definiti da una serie numerica di simboli, sono impulsi tradotti in codice binario, dunque inequivocabilmente non eventi e la prima contestualizzazione informazionale è data dalla capacità di macchine di leggere quei simboli, aggregarli e trasformali appunto in informazioni intelleggibili.
Non di meno ciò che è importante osservare con Floridi è che tali differenze sono mancanze di uniformità, meri punti di discontinuità[132] «puri dati o dati protoepistemici, vale a dire, dati prima di essere interpretati in chiave epistemica. In quanto fratture nella fabbrica dell'essere, possono essere posti solo come ancoraggio esterno per le nostre informazioni, poiché non è possibile accedere o elaborare tali dati indipendentemente da un livello di astrazione»[133].
E' necessario dire però che queste suggestioni ci suggeriscono delle riflessioni di natura epistemologica e non ontologica, nella differenza che descrive Ferraris, dove l'epistemologia è osservata come sfera del sapere e l'ontologia come sfera dell'essere[134].
Quello che sto proponendo dunque non è tanto una comparazione tra dato e evento e tra fonte e informazione, ma più semplicemente la possibilità di farsi guidare dalla filosofia dell'informazione, così come è già avvenuto con la semiotica e la linguistica, nel provare a delineare delle definizioni o forse descrizioni o forse spiegazioni, di fatto, fonte, storia e storiografia all'interno di un processo di datificazione della società.
Se i dati si mostrano come l'elemento puro, non interpretato, dell'agire umano all'interno di un ambiente digitale, e, se come teorizza Ferraris la registrazione, dunque la produzione di dati, è alla base della costruzione di un oggetto sociale che deriva da un atto sociale, in che modo storiche e storici possono o potranno definire l'importanza del dato all'interno della loro disciplina?
È nella filosofia di Ferraris, specificatamente in *Scienza nuova. Ontologia della trasformazione digitale* che ho rintracciato i prodromi di questo ragionamento.

L'ipotesi di fondo che mi ha guidato è che il web sia un apparato di registrazione, come la scrittura, il denaro, l'inventario e l'archivio. Il più grande apparato di registrazione che la storia abbia conosciuto fino a qui e il più potente.
D'accordo con la legge fondamentale dell'ontologia sociale, la definizione degli oggetti sociali è oggetto=atto registrato: un oggetto sociale è il risultato di un atto sociale (tale da coinvolgere almeno due

132 M. Durate, *Potere computazionale. l'impatto delle ICT su diritto società, sapere*, Meltemi, Milano, 2019 p. 154
133 L. Floridi, The Philosophy of information, Oxford University Press, Oxford 2011, p. 85
134 M. Ferraris, *Documentalità. Perché è necesario lasciare tracce*, Laterza, Bari Roma 2019, p. 59

soggetti, un soggetto e una macchina delegata, o due macchine delegate) che ha la caratteristica di essere registrato su un qualche supporto.
L'informazione è solo un effetto secondario del web, la vera forza motrice che lo anima è la registrazione, che nel web attraverso il processo di spacchettamento dei documenti è prioritaria rispetto all'informazione.[135]

Il pensiero di Ferraris ci offre diversi sopunti, ma porta a nche a doverci porre delle domande in quanto storici. In prima istanza l'idea del web come apparato di registrazione, dunque come il più grande e potente archivio della storia, ma cosa contiene questo archivio, certo, dati, ma appunto, come definire questi dati, quando sono intellegibili? Quando sono informazione? Quando sono utilizzabili? In che modo lo storico del presente e del futuro si può relazionare con questi nuovi documenti? E soprattutto, i dati hanno davvero un valore per una indagine storica che non voglia essere quantitativa? E' sempre in queste poche righe che Ferrrais indica la strada per trovare delle risposte ad alcune domande, e ciò avviene grazie al concetto di registrazione come processo propedeutico all'informazione e che trova una ricaduta nella nostra disciplina, sempre in correlazione al rapporto con i sistemi digitali, nel suo compenetrare in un concetto complesso e affascinante come quello di oggetto sociale, che inevitabilmente, ripensando a Bloch, non può che cogliere il nostro interesse.

135 M. Ferraris, *Scienza nuova. Ontologia della trasformazione digitale*, Rosenberg e Sellier, Torino 2018, p. 21

2 – Fonti e rivoluzione digitale

2.1 Da artefatti tangibili a oggetti digitali. La nuova forma del documento

Le società contemporanee sono indubbiamente società informazionali, gli esseri umani che le abitano passano molto tempo della loro giornata, per diletto o per lavoro, a produrre, scambiare e archiviare informazioni.

Ogni interazione su un social network, ogni articolo di giornale o blog, ogni pubblicazione scientifica, ogni mail, ogni messaggio su Whatsapp o Telegram, ogni call su Zoom, Meet o Skype, ogni volta che si fa una ricerca su Google o si utilizza il navigatore, ogni video pubblicato o visto su Youtube o Vimeo, ogni volta che giochiamo a Candy Crush o che apriamo una qualsiasi altra app che ci connette con un server esterno al nostro telefono, ogni azione che nello svolgersi coinvolge uno strumento digitale, genera un dato, che se elaborato può diventare informazione.

L'uomo produce dati e informazioni da decine di migliaia di anni, lo sono le pitture rupestri, i simboli cuneiformi sumeri, i geroglifici, le sculture totemiche, il sanscrito, l'alfabeto runico, greco e latino; dati che nel loro evolversi, nell'essere contestualizzati e aggregati, hanno generato informazioni, che sono esistite e che esistono grazie all'utilizzo di diverse tecnologie con le quali è stato possibile generare artefatti tangibili.

Il papiro, la pergamena, l'invenzione della carta, l'inchiostro e le penne, la stampa a caratteri mobili, le macchine da scrivere, sono solo alcune delle tecnologie che nel passato hanno consentito la conservazione di un dato in forma di informazione.

Ciò che sicuramente differenzia la produzione e lo scambio di informazioni tra il passato e il presente è che in passato l'informazione scambiata e registrata aveva una forma tangibile, si presentava a noi come elemento materico, esisteva come oggetto fisico.

Un contratto che attesta un matrimonio, un verbale di polizia, la sentenza di un tribunale, una lettera, una registrazione audio, un certificato di morte, un filmato, un regio decreto o una bolla papale, sono tutti documenti, artefatti che ci è possibile toccare.

L'esistenza fisica del dato e dell'informazione ne attestava il valore in quanto documento, nulla di formale esisteva al di fuori di un artefatto documentale fisico. L'essere materico era prova dell'esistenza stessa del documento, che nel suo essere tangibile diveniva parte dell'esperienza

umana in società, non solo per il contenuto, ma per il fatto stesso che esistesse nel nostro mondo dell'esperienza.

Il dato dunque per avere possibilità di divenire informazione ed essere conservato, doveva passare per un processo di iscrizione, che nell'era pre-digitale, mutava a seconda del medium con il quale si desiderava conservare e rendere fruibile l'informazione.

Non è mio interesse qui confrontarmi con il pensiero di MacLuahan rispetto al fatto che il medium sia il messaggio[136], ma è indubbio che non può esistere informazione se non vi è un medium a veicolarla e che ogni dato ha bisogno di essere iscritto per diventare informazione[137].

L'iscrizione di un dato produce un documento che interessa sempre l'esistenza di almeno due esseri umani, il documento era dunque la rappresentazione fisica di una interazione sociale, per dirla con Ferraris «nulla di sociale esiste al di fuori del testo»[138].

La trasformazione digitale che sta coinvolgendo le società contemporanee sta cambiando in modo profondo la nostra relazione con i documenti, la posta certificata PEC, la firma digitale e l'identità digitale consentono di generare e sottoscrivere documenti formali, senza che questi assumano necessariamente una forma fisica.

Il documento digitale si manifesta a noi come rappresentazione intellegibile di un codice binario, un linguaggio macchina elaborato da esseri umani e interpretato da macchine digitali grazie alla codifica binaria della tabella ASCII, American Standard Code for Information Interchange[139]

Il digitale ha uniformato i processi di iscrizione che possono assumere differenti livelli di complessità, dati dai linguaggi di programmazione con cui l'umano interagisce con la macchina, ma che sono ontologicamente costituti da un serie di uno e di zero.

Così un video digitale, una foto digitale, un suono digitale o un certificato di morte digitale sebbene si manifestino alla nostra realtà in modo sensibilmente differente sono tutti composti di bit.

Omologando in senso computazionale i sistemi di iscrizione, si ha la possibilità di generare e fruire dell'informazione in modo assolutamente innovativo. Dal momento che per le macchine digitali ogni dato viene registrato con lo stesso linguaggio non sarà un problema per molte di queste riprodurre documenti anche molto differenti. Un computer, uno smartphone o un tablet, possono generare e veicolare documenti che fin tanto che esistevano in senso materico necessitavano di tecnologie specifiche per essere generati e di medium differenti per essere fruiti.

Un video necessitava di cellulosa e di macchine da presa per essere generato e di un proiettore per essere fruito, un suono necessitava di un registratore magnetico per essere registrato e di un sistema

136 M. McLuhan, *gli strumenti del comunicare*, Il Saggiatore, Milano 2015
137 M. Feraris, *Documentalità. Perché è necessario lasciare tracce*, Editori Laterza, Bari-Roma, 2020, p. 201
138 M. Feraris, ivi p. 176
139 https://www.lingue.unich.it/sites/st07/files/04_-_codifica_informazioni.pdf ultima consultazione 22/07/2021

in grado di tradurre e amplificare i segnali magnetici per essere fruito (la puntina di un grammofono o di un giradischi, la testina di un lettore di cassette), un certificato di morte necessitava di carta e di inchiostro per essere realizzato e doveva essere tangibile per poter essere fruito.

Ad oggi con uno smartphone, per esempio, è possibile generare video, immagini, suoni e documenti e sempre utilizzando lo stesso strumento digitale possiamo fruire di ognuno di questi oggetti digitali, così come possiamo archiviarli o condividerli.

Ciò che era ontologicamente differente ad oggi è composto della stessa sostanza, bit.

Non solo, i programmi che consentono una iscrizione e fruizione dell'informazione, escludendo chiaramente l'hardware in cui sono inseriti, sono anche essi formati dalla stessa sostanza dei dati, sempre dei bit.

L'uniformità che determina la sostanza tramite la quale viene realizzato ogni forma di documento digitale, consente di immaginare dei sistemi di fruizione di tali documenti completamente innovativi.

Internet ci consente di accedere a zettabyte di informazioni digitali fortemente eterogenee ed è incredibile pensare che questa mole di informazioni stia nelle nostre tasche e che, se si è connessi alla rete le si possa consultare in qualsiasi momento, fruendo contemporaneamente di più documenti. Ci è possibile archiviarli come condividerli sia pubblicamente che privatamente e come già detto in precedenza ognuna di queste azioni genera a sua volta dei dati che vengono archiviati, aggregati condivisi e rivenduti da società di data mining e di big data.

Oggi fruiamo prevalentemente documenti e contenuti digitali, come abbiamo potuto vedere nel capitolo precedente anche le amministrazioni pubbliche si stanno impegnando in questa trasformazione della società, così che sarà sempre più raro doversi recare fisicamente in un ufficio per ottenere un documento.

Il fatto che la nostra relazione con i documenti sia mediata da strumenti digitali portabili ci consente di fruire di questi contenuti ovunque si voglia, per esempio mentre siamo seduti ad una scrivania o mentre siamo su un mezzo pubblico, non si è più costretti a dover accettare un dove e un quando predefiniti.

Come vedremo più avanti sempre in questo capitolo, il tema della presenza, intesa o come tele-presenza o come telepistemica[140] è centrale per comprendere come e quanto la digitalizzazione e la conseguente dematerializzazione dell'informazione possa e stia impattando sulle nostre vite.

140 L. Floridi, *Pensare l'infosfera, La filosofia come design concettuale,* Raffaello Cortina Editore, Milano, 2020 p. 83

2.2 Sovrapproduzione e metodo

La nostra relazione con l'informazione è profondamente cambiata, non solo ne generiamo moltissima, ma siamo anche abituati a fruirne in continuazione, su questo tema ci sono analisi, come il libro finalista al premio Pulitzer del 2001 di Nicholas Carr[141], o i lavori del neuroscienziato e psichiatra tedesco Manfred Spitzer[142] che pongono l'attenzione su quanto e se questa nostra costante connessione, possa trasformarsi in un problema sociale fino a quando non clinico. Nel mio lavoro non vi è interesse a definire se la rivoluzione digitale sia un bene o un male per la società, mi interessa provare a far emergere pro e contro dei processi di digitalizzazione, in relazione specificatamente al lavoro dello storico e alla produzione e alla conservazione di documenti.

Il Web 2.0 ci ha dato la possibilità di esprimere molto di più le nostre opinioni pubblicamente, ci scambiamo molte più informazioni di un tempo, interveniamo quotidianamente nel dibattito culturale e politico non solo del nostro paese e abbiamo accesso ad una base documentaria quasi sconfinata. Per non parlare poi di come il Partito Pirata prima e subito dopo il Movimento 5 Stelle abbiano provato a fare del Web il luogo della democrazia e del dibattito politico anche istituzionale[143]. Tutto ciò è possibile solo perché c'è un mezzo tecnologico che ce lo consente.

Senza avere nessuna ambizione di poter qui suggerire un processo metodologico che possa garantire una fruizione corretta dell'informazione digitale, possiamo azzardare che il metodo della ricerca storica possa essere utile per contribuire al dibattito, così come non possiamo escludere che la rivoluzione digitale possa suggerire dei nuovi o rivisti processi metodologici all'indagine storica.

A mio avviso il ruolo dello storico si manifesta in due diverse azioni, da una parte attraverso il suo lavoro di ricerca tramite il quale crea storiografia, così da consegnare al presente un punto di vista sul passato costruito attraverso l'analisi critica delle azioni, del pensiero, e delle arti di altri esseri umani che hanno vissuto in uno spazio tempo differente dal nostro ma che attraverso le loro azioni, il loro pensiero e le loro arti, sono stati artefici del nostro presente; e dall'altra la necessità di insegnare e diffondere il proprio metodo.

Nel processo di rielaborazione epistemica che sta coinvolgendo l'informazione, entrambe le azioni dello storico trovano un nuovo e diverso spazio di attuazione.

In base alla letteratura a disposizione riguardo i cambiamenti generati dalla rivoluzione digitale e soprattutto in relazione alla produzione condivisione e archiviazione delle fonti, possiamo penare che come storici si possa essere chiamati a chiedersi come organizzare i big data, accettando la sfida di non usarli solo in senso quantitativo, ma anche qualitativo; così come siamo coinvolti nel pensare

141 N. Carr, *Internet ci rende stupidi? Come la rete sta cambiando il nostro cervello,* Cortina Raffaello, Milano 2001
142 M. Spitzer, *Demenza digitale. Come la nuova tecnologia ci rende stupidi*, Corbaccio, Milano 2019
143 Ivi, p.

il funzionamento degli archivi digitali, soprattutto in relazione all'utilizzo dei metadati tramite i quali è possibile modificare in modo sensibile la possibilità di ricerca e interazione con la fonte; siamo altresì chiamati a riflettere sulle possibilità date dall'analisi delle fonti attraverso metodi mutuati da altre discipline, come per esempio il metodo di indagine della linguistica dei corpora, così da poter compiere analisi computazionali sulle parole presenti nelle informazioni che costituiscono le fonti.
Se queste riflessioni riguardano unicamente chi si occupa di storia, è di interesse pubblico chiedersi come ci si possa confrontare con la mole di informazioni con la quale interagiamo quotidianamente, e su questo sono convinto che le competenze di uno storico o di una storica siano di grande utilità sociale, essendo che il nostro lavoro si concentra sull'analisi delle fonti in senso critico.

2.3 Condivisione fisica e digitale

Uno dei più importanti cambiamenti dati dalla digitalizzazione e dallo sviluppo del Web 2.0 è l'accessibilità alle informazioni e la possibilità di condividerle.
Per comprendere la mutazione epistemica del concetto di accessibilità sarà necessario sviluppare un ragionamento più articolato che tenga in considerazione le teorie sulla presenza[144], ragionamento al quale arriveremo, ma ora mi voglio concentrare sul tema della condivisione.
L'informazione oggi non si esaurisce in se stessa attraverso la comunicazione fatta dal soggetto produttore con i possibili fruitori, ma grazie alla dematerializzazione[145], le informazioni prodotte vengono rilanciate e condivise da migliaia di persone in diverse parti del mondo, andando così a creare un effetto moltiplicatore.
Prima dello sviluppo del web per fruire una informazione si doveva necessariamente impossessarsi di un documento tangibile che la contenesse, si doveva dunque raggiungere un luogo fisico dove trovare il documento, mettiamo un'edicola, comprare ciò che si desiderava e solo a quel punto poter fare nostra l'informazione.
Vi sono delle eccezioni dove l'informazione in senso materico la si può fruire perché delle istituzioni quali possono essere un archivio o una biblioteca la mettono a disposizione dei loro fruitori. In questo caso l'informazione rimane legata ad un luogo fisico, dal quale non può uscire e dunque può essere fruita solo in uno spazio tempo determinato da un regolamento. Un'altra

144 International Society for Presence Resource https://ispr.info/about-presence-2/about-presence/ ultima consultazione 23/07/2021
145 Torneremo ad affrontare la complessità del termine dematerializzazione in relazione ai processi di digitalizzazione

eccezione può essere un bar, dove i proprietari decidono di mettere a disposizione dei loro clienti i giornali del giorno, ma anche in quel caso l'informazione rimane legata ad un luogo.

Medium come televisione e radio veicolano le informazioni in modo molto differente rispetto a ciò che avviene nel Web, pensiamo alle informazioni di un telegiornale, o le si fruisce nel momento scelto dall'editore oppure non è più possibile fruirne, così come per conservarle possiamo al massimo decidere di videoregistrare le notizie; lo stesso discorso vale per le informazioni veicolate attraverso la radio. Il broadcasting permette lo sviluppo di una comunicazione di massa, ma è solo l'ente produttore o conservatore che può decidere di condividere l'informazione, salvo appunto alcuni escamotage che possono consentire una condivisone delle informazioni, ma in modo direi molto poco incisivo.

Prima dell'avvento del Web, per esempio, una volta esaurito l'interesse per l'informazione che in qualche modo siamo riusciti a fare nostra come oggetto tangibile, e deciso di non archiviarla, sarebbe stato possibile condividerla con altri utenti interessati prevalentemente grazie ad una compresenza fisica di almeno due persone. Una eccezione è data dal sistema delle Poste, che consentono di condividere informazioni senza che vi sia una compresenza fisica.

Prima del Web 2.0 era possibile trovare le informazioni online, prevalentemente su dei siti, in questo caso, per accedere all'informazione non eravamo costretti ad uscire di casa, non ci saremmo dovuti relazionare con altri esseri umani o accettare i tempi dettati da un editore di un media broadcasting, ma dalla nostra postazione computer avremmo navigato la rete fino a trovare il sito di interesse, cercare all'interno l'informazione e a quel punto fruirla.

Questo sistema non ci legava ad uno spazio tempo prestabilito, in qualsiasi momento del giorno e della notte e da qualsiasi parte del mondo, se in possesso di una connessione internet potevamo accedere all'informazione. Una volta trovata, potevamo "appropriarci" dell'informazione salvando la pagina web o se possibile scaricando il documento digitale.

La condivisione poteva avvenire o tramite il servizio di posta elettronica, che prima dell'avvento di Gmail nel 2004, aveva molti limiti nella capacità di archiviazione e ancora oggi non consente di condividere allegati troppo pesanti (salvo reindirizzarci all'utilizzo di un sistema in cloud come Drive), o solo tramite un supporto fisico, un floppy disc, un cd, un hard disk esterno o una pennina usb, ciò implicava che nonostante l'informazione fosse digitale, quindi intangibile, per poterla condividere era necessario condividere il supporto fisico che la conteneva.

Rispetto all'era pre-digitale che abbiamo prima descritto, la condivisione dell'informazione digitale attraverso la condivisione di un supporto fisico, difficilmente poteva avvenire se non tramite una compresenza degli interessati.

Nel 1999 grazie a Napster, cambia in modo sensibile il modo di condividere le informazioni, il programma utilizzava un protocollo di rete peer-to-peer centralizzato, che si basava su un sistema di server centrali che mantenevano la lista dei sistemi connessi e dei file condivisi, mentre le transizioni avvenivano direttamente tra i vari utenti.

In poche parole se si utilizzava Napster come client, la lista dei dati che mettevamo in condivisione veniva acquisita da uno o più server centrali che contemporaneamente ti mettevano a disposizione la lista dei contenuti degli altri client, così facendo una volta individuata l'informazione di cui avevamo bisogno la ottenevamo direttamente da chi già la possedeva.

Questo sistema di condivisione si basava sulla costruzione di una comunità, su un rapporto paritario tra i client e i server, ma apriva una questione complessa da un punto di vista legale in rapporto al diritto d'autore[146], in particolar modo perché Napster dava la possibilità di condividere tra gli utenti Mp3, dunque brani musicali protetti da copyright che nel loro perdere struttura fisica, sfuggivano *de facto* al controllo dell'industria musicale.

Oggi dati e informazioni nascono prevalentemente in forma digitale e in seguito allo sviluppo del Web 2.0 le informazioni vengono veicolate soprattutto attraverso i social network, piazze digitali dove informazioni provenienti da altre piattaforme e siti vengono condivise tra gli utenti.

Con lo sviluppo del Web e la conseguente digitalizzazione del dato è mutato il modo che si ha di condividere le informazioni, queste sono diventate parte integrante della nostra esistenza e nonostante i documenti abbiano perso la loro natura fisica o forse proprio perché l'hanno persa, oggi ognuno di noi, se interessato, può avere un rapporto molto più diretto con il documento, lo può infatti fruire nella sua originalità senza aver bisogno di intermediazioni umane.

Prima dell'avvento del digitale un disegno di legge del Governo Italiano era fruibile prevalentemente da addetti ai lavori come i Senatori o i Deputati che ci stavano lavorando o i giornalisti che si occupano di politica, difficilmente comunque il testo nella sua interezza diventava un documento pubblico non in senso giuridico, ma in senso sociale. Quindi una complessità informativa di pubblico dominio.

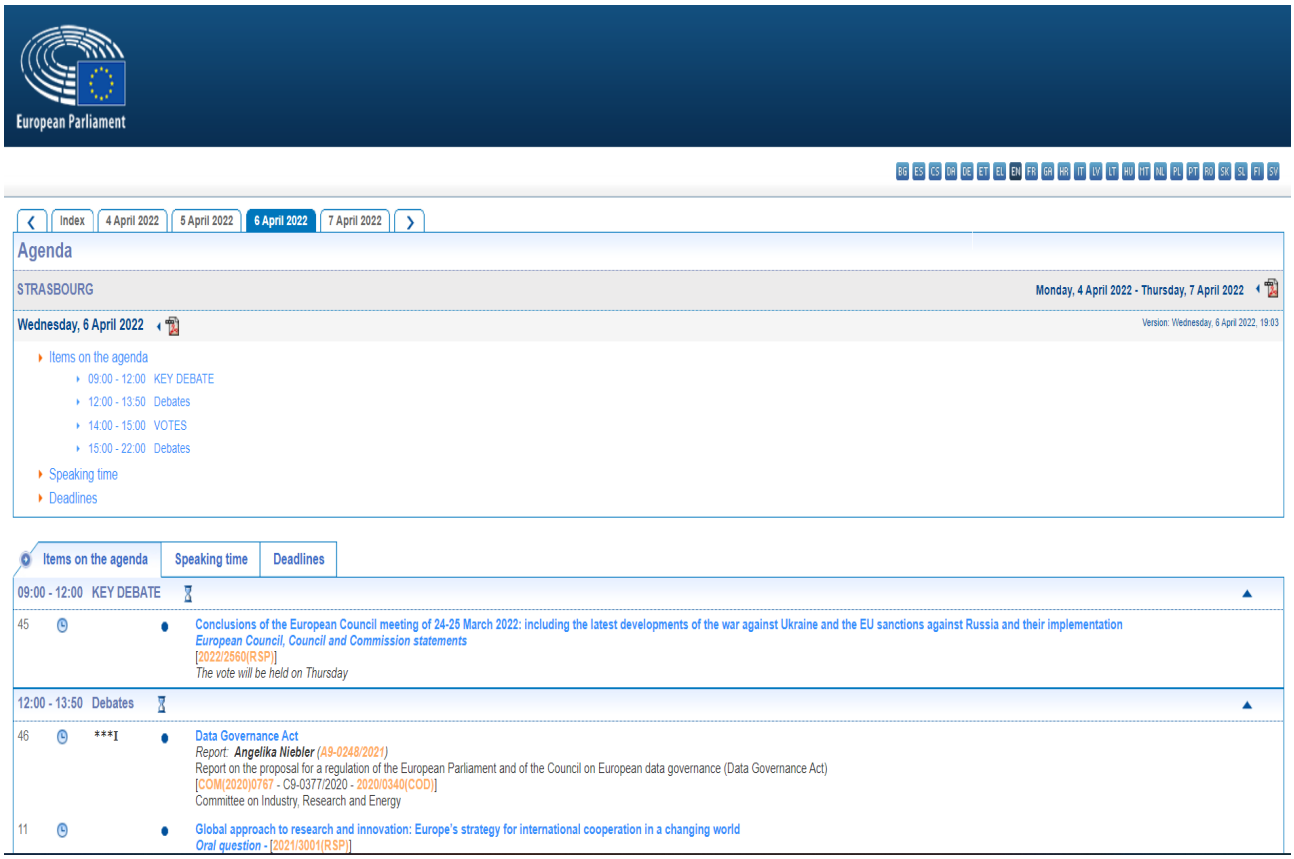
I dati e le informazioni presenti nel documento erano osservabili nella loro forma originale, in quello che dopo proverò a definire come medium-documento, solo da un ristretto numero di persone che a loro volta potevano riportare delle opinioni a riguardo a mezzo stampa. I cittadini così non si confrontavano quasi mai con il testo nella sua interezza e le informazioni arrivavano mutuate dalle idee di politici e giornalisti, la comunicazione si svolgeva dunque in broadcasting, da uno a tanti e una volta espressa sui giornali, alla radio o in televisione, contribuiva a creare opinione pubblica, ma difficilmente generava altra informazione.

146 https://it.wikipedia.org/wiki/Napster ultima consultazione 27/07/2021

Oggi, un disegno di legge in discussione al Senato, dunque già approvato alla Camera, può essere liberamente fruito in versione digitale da qualsiasi cittadino, che può scaricare il PDF o leggere il testo online. Non solo, il documento, nella sua forma intangibile, può essere condiviso all'interno di ambienti informali, ma nella sua assoluta formalità e interezza, può dunque essere postato su Facebook e consentire di creare un confronto tra cittadini e cittadine non tanto basato sull'opinione espressa da terzi, giornalisti o politici, ma a partire dal documento originale.

Questo fa sì che l'informazione presente e non modificata all'interno del testo originale possa diffondersi in modo profondo e a sua volta generare una serie di reazioni positive o negative, reazioni che nell'ambito digitale, si manifestano sotto la forma di dati e di informazioni che a loro volta verranno condivise, fruite, discusse criticate, rilanciate.

Il portale del Parlamento Europeo è un ottimo esempio per comprendere come sia mutata l'accessibilità ai documenti. Se su un qualsiasi motore di ricerca scriviamo le parole "eu parliament russian aggression", il sistema ci restituirà una serie di link che ci indirizzeranno all'interno del portale del Parlamento Europeo, dove potremmo accedere ad ogni risoluzione, votazione o dibattito in relazione all'invasione Russa dell'Ucraina. Ogni documento è fruibile in ventiquattro lingue e oltre ad essere legato ad altri documenti attraverso un'intricata rete di collegamenti ipertestuali, può essere liberamente scaricato da qualsiasi utente e dunque archiviato o condiviso[147].



147 https://www.europarl.europa.eu/doceo/document//OJQ-9-2022-04-06_EN.html#TF7 ultima consultazione 19/07/2022

Il Web per come si è strutturato con l'avvento del 2.0 non si limita a rendere possibile uno scambio informazionale unidirezionale, ma al contrario, un po' come funzionava Napster, ha dato forma ad un sistema che spinge a creare una comunità all'interno della quale i contenuti si spostano da un utente all'altro.

La nostra relazione con le informazioni digitali è diventata intima e caotica, gli smartphone oggetti che per eccellenza utilizziamo nel fruire e condividere informazioni, sono, con le chiavi e il portafoglio, uno di quegli oggetti senza i quali non si può uscire di casa.

All'interno del Web 2.0 la condivisione dell'informazione avviene in modo completamente diverso da come era possibile quando le informazioni esistevano in senso materico o quando il Web non era social. Oggi per condividere una informazione basta essere iscritti ad un qualsiasi social network o programma di messaggistica, dunque avere un'identità digitale informale che esiste solo in seguito ad una azione di iscrizione che è determinata da uno scambio di informazioni tra noi e la piattaforma che ci ospita[148]. Una volta entrati in una di queste comunità per condividere informazioni di cui siamo o meno i soggetti produttori, ci basterà premere un tasto. L'azione di condivisione avverrà senza il bisogno di una compresenza fisica o temporale con altri esseri umani, l'informazione rimarrà fruibile a tempo indeterminato e potrà essere fruita contemporaneamente da un numero anche molto elevato di persone, perché ciò possa avvenire, oltre alla necessità di iscriversi ad una comunità ognuno di noi è vincolato all'utilizzo di un device digitale, deve esistere digitalmente.

La tecnologia digitale ha indubbiamente aumentato le nostre possibilità di informarci, questo non necessariamente ci rende più intelligenti o sapienti, sicuramente non siamo in grado di sfruttare al meglio gli strumenti che abbiamo, ma non si può sostenere che oggi la società sia meno informata di dieci anni fa, possiamo dire che ci sono molte informazioni imprecise, alcune del tutto false e che in tanti vi credono in modo acritico, ma questo non toglie che oggi si abbia una consapevolezza maggiore, magari fuorviata, dei fatti che accadono nel mondo.

2.4 Condivisione e Infodemia

Secondo gli studi svolti da Walter Quattrociocchi[149] e dal suo team di ricerca, nella eccessiva quantità di informazione prodotta e scambiata, risiede il virus della disinformazione, o meglio della misinformation, questo però non ci deve spingere solo a chiederci come limitare la quantità di

148 M. Ferraris, *Scienza Nuova. Ontologia della trasformazione digitale,* Rosenberg e Seller, Torino 2018, p. 33
149 M. Cinelli, W. Quattrociocchi, A. Galeazzi, *et al. The COVID-19 social media infodemic. Sci Rep* 10, 16598 (2020)

informazione prodotta e condivisa, che inevitabilmente porta il dibattito verso il complesso tema della libertà di espressione e della censura, sto parlando dei tentativi di limitare le fake news, ma ci dovrebbe stimolare a proporre delle soluzioni di metodo che aiutino a gestirla.

Come ha affermato anche Chris Wetherell, sviluppatore dell'aggregatore Google Reader che nel 2009 ha sviluppato il tasto retweet per Twitter, la possibilità di condividere in modo acritico e immediato il pensiero o le informazioni di altri, è ciò che maggiormente ha reso i social network un ambiente che in alcune fasi può essere definito tossico[150]. Quando Wetherell creò il tasto retweet era convito di aver fornito uno strumento importante per tutte le comunità che avevano meno voce, poter rilanciare istantaneamente dei contenuti aveva il poter di moltiplicare il flusso informativo e conseguentemente creare dei presupposti per aumentare l'interesse sociale su questioni che altrimenti avrebbero avuto poco risalto, oggi è lo stesso creatore che si rende conto di come questo strumento sia stato utilizzato in un modo disfunzionale.

L'algoritmo di feed è quello che seleziona ciò che effettivamente visualizziamo rispetto all'intera offerta dei contenuti della nostra rete sociale. Facebook, in particolare, funziona secondo un algoritmo chiamato *EdgeRank* che tiene conto delle similarità delle interazioni che un utente ha con i suoi amici e stabilisce delle "corsie preferenziali" per i contenuti postati dagli utenti con i quali si interagisce di più. Questo processo, unito alla tendenza degli utenti a cercare notizie che confermino e che non mettano in discussione le loro idee, ha generato ciò che è conosciuto come eco chamber. Quello che in modo improprio alle volte definiamo come "la mia bolla social" comporta una sorta di autosegregazione informazionale dove si è in contatto con utenti che hanno posizioni simili da un punto di vista politico sociale e culturale e i contenuti condivisi tendono a confermare le posizioni dominanti e vengono rilanciati continuamente dagli utenti, così da dare la sensazione che determinate informazioni acquistino sempre più significato e risonanza[151].

La sovrapproduzione di informazioni in un sistema massificato di diffusione, ha reso il web dell'autorità condivisa, un luogo molto complesso all'interno del quale muoversi. Spesso ci troviamo immersi in un ambito saturato da informazioni, documenti e opinioni, luoghi digitali dove è difficile orientarsi, dove *episteme* e *doxa* non si affermano più in una differenza, ma quasi si equivalgono.

Come riuscire a non essere sopraffatti quantitativamente e contemporaneamente a essere in grado di gestire le informazioni che ci arrivano così che ci possano aiutare a essere informati?

150 https://www.buzzfeednews.com/article/alexkantrowitz/how-the-retweet-ruined-the-internet ultima consultazione 23/07/2021

151 M. Cinelli, G.De Francisci Morales, A. Galeazzi, W. Quattrociocchi, M. Starnini, *The echo chamber effect on social media*, PNAS 2021

In una intervista rilasciata a «Cfnews» dell'Università Ca' Foscari di Venezia, Quatrociocchi suggerisce che «Servirebbero programmi di formazione nelle scuole per docenti e studenti su scala nazionale mirati a comprendere come ci comportiamo quando siamo online. Su questo stiamo lavorando con Facebook sviluppando un programma rivolto ai giovani»[152].

Eppure il nostro ragionamento è partito dall'osservare come le tecnologie digitali e la dematerializzazione degli oggetti tangibili, comporti la possibilità di accedere e condividere in rete documenti originali e ufficiali, processo che consente ad ognuno, ora come mai nella storia, di potersi confrontare con delle fonti primarie, dunque con una informazione meno contestualizzata ma più "pura".

Dando seguito al pensiero di Quatrociocchi, è lecito pensare che anche noi come storiche e storici si possa avere un ruolo per provare a ridurre i rischi della misinformation; ciò che possiamo presupporre è che, se si mettesse all'interno della didattica della storia, anche l'insegnamento del metodo storico di indagine, critica e comparazione delle fonti e non solo delle nozioni della storia, si potrebbe contribuire a far sviluppare in studenti e studentesse uno spirito critico utile ad osservare le fonti in modo più accurato e meno passivo.

Quando si vuole osservare un evento, in quanto storiche e storici, il nostro primo compito è di cercare delle fonti/informazioni il più attendibili possibile, o comunque che possano essere inserite in un contesto critico di analisi. La conoscenza arriva solo dopo una attenta analisi delle fonti. Come sappiamo questo implica che la storica per fare il suo lavoro si impegna a fare ricerca sulle singole fonti così da validarne il valore informazionale. Questo stesso processo può essere replicato per la fruizione delle informazioni online anche da chi non è uno storico.

Il metodo di indagine storico richiede una costante analisi critica delle fonti, analisi che avviene attraverso la comparazione e lo studio della fonte stessa. Quando ci troviamo davanti ad un documento siamo chiamati a porci delle domande: da chi è stato prodotto il documento? In quale archivio o fondo si trova? Ci sono altri documenti che possono confutare o confermare? E per dare risposta ad ognuna di queste domande sviluppiamo delle ricerche ad hoc, così da confermare o confutare la validità del documento stesso.

Per attualizzare questo processo metodologico all'interno dello scambio informazionale digitale in prima istanza non possiamo banalizzare il problema della sovrabbondanza.

Con la presenza di tante informazioni è facile che due utenti con posizioni opposte possano trovare differenti fonti che confermino le posizioni di entrambi e ciò può minare il sistema di affidabilità delle fonti, portandoci nel paradosso che non si sa più di chi fidarsi. Ma anche per scongiurare un

152 https://www.unive.it/pag/14024/?tx_news_pi1%5Bnews%5D=8464&cHash=8fb0ead51f3e55bdbbe155b4779281bb ultima consultazione 23/07/2021

processo che porti al relativismo assoluto dove l'opinione di chiunque ha lo stesso valore dell'affermazione di un esperto riconosciuto che si esprime solo in seguito ad un'analisi realizzata in modo scientifico, è importante promuovere un metodo di critica della fonte, che in prima istanza sottolinei l'importanza di riconoscere il soggetto produttore così da poterne considerare l'attendibilità.
Con questo non sto suggerendo che tutto ciò che leggiamo online debba passare al vaglio di una analisi di ricerca critica, ma se insegnassimo già nelle scuole che le fonti possono essere criticate, soprattutto se si ha intenzione di condividerle, potremmo aiutare a diffondere un metodo utile per riuscire a costruire pensiero critico.
Come riporta Hilda Girardet citando Paola Di Ciori

> Abituare gli studenti all'idea che la identificazione e la contestualizzazione (chiedersi di fronte a qualsiasi tipo di materiale – un documento di archivio, un manuale, una notizia di telegiornale, una pubblicità, un rito – da chi, quando e dove è stato prodotto, in che modo è giunto a noi, per quale destinazione e pubblico) sono esigenze basilari della conoscenza in generale e di quella storica in particolare.[153]

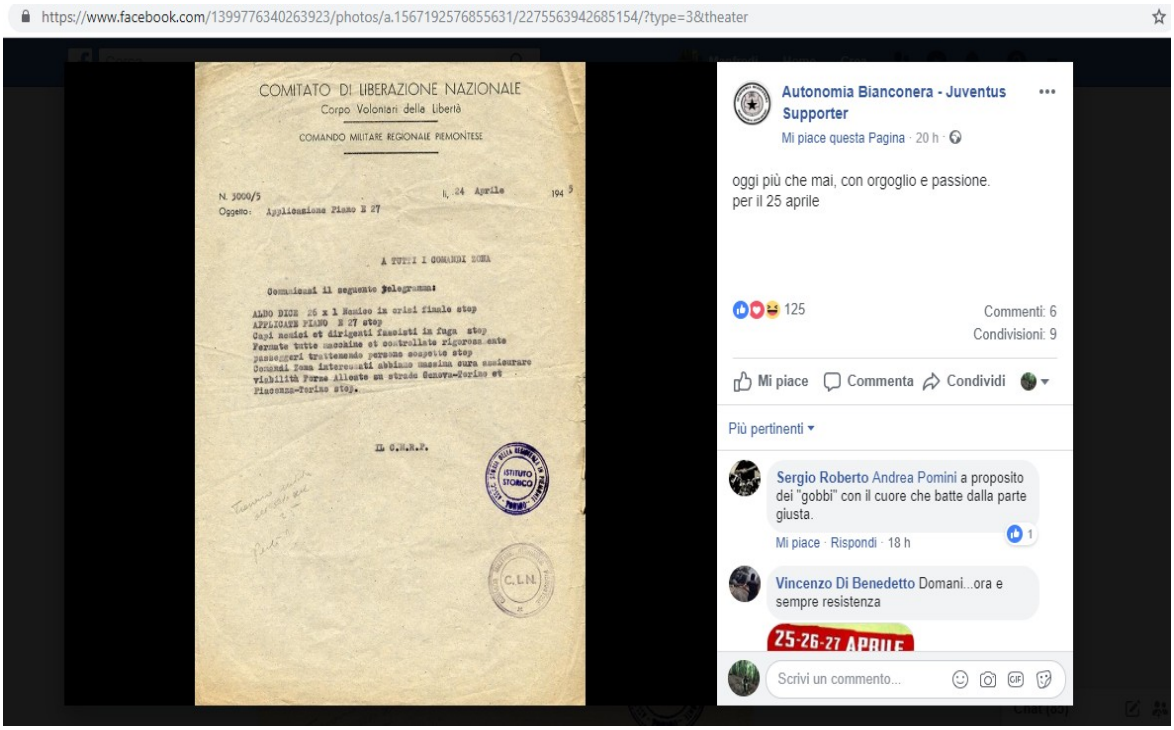
Questo fu l'approccio che guidò l'Istituto Storico di Modena e Future Education Modena, nello sviluppare un progetto che vide il coinvolgimento di tre diversi istituti scolastici della città. Durante il progetto si andò a lavorare sull'analisi delle fonti a partire dalle informazioni che studenti e studentesse fruivano attraverso i social network per poi arrivare ad analizzare delle fonti primarie relative al fascismo. La prima fase di elaborazione vide gli studenti e le studentesse delle classi coinvolte condividere una serie di contenuti molto eterogenei tra di loro: pubblicità, video di influencer, post di politici, di organizzazioni politiche o di semplici cittadini, per ogni informazione condivisa con la classe dovevano individuare e analizzare la provenienza dell'informazione, il target a cui era riferita, il significato implicito oltre a quello esplicito e se vi erano o meno forme di comunicazione assertiva. Questo passaggio che ha visto come unici protagonisti studenti e studentesse è stato propedeutico per riuscire a sviluppare al meglio l'analisi delle fonti storiche fornite dall'Istituto Storico di Modena.
Se l'output del progetto ha visto la realizzazione di un prodotto multimediale, l'outcome lo si è raggiunto con la condivisione e l'applicazione pratica di un metodo di analisi dell'informazione utile a studenti e studentesse, non tanto ad attuare un processo di debunking – utilizzato

153 H. Girardet, *Vedere, toccare, ascoltare. L'insegnamento della storia attraverso le fonti*, Roma, Carocci, 2015, pp. 21, 22

prevalentemente per la confutazione di notizie o affermazioni false o antiscientifiche -, quanto per migliorare la loro capacità di analisi critica mentre fruiscono informazioni online.[154]

Ancora non si hanno dati certi per poter definire se una così alta condivisione di informazioni nella rete possa contribuire o meno a generare conoscenza, così come non è del tutto chiaro se le eco chamber contribuiscano a orientare il pensiero degli utenti al loro interno, tema centrale per capire come affrontare la misinformation, ma se questo dibattito è ancora in una fase iniziale e dunque non è possibile trarre conclusioni, in quanto storici non possiamo non chiederci cosa succede quando navigando su Facebook ci troviamo davanti ad un post di questo tipo:

Tabella 2.4.1



Il 25 aprile 2019, un giornalista e scrittore che ho tra i miei contatti, tanto juventino quanto fermamente antifascista, decise di celebrare il giorno della liberazione, condividendo sulla sua pagina Facebook il post del gruppo legato alla Juventus“Autonomia Bianconera – Juventus supporter”[155].

154 Per approfondimenti rispetto al progetto rimando a M. Scanagatta, *Archivi e fonti digitali per la didattica della storia*, «Rivista di Ricerca e Didattica Digitale» Anno I – 1/2021 https://studiumeditore.it/riviste/rivista-di-ricerca-e-didattica-digitale/rrdd-1-2021/ ultima consultazione 28/07/2022

155 https://www.facebook.com/Autonomia-Bianconera-Juventus-Supporter-1399776340263923/ ultima consultazione 28/07/2021

Per gli amministratori della pagina le intenzioni del post sono duplici, da un parte richiamare l'attenzione sul 25 aprile come data fondamentale nella storia italiana, dall'altra denotare in modo chiaro il posizionamento della pagina stessa rispetto al fascismo, marcando dunque il dato che lo juventino vicino all'autonomia bianconera è uno juventino antifascista.

Il testo del post è molto sintetico e diretto: «Oggi più che mai, con orgoglio e passione per il 25 aprile»

Ciò che attirò la mia attenzione però fu l'immagine allegata al post, una rappresentazione simulacrale di un documento materico che ha subito un processo di dematerializzazione e rimediazione e che viene utilizzato come strumento narrativo, all'interno di un social network generalista e con fini ricreativi.

Il documento scelto non è un documento qualsiasi, è infatti l'ordine dato dal Comitato Regionale Piemontese (CMRP), inviato il 24 aprile del 1945 via telegramma a tutti i comandi di zona per rendere operativo il piano E 27 (Emergenza 27). Un piano insurrezionale che dal 1944 veniva elaborato dal CLN, per organizzare l'insurrezione nel nord Italia contro l'occupazione nazifascista.

Come scrive Claudio Dellavalle, l'attuazione del Piano E 27 è di particolare interesse perché mette al centro dell'insurrezione sia l'azione dei partigiani che della cittadinanza antifascista. Dal 18 aprile fu organizzato uno sciopero di massa, gli operai bloccarono la produzione e uscirono dagli stabilimenti. Soltanto alla Grandi Motori, alla Mirafiori e alle Fonderie Ghisa le milizie fasciste impedirono l'uscita ma non la sospensione del lavoro. Le scuole, di ogni ordine e grado, compresa l'Università, vennero chiuse e l'ordine del provveditore agli studi di riprendere le lezioni fu ignorato. I tram si fermarono e i tentativi dei fascisti di ripristinarne il funzionamento non raggiunsero gli scopi sperati.

I giorni seguenti furono carichi di tensione, non solo per l'avvicinamento delle colonne tedesche verso il capoluogo, due divisioni completamente efficienti, dotate di mezzi pesanti, accompagnate da truppe della Rsi, in tutto circa 75.000 uomini al comando del generale Schlemmer, ma anche per l'atteggiamento del comandante della missione alleata in Piemonte, colonnello Stevens che, nell'intenzione di ritardare, se non addirittura impedire un'insurrezione difficile da controllare, sottolineva i gravi rischi derivanti da un'azione anticipata e priva del necessario supporto alleato.

L'insurrezione ha due soggetti distinti e due fasi: gli operai e i partigiani; la struttura paramilitare costituita dalle Sap che dà l'avvio all'insurrezione e regge soprattutto alla periferia della città l'urto con tedeschi e fascisti tra il 25 e il 27 aprile e le formazioni partigiane foranee che entrarono in città il 27, con una vistosa sfasatura rispetto al piano previsto dal CLN[156].

156 C. Dellavalle, *Economia e industria nella guerra. Le fonti e gli archivi in Piemonte*, Franco Angeli, 1987 p. 192

Il CLN in un primo momento decise di attendere lo sviluppo di una operazione congiunta con gli alleati che stavano risalendo, ma per il CMRP era importante che l'insurrezione avvenisse con il coinvolgimento diretto delle formazioni resistenziali cittadine e di quelle foranee.
Il piano risponde sia ad esigenze militari sia politiche, perché il problema della liberazione del territorio è inscindibile da quello del suo successivo controllo, è necessario dunque sostenere la battaglia contro i nazifascisti, ma anche superare le diffidenze inglesi e accelerare i tempi dell'insurrezione. Non si possono aspettare le armate angloamericane, col rischio di veder liquidato il movimento resistenziale. Non secondaria è l'esigenza da parte del CLNRP di prendere il comando temporaneo, così da poter affermare anche un'azione politica.
Il gruppo Facebook che ha condiviso questo documento è stato fondato in memoria di uno dei primi gruppi di tifoseria organizzata della Juventus, nato a metà degli anni settanta, che affermava delle posizioni politiche di estrema sinistra[157].
La stretta correlazione tra il gruppo organizzato e il pensiero antifascista ci fanno meglio comprendere come mai abbiano deciso di condividere proprio il comando E27 per celebrare il 25 aprile, non di meno il fatto che la Juventus sia una delle due squadre del capoluogo piemontese rende ancora più carico di significato il raccontare come avvenne la liberazione del Piemonte, ma la cosa davvero interessante è che appunto abbiano deciso di raccontarla non con un brano estratto da un libro di storia sulla liberazione di Torino, ma condividendo il documento stesso che diede l'ordine di iniziare l'insurrezione.
Possiamo immaginare che la maggior parte delle 125 persone che a 20 ore dalla pubblicazione del post avevano messo un "mi piace" non fossero storici, come non lo fossero tutti e 9 gli utenti che a loro volta hanno deciso di condividere il post e dunque il documento.
Questo ci deve far riflettere su quelle che sono le reali implicazioni della digitalizzazione del dato e dell'informazione e di come questo processo possa mostrarsi rivoluzionario nella condivisione e analisi, già oggi, di contenuti storici attraverso la fruizione del documento digitalizzato.
Ma procediamo con ordine cercando di capire in che modo un post su Facebook possa contribuire all'approfondimento della storia attraverso un processo di apparente disintermediazione, ma che in realtà si verifica grazie ad una neo intermediazione digitale che non avviene nel momento della fruizione del documento, ma la precede.

Nel documento troviamo queste informazioni:

COMITATI DI LIBERAZIONE NAZIONALE

157 https://it.wikipedia.org/wiki/Tifoseria_della_Juventus_Football_Club ultima consultazione 28/07/2021

Corpo volontari della libertà

COMITATO MILITARE REGIONALE PIEMONTESE

N. 300/5
Oggetto: Applicazione Piano E27

li, 24 aprile 1945

A tutti i comandi di zona
comunicasi il seguente telegramma

«Aldo dice 26 x 1. Nemico in crisi finale. Applicate Piano E 27. Capi nemici et dirigenti fascisti in fuga. Fermate tutte le macchine et controllate rigorosamente passeggeri trattenendo persone sospette. Comandi Zona interessati abbiano massima cura assicurare viabilità forze alleate su strada Genova-Torino et Piacenza-Torino».

Prendiamo il caso che un utente che non conosce nulla sulla liberazione di Torino, rimanga colpito dal post perché condiviso da una pagina Facebook attinente al calcio di cui è molto tifoso.
Quello che può avvenire, se l'utente dimostra un minimo di curiosità è che per capire meglio il contenuto condiviso cerchi su Google le parole che troviamo nel documento.
Come è noto Google, come tutti i motori di ricerca presenti sul Web, si basa su tre diverse componenti. La prima è costituita da appositi "robot" o "spider" che, ciclicamente, esplorano la rete procedendo alla "cattura" e alla memorizzazione delle pagine Web o di loro elementi o parti significative. La seconda è costituita da procedure automatiche di estrazione dei termini presenti nelle pagine "catturate", che confluiscono in apposite basi di dati, grazie alle quali le parole indicizzate vengono messe in relazione con i "documenti" da cui sono state estratte. La terza è il motore di ricerca vero e proprio, cioè un software che permette di interrogare la banca dati dei termini indicizzati e che restituisce, sulla base di criteri predefiniti, i "documenti" che contengono il termine ricercato. Dal punto di vista dei loro essenziali meccanismi di funzionamento, quindi, i motori di ricerca appartengono alla famiglia dei sistemi di *information retrieval*, cioè di sistemi deputati al reperimento efficiente di informazioni e documenti.[158]

158 S. Vitali, *Ordine e caos: Google e l'arte della memoria*, in A. Spaziante, a cura di, *Il futuro della memoria: la trasmissione del patrimonio culturale nell'era digitale*, CSI Piemonte, Torino, p. 81

Google è un motore di ricerca completamente automatizzato che utilizza software chiamati *web crawler* per esplorare regolarmente il Web e trovare siti da aggiungere al suo indice. La maggior parte dei siti riportati nei risultati non viene inviata manualmente per l'inclusione, ma viene trovata e aggiunta automaticamente quando web crawler eseguono la scansione del Web.
La Ricerca Google prevede essenzialmente tre fasi:

- Scansione: Google usa programmi automatizzati chiamati *crawler* per cercare sul Web pagine nuove o aggiornate. Google memorizza gli indirizzi di tali pagine (o *URL delle pagine*) in un vasto elenco da esaminare in un secondo momento. Troviamo le pagine con molti metodi diversi, ma quello principale è seguire i link da pagine che già conosciamo.
- Indicizzazione: Google visita le pagine che ha trovato tramite la scansione e cerca di analizzare il contenuto di ogni pagina. Analizza i contenuti, le immagini e i file video per cercare di capire l'argomento della pagina. Queste informazioni vengono archiviate nell'*Indice Google*, un enorme database archiviato su molti server.
- Pubblicazione dei risultati di ricerca: Quando un utente esegue una ricerca su Google, Google cerca di determinare i risultati di qualità ottimale. I risultati "migliori" vengono stabiliti in base a molti fattori, ad esempio la posizione, la lingua, il dispositivo (desktop o telefono) e le query precedenti dell'utente. Ad esempio, la ricerca di "officine per la riparazione di biciclette" mostra risultati diversi a un utente di Parigi rispetto a un utente di Hong Kong[159].

La ricerca avviene quindi grazie a un sistema di intelligenza artificiale capace di incrociare i dati derivati dalla richiesta dell'utente con un database che possiamo immaginare come una biblioteca, ma che contiene più informazioni di tutte le biblioteche del mondo messe insieme.[160]
I caratteri dello strumento informatico permettono di moltiplicare le chiavi di accesso e di svincolarsi da una ricerca che rimane tutto sommato monodimesionale e sequenziale come quella in biblioteca. In un sistema di *information retrieval*, tipicamente, «l'utente specifica i propri bisogni informativi, indicando alcune parole chiave e il sistema recupera l'insieme dei documenti che si avvicinano alla domanda dell'utente»[161], cioè i documenti che, direttamente o tramite le loro rappresentazioni, sono associati nell'indice alle parole ricercate.[162]
Possiamo dire che Google nel momento dell'indicizzazione avvia un processo automatico di estrazione di dati che poi vengono attribuiti, in qualità di metadati, alla risorsa digitale, rendendo

159 https://developers.google.com/search/docs/basics/how-search-works?hl=it ultima consultazione 30/07/2021
160 https://www.google.com/search/howsearchworks/?hl=it ultima consultazione 30/07/2021
161 R. Baeza-Yates e B. Ribeiro-Neto, *Modern infomation retrieval*, Addison Wesley, Reading (US) 1999, p. 12
162 S. Vitali, *Ordine e caos: Google e l'arte della memoria*, in A. Spaziante, a cura di, *Il futuro della memoria: la trasmissione del patrimonio culturale nell'era digitale*, CSI Piemonte, Torino, p. 83

possibile rintracciare il link in seguito all'utilizzo, da parte dell'utente, di parole chiave corrispondenti a quelle selezionate dal sistema.

Fatte queste dovute premesse proviamo ora ad avanzare con un piccolo esperimento[163] e immaginiamo di essere l'utente che incuriosito dal post di Facebook voglia approfondire la sua ricerca sul documento che è stato condiviso, e immettiamo sulla barra di ricerca di Google le prime parole che troviamo all'interno del documento: "comitato di liberazione nazionale".

La ricerca porterà come primo risultato la pagina Wikipedia, che come osserva Deborah Paci è una fonte secondaria molto utile per poter accedere a una conoscenza storica di base, ma è anche complessa nel suo funzionamento[164], poiché la cocreazione di contenuti su cui si basa il sistema implica che ci sia una costante revisione dei contenuti stessi portata avanti non necessariamente da degli esperti ma da qualsiasi tipo di utente: questo garantisce una pluralità di interventi, ma la scientificità dei contenuti può essere messa in discussione, anche se il testo discretamente articolato è contrappuntato da note e citazioni, e per ogni voce di Wikipedia scorrendo la pagina verso il fondo viene riportata una bibliografia[165].

Il secondo risultato ci indirizza verso la pagina del'Associazione Nazionale Partigiani d'Italia che nel testo, se pur breve, riporta alcuni riferimenti bibliografici.

Ma ecco che al quarto risultato, subito dopo la Treccani, Google suggerisce di visitare il sito del SIUSA, il Sistemo Informativo Unificato per le Sopraintendenze Archivistiche[166], e rispetto ai risultati che lo precedono questo si distingue in modo deciso; offre infatti all'utente, che ricordiamo è partito da un post su Facebook, la possibilità di accedere all'interno di un sistema archivistico complesso che oltre a fornire alcuni testi di riferimento, consente di approfondire la storia del CLN attraverso dei link che rimandano a ulteriori schede di tutte le sedi del CLN attive sul territorio Italiano.

---

163 Per procedere in questo piccolo esperimento empirico, considerando che l'algoritmo di Google indicizza i risultati della ricerca in base alle interazioni precedenti dell'utente con il motore di ricerca, ho compiuto la stessa ricerca sia accedendo a Google dal mio pc sia utilizzando il pc di un'altra persona che ha installato *Brave*, un browser differente da Google crome e che utilizza come motore di ricerca primario Duck Duck. I risultati ottenuti non sono identici ma molto simili. Alla ricerca delle parole "comitato nazionale di liberazione" il computer dell'altro utente mi da come primo risultato sempre la pagina Wikipedia e mi propone l'accesso al SIUSA come sesto risultato. Per la ricerca delle parole "piano e27" il primo risultato è sempre militarystory e mi propone di accedere ad ARCHOS come quinto risultato, ma proponendomi come terzo e quarto risultato di accedere ad Amazon per comprare una lampada che si chiama e27.

164 D. Paci, *Conoscere è partecipare: digital public history, wiki e citizen humanities*, Umanistica Digitale, in corso di pubblicazione

165 D. Paci, *La digital public history e Wikipedia: tra partecipazione e conflittaualità,* in, S. Noiret, M. Scanagatta, D. Paci, M. Ravveduto, *La storia come bene comune, le nuove frontiere della public history digitale,* Passato e Presente, F. 113, Franco Angeli, Milano 2021, p. 126-130

166 https://siusa.archivi.beniculturali.it/cgi-bin/pagina.pl?TipoPag=profist&Chiave=238 ultima consultazione 30/07/2021

Proviamo ora a ripetere la stessa operazione cambiando le parole della nostra ricerca, e inseriamo quello che è l'oggetto del documento condiviso: "Piano E 27".

Il primo risultato che ci fornisce il sistema di ricerca ci rimanda al sito Militarystory.org[167], sito che non propone all'utente un resoconto storiografico di cosa sia stato il piano E27, ma lo mette di fronte al documento originale digitalizzato, non solo al comando operativo, ma al documento che fu elaborato dal CLN dal 1944. Scorrendo la pagina verso il basso troviamo il documento degli ordini che resero operativo il piano, fino ad arrivare in fondo dove ritroviamo il documento con l'ordine di agire, lo stesso da cui eravamo partiti leggendo il post della pagina Facebook.

Il sito su cui siamo approdati si presenta come un archivio inventato[168], non è la rappresentazione digitale di un archivio che esiste in senso materico, i documenti dunque non sono organizzati in fascicoli e buste con un vincolo di archiviazione, ma più semplicemente, grazie a tag, parole chiave e metadati, il sito organizza e condivide documenti digitalizzati relativi alla storia militare italiana, con focus specifici su aviazione, esercito e marina.

Il sito offre anche degli approfondimenti *ad hoc* su specifici argomenti, condividendo con l'utente non solo il documento cartaceo digitalizzato, ma anche documenti fotografici e audiovisivi. Per ogni documento è riportata la collocazione delle fonti all'interno degli archivi fisici, così da assicurare sempre la tracciabilità delle fonti.

Una cosa che trovo interessante è che il sito è amministrato da Michele Giuliani che di professione non è né storico né archivista, ma architetto urbanista[169]. A questo punto ci dovremmo chiedere se è possibile fidarsi di informazioni condivise da un non storico, ma sarebbe una domanda superflua, sul sito infatti il racconto non è lasciato a interpretazioni o a resoconti storiografici, ma unicamente ai documenti stessi e ci porta al riflettere sul tema delle *citizen sience*, che come scrive Paci

> la citizen science (CS), "scienza dei cittadini" o "scienza partecipata", indica la partecipazione e il coinvolgimento attivo dei cittadini in attività di ricerca scientifica. Mentre la CS include le scienze naturali, come la biologia, la chimica e la fisica e la citizen social science si occupa delle società, le citizen humanities si applicano alle discipline storiche, letterarie, linguistiche e filosofiche. La digital public history, per la sua vocazione pubblica, partecipativa e collaborativa in un contesto digitale, si presenta come uno dei campi di applicazione privilegiata dalle citizen humanities[170].

167 http://www.militarystory.org/il-piano-emergenza-27-del-comitato-di-liberazione-nazionale/ ultima consultazione 30/07/2021

168 Archivi inventati – il web e gli studi storici

169 https://www.linkedin.com/in/michele-giuliani-31185422/ ultima consultazione 30/07/2021

170 D. Paci, *Conoscere è partecipare: digital public history, wiki e citizen humanities*, Umanistica Digitale, in corso di pubblicazione

Il tema degli archivi inventati e della cocreazione di contenuti attraverso il coinvolgimento della cittadinanza è centrale all'interno dell'elaborazione teorica della Public History, in particolar modo in riferimento alla possibilità della cocreazione di archivi grazie alla condivisione dal basso di fonti storiche. Su questo tema, che trova posizioni discordanti all'interno della comunità di storici e storiche, torneremo con maggior attenzione, ma adesso sono interessato a sviluppare la questione della condivisione come forma assolutamente innovativa di accesso alle informazioni e di come la condivisione digitale possa essere uno strumento molto potente di avvicinamento alla storia e ai suoi documenti per chi storico non è.

Il secondo risultato offerto da Google in seguito alla ricerca "piano E 27".

Non senza sorpresa veniamo indirizzati ad accedere su www.metarchivi.it e cliccando sul link ci troviamo davanti a questo:

tabella 2.4.2

ARCHOS è un sistema integrato di catalogazione, archiviazione e ricerca, che opera sulle differenti tipologie di documenti appartenenti ai fondi e alle collezioni degli enti conservatori, in questo caso prevalentemente Istituiti Storici del Novecento e della Resistenza.

Il sistema è costituito da due ambienti: Biografie e Metarchivi, ciò che a noi interessa è la parte relativa ai Metarchivi, che si presenta come un multi-catalogo in cui l'accesso alle informazioni può avvenire esplorando la struttura gerarchica degli archivi, organizzata in livelli: fondi, serie, sottoserie, fascicoli, sottofascicoli, documenti. E dove ogni livello archivistico può essere aperto cliccando sulla relativa icona, mentre attivando il collegamento del suo titolo si accede alla relativa scheda descrittiva. Il catalogo conduce alla lista degli enti conservatori, dove scegliendo l'ente di interesse con un click si accede al fondo digitalizzato di quell'ente.[171]

ARCHOS si presenta come un punto di accesso unificato a una confederazione di archivi digitalizzati[172]; i documenti che è possibile osservare attraverso ARCHOS in forma digitale esistono in forma tangibile all'interno di archivi fisici, e la loro trasposizione in forma digitale, come è possibile vedere anche dall'immagine, mantiene fede a specifici processi di ordinamento archivistico.

Il primo e forse più importante, e completamente assente all'interno degli archivi inventati, è il vincolo archivistico, che come scrive Pigliapoco è l'insieme delle relazioni logiche e formali che esistono tra i documenti di un archivio. Il vincolo archivistico nasce spontaneamente, senza una precisa volontà del soggetto produttore, e si forma nel momento stesso in cui vengono prodotti i documenti[173].

Per dirlo con le parole di De Felice, il vincolo archivistico «non è frutto di una volontà esterna, ma è naturale, caratterizzato dalla necessità e originarietà e che determina una lunga catena di precedenti e susseguenti, catena che non può essere spezzata senza compromettere l'essenza stessa dell'archivio»[174].

Queste parole non possono che farci riflettere rispetto al processo di indagine che stiamo descrivendo, noi siamo partirti da un post di Facebook in cui si condivide l'immagine di un documento, l'utente che si trova di fronte al documento probabilmente non ha nessuna competenza in campo archivistico, non ha interesse verso i vincoli le classificazioni e l'ordinamento, il documento si presenta a lui come complessità narrativa atomica, nel senso che basta a se stessa per veicolare l'informazione. Ciò non di meno però, l'utente che incuriosito si propone di fare una breve e semplicissima ricerca (ricordiamoci che abbiamo usato solo le parole presenti nell'oggetto del documento per compiere la ricerca su Google) potrà trovarsi di fronte a una struttura gerarchica

171 http://www.metarchivi.it/ ultim consultazione 30/07/2021

172 Ancora in realtà in questo caso sono prevalentemente gli indici archivistici a essere digitalizzati, mentre sono poche le fonti digitalizzate possibili da fruire

173 S. Pigliapoco, *Progetto archivio digitale. Metodologia, sistemi, professionalità*, Civita Editori, Totre del lago 2018, p. 15

174 R. De Felice, *L'archivio contemporaneo. Titolario e classificazione sistematica di competenza nei moderni archivi di enti pubblici e provati*, Roma, La Nuova Italia Scientifica, 1988, p. 21

in cui si manifesta in modo chiaro la concatenazione di precedenti e susseguenti in cui i documenti stanno in relazione. Questa concatenazione è visibile e soprattutto navigabile, implica quindi che l'utente solo con un click possa accedere ad altri fascicoli e dunque aumentare il contesto storico e documentale all'interno del quale è inserito il documento, che abbiamo visto per la prima volta su una pagina di tifosi della Juventus.

Per far ricerca con Google non c'è bisogno di conoscerne i meccanismi di funzionamento né sapere alcunché di ciò che si vuol trovare: basta digitare una o poche parole nel suo *form* d'interrogazione per ottenere quantità a volte sorprendenti di risposte, nelle prime dieci o venti delle quali, c'è quasi sempre qualcosa che ci soddisfa o, più realisticamente, qualcosa di cui si finisce per accontentarsi. Non solo non occorre conoscere alcunché - o conoscere molto poco - per condurre ricerche su Google e trovare i documenti che ci interessano, non c'è neanche bisogno che qualcuno li abbia preventivamente organizzati, ordinati e classificati a nostro beneficio. I meccanismi di funzionamento di Google non solo non si basano affatto su classificazioni, ordinamenti, sistemazioni preventive dei documenti, ma, in un certo senso, evitano proprio che ciò possa accadere[175].

Per dirla con le parole dei creatori di Google, Brin e Page, «è una estesa collezione di documenti eterogenei completamente incontrollabili, vari dal punto di vista del formato (testo, immagini, suono, ecc.), da quello della codifica (html, pdf, ecc.), delle lingue utilizzate e dei vocabolari adottati»[176].

L'eccesso di produzione e condivisione di informazioni, che prima abbiamo osservato come uno dei problemi primari della nostra società, contiene al suo interno anche il post che abbiamo analizzato. Pensiamo di essere a un qualsiasi 25 aprile e al dibattito che inevitabilmente ogni anno si sviluppa sui social network e non solo, riguardo il ruolo della resistenza nel processo di liberazione dell'Italia dal nazifascismo. Le informazioni con le quali potremmo confrontarci sarebbero di tre livelli: storiografiche, politiche e documentali.

Prima dello sviluppo dei processi di digitalizzazione i sistemi di comunicazione consentivano di avere accesso prevalentemente a informazioni di tipo storiografico e politico, dunque alla storiografia riguardo la liberazione, attraverso libri, saggi e articoli che era possibile rinvenire in formato materico in biblioteche, librerie e edicole e informazioni politiche che fruivamo attraverso

175 S. Vitali, *Ordine e caos: Google e l'arte della memoria,* in A. Spaziante, a cura di, *Il futuro della memoria: la trasmissione del patrimonio culturale nell'era digitale*, CSI Piemonte, Torino, p. 85

176 S. Brin, L. Page, *The Anatomy of a Large-Scale Hypertextual Web Search Engine*, cit.; cfr. anche M. Agosti, M. Melucci, *Information Retrieval on the Web*, in M. Agosti, F. Crestani, G. Pasi (a cura di), *Lectures on Information Retrieval: Third European Summer-School, ESSIR 2000, Varenna, Italy, September 11-15, 2000*, Springer Verlag, Heidelberg 2001, in particolare pp. 276 e segg.

media di massa come la televisione, la radio e i giornali. Da non sottovalutare la figura dello storico televisivo che si pone a metà tra il tipo di informazione storiografica e quella politica.

La differenza tra questi tipi di informazione è che la prima si basa su un metodo di analisi delle fonti che si manifesta al fruitore attraverso l'organizzazione del discorso sviluppata da storici e storiche, ma che come sostiene Topolski si amalgama con il *savoir* quotidiano dello storico o della storica, che inevitabilmente domina lo studio del passato e la costruzione del racconto[177], e per questo costantemente messa in discussione da processi di revisione, fondamentali per l'approfondimento storico, ma deleteri quando frutto di volontà politica.

Il secondo tipo di informazione, quello politico appunto, si basa prevalentemente sull'analisi critica della storiografia e si amalgama con le credenze politiche e ideologiche di chi produce l'informazione stessa. In questo caso non vi è un metodo che possa definire l'informazione come *episteme*, ma ci si ferma necessariamente alla *doxa*, che assume più o meno valore unicamente in base all'autorità del mittente. Questo tipo di informazione ha più un valore di intrattenimento che di diffusione della conoscenza, è la base dell'utilizzo politico della storia, dove i contenuti vengono piegati a specifiche volontà e veicolati unicamente attraverso mezzi di comunicazione di massa, giornali, radio, televisioni.

La terza opzione, quella documentale, fino a prima della rivoluzione digitale era accessibile prevalentemente a storici e storiche e a qualche giornalista particolarmente motivato. Per potersi confrontare direttamente con le fonti primarie, unica reale rappresentazione della storia alla quale abbiamo accesso, era necessario accedere ad archivi di varia natura, comprenderne il funzionamento e faticosamente confrontarsi con la massa documentale materica. Il che richiedeva competenze specifiche e molto tempo.

Oggi, grazie alla digitalizzazione dei documenti materici e alla loro rimediazione, l'accesso all'informazione documentale cambia sensibilmente: se nel post che stiamo analizzando gli amministratori della pagina avessero riportato il link del documento digitalizzato, per accedere alla fonte primaria sarebbe bastato un *click*.

Le fonti, parafrasando Rossana Lista, sono l'unico medium attraverso il quale la realtà storica diventa accessibile[178] e oggi, proprio grazie alla rimediazione che, come sostengono Bolter e Grusin, così come prima di loro da McLuhan[179], «è la rappresentazione di un medium all'interno di un

177 J. Topolski, *Narrare la storia, Nuovi principi di metodologia storica*, Bruno Mondadori, Milano, 1997 p. 112

178 R. Lista, *Geschichtsbegriff. Storicizzazione del tempo e temporalizzazione della storia*, prefazione a R. Koselleck, *Storia la formazione del concetto moderno,* Clueb, Bolgna 2009, p. 14

179 M. McLuhan, *gli strumenti del comunicare,* Il Saggiatore, Milano 2008

altro»[180], «una caratteristica fondamentale dei nuovi media digitali»[181], gli archivi digitali possono contenere ogni tipo di fonte che sia mai stata prodotta, non tanto e solo come quantità, ma come tipologia. La principale caratteristica della rimediazione consiste nel fatto che l'evoluzione dei nuovi media mira alla realizzazione di numerose convergenze in una pluralità di forme mediali piuttosto che la loro omogeneizzazione in un singolo medium[182].

2.5 E le fonti materiche?

L'occasione che si sta mostrando ai nostri tempi è dunque di quelle irripetibili, la possibilità di portare la storia, nel suo manifestarsi e superare il tempo, alla portata di chiunque, una delle sfide che la rivoluzione digitale ci mostra e di non marginalizzare il mondo analogico e inq uesto specifico caso tutte le fonti materiche che risiedono in archivi fisici.

E' indubbio che nel nostro contemporaneo produciamo una quantità incommensurabile di dati e di informazioni, ma questo non è indicativo solo dei nostri giorni, da sempre l'essere umano produce una enorme mole di informazione, solo che l'essenza materica di tale informazione ne rende la fruizione complessa.

Per dare un'idea di natura pratica sulla documentazione che gli archivi contengono è opportuno, come suggerisce Isabella Zanni Rosiello, fare qualche esempio con riferimento alla nostra vita di tutti i giorni.

Nascere, frequentare scuole primarie, secondarie e università, sposarsi, acquistare un appartamento, pagare le tasse e fare ricorso contro eventuali contestazioni, servirsi di prestazioni sanitarie o essere ricoverati in un ospedale, svolgere un lavoro, fare testamento, morire ecc. sono tutti eventi che danno luogo a documentazione archivistica. Essa è posta in essere da determinati uffici, enti, persone: dagli uffici di stato civile per quanto riguarda nascite, matrimoni e morte; da scuole, uffici scolastici regionali, università per quanto riguarda la sfera scolastica; da uffici che si occupano di imposte e comuni per quanto riguarda tasse ed eventuali contestazioni; da enti mutualistici e ospedalieri per quanto riguarda le cure mediche; da imprese e aziende quando si tratta di attività lavorative, da notai per quanto riguarda acquisti di immobili o volontà testimoniali. Tutta questa documentazione serve sia agli stessi soggetti che l'hanno posta in essere, per propria memoria, sia

180 J.D. Bolter, R. Grusin, *Rimediation. Competizione e integrazione tra media vecchi e nuovi*, Guerini, Milano, 2002 p. 73

181 P. Montani, *Limmaginazione intermediale, Perlustrare rifigurare, testimoniare il mondo visibile*, Editori Laterza, Bari 2010, p. 10

182 J.D. Bolter, R. Grusin, *Rimediation. Competizione e integrazione tra media vecchi e nuovi*, Guerini, Milano, 2002 p.

alle persone eventualmente interessate a prendere visione delle pratiche che le riguardano o per avere copie, certificazioni attestati, al fine di soddisfare i propri interessi.

Se guardiamo ai secoli passati, comuni, signorie, ducati, repubbliche, regni, stati regionali, registravano dati e archiviavano informazioni[183].

La Biblioteca del Congresso degli Stati Uniti contiene 171 milioni di volumi[184], così come la British Library ne contiene altri 170 milioni per 746km di scaffali[185], in Italia ci sono 101 Archivi di Stato e a questi vanno aggiunti gli archivi di impresa, gli archivi di persona, gli archivi degli istituti storici, gli archivi di famiglia, degli enti pubblici territoriali e non territoriali, dei notai, gli archivi ecclesiastici, le emeroteche, i musei e le pinacoteche. Proviamo a immaginare la quantità di documenti, dunque informazioni, dunque dati, che sono contenuti all'interno di questi archivi e biblioteche.

In Italia nel 1990 i quotidiani tiravano un totale di circa 9.763.197 copie giornaliere[186].

Dopo aver osservare sia da un punto di vista quantitativo che qualitativo in cosa consista l'infosfera e in che termini una società si possa definire iperstorica; dopo che abbiamo potuto osservare come la condivisione di documenti dematerializzati, per quanto inserita in un contesto complesso come quello della sovrapproduzione di dati, consenta un nuovo processo di interazione con le fonti, come per esempio partendo da un documento storico Google possa contribuire in modo sensibile a creare un contesto di riferimento in cui inserire il documento stesso; visto che ci son chiare alcune delle possibilità maggiormente rilevanti dell'accesso all'informazione attraverso strumenti digitali, credo che non sia possibile immaginare di non coinvolgere le fonti materiche e gli artefatti tangibili, conservati nella loro fisicità, in un processo di dematerializzazione e rimediazione che non elimini ciò che già esiste e che sappiamo come fruire, ma che consenta a nuove e vecchie generazioni di iniziare un rapporto completamente nuovo con le tracce del passato, a partire dall'incredibile possibilità di "trasportare" un intero archivio digitalizzato all'interno del nostro spazio di lavoro.

2.6 Presenza

«Negli studi attuali la presenza è spesso compresa come un tipo di esperienza dell'"essere qui", che coinvolge genericamente qualche mediazione tecnologica e spesso dipende da ambienti virtuali.»[187]

183 I. Zanni Rosiello, *Gli archivi nella società contemporanea*, Il Mulino, Bologna 2009, p. 33
184 https://www.loc.gov/about/general-information/ ultima consultazione 15/07/2021
185 https://www.bl.uk/about-us/our-story/facts-and-figures-of-the-british-library# ultima consultazione 15/07/2021
186 *La stampa in Italia. 1998-2001*, Federazione Italiana Editori e Giornali, 2001 disponibile online https://www.fieg.it/upload/studi_allegati/studio19982001.pdf
187 L. Floridi, *Pensare l'infosfera, la filosofia come design concettuale*, Raffaello Cortina editore 2020, Milano p. 66

Immaginiamo di essere sulla soglia di un archivio senza pareti, all'interno del quale le fonti, ossia le informazioni, non sono altro che codici numerici, mentre macchine intelligenti si adoperano per renderci questi codici intellegibili e contemporaneamente ci propongono strade di ricerca[188].

Con la digitalizzazione delle fonti materiche e la costruzione di archivi digitali consultabili online, si rende l'archivio un luogo diverso da ciò che è stato fino a pochi anni fa.

La parola "archivio", che come sappiamo ha almeno un duplice significato, indica sia il complesso documentario sia il luogo fisico di conservazione, con la digitalizzazione subisce un processo probabilmente controintuitivo: raddoppia il proprio significato fisico.

Vi sarà infatti il luogo di conservazione delle fonti materiche e il luogo di conservazione delle fonti digitalizzate. Questa doppia fisicità rappresenta un doppio livello di conservazione delle fonti, e un doppio sistema di consultazione. Si potrà dunque accedere all'archivio materico, recandosi fisicamente all'interno di un edificio, e si potrà accedere alle fonti entrando in un archivio senza che il corpo si sposti. Non è da escludere l'ipotesi che una ricerca svolta attraverso un archivio digitale e la stessa ricerca svolta in un archivio materico possano non portare allo stesso risultato. Non è la presenza o l'assenza del nostro corpo all'interno di un archivio a modificare i risultati di una ricerca, ma è il nostro livello di presenza esperenziale, l'interazione con l'archivio e i documenti a esso interni a mutare.

Questa discrepanza che non dico vada in favore di uno o dell'altro scenario è presumibile dal fatto che gli archivi digitali, se sviluppati attraverso le logiche e le funzionalità del Web semantico possono rendere interoperabili[189] diverse fonti e possono fornire al fruitore dei percorsi precostituiti di ricerca. I tempi di consultazione per il ricercatore non saranno più vincolati ai tempi tecnici di un archivio fisico e la ricerca sarà coadiuvata da sistemi tecnologici intelligenti. Nell'archivio materico invece si è guidati sempre dall'istinto e dalle competenze dello storico, coadiuvato dall'archivista che utilizzando gli indici, i cataloghi e le proprie conoscenze del fondo archivistico sarà in grado di guidare la ricerca e indicare fonti correlate all'interesse della ricercatrice. D'altra parte saremo però legati ai tempi dell'archivio che in quanto entità fisica è regolamentato da procedure burocratiche complesse che non necessariamente aiuteranno il ricercatore o la ricercatrice.

Attraverso la digitalizzazione, i documenti e dunque le fonti storiche assumono in modo ancora più definito il significato di informazioni[190].

Ciò che il percorso di ricerca richiede di fare consta anche nel comprendere in che modo i processi di digitalizzazione che coinvolgono la disciplina della storia possano essere osservati attraverso la

188 Vedremo nei prossimi capitoli come l'utilizzo dei metadati, se inseriti in un vocabolario comune, consenta la realizzazione di ontologie che grazie alle possibilità del Web semantico possono arrivare a proporre dei percorsi di ricerca interni all'archivio digitale.

189 S. Vitali, *Passato digitale. Le fonti dello storico nell'era del computer,* Bruno Mondadori, Milano 2004, p. 121

190 Tra gli altri: Floridi (2020,2006); T.W. Bynum and J.H. Moor (2003); Levy (1995); Maldonato (2005)

lente dell'infosfera, un luogo fisico e virtuale all'interno del quale non siamo entità isolate, quanto piuttosto agenti informazionali interconnessi, che condividono con altri agenti biologici e artefatti ingegneristici un ambiente globale, costituito in ultima istanza da dati.[191]

Visto che in queste pagine si vuole analizzare come mutano i processi relativi all'accessibilità alle fonti digitalizzate interne ad archivi digitali, dovremmo prendere in considerazione le teorie riguardo la *tele-presenza* e la *tele-epistemica*.[192]

Ad oggi per quanto riguarda la relazione con gli archivi digitali non si può parlare di tele-presenza, almeno fino a quando non si arriverà a interagire attraverso dei nostri avatar digitali all'interno di un metaverso in cui è possibile accedere alla ricostruzione virtuale dell'ambiente fisico di un archivio, o la nostra presenza all'interno di un archivio non sarà mediata da delle infrastrutture tecnologiche (robot) presenti fisicamente nell'archivio, mentre noi siamo presenti nel nostro spazio di lavoro e non interagiamo con lo spazio locale di osservazione (SLO) nel quale siamo situati, ma con lo spazio remoto di osservazione (SRO) nel quale è situato il nostro avatar.

Questi processi sono particolarmente utilizzati in medicina dove un medico, con l'aiuto di un robot, può interagire e finanche operare un paziente e collaborare con il personale medico in SRO mentre risiede in SLO[193].

Per dirlo con Riva «Il senso di presenza è la percezione e l'esperienza soggettiva che deriva dal trovarsi in un ambiente. Nel nostro caso è "*l'esserci, cioè la sensazione di essere nel mondo sintetico generato dal pc*", esserci davvero, perdendo la consapevolezza della mediazione della tecnologia»[194]

La digitalizzazione dei documenti comporta l'ordinamento di questi in senso unicamente logico e non fisico[195], mentre negli archivi materici l'ordinamento dei documenti è sia logico che fisico, questo doppio livello garantisce una prassi scientifica di archiviazione che consente al ricercatore di pensare di individuare un documento e una informazione, o la loro assenza, anche grazie alla posizione che dovrebbe avere fisicamente all'interno dei complessi documentali. L'ordinamento logico non consente questa strategia di pensiero, ma si basa unicamente sulle relazioni informative attraverso le quali sono messi in collegamento tra di loro i documenti e il sistema di gestione. Se riproducessimo in modo virtuale anche la composizione fisica di un archivio, andando a

191 L. Floridi, *Pensare l'infosfera. La filosofia come design concettuale*, Raffaello Cortina editore 2020, Milano p. 15

192 K. Goldberg, (a cura di) *The robot in the garden: telerobotics and telepistemology in the age of the internt, MIT Press, Cambridge, MA-London*

193 https://www.youtube.com/watch?v=fU7Dk8ZZSLY, il video è stato realizzato grazie alla collaborazione tra Hopital Avicenne AP-HP, Assistance Pubblique Hopitaux de Paris, Evolutis France, Universitè Sorbonne Paris Nord, Microsoft, Ultima consultazione 09/08/2021

194 G. Riva, *Tecnologie della Presenza. Concetti e applicazioni*, Mondo Digitale, 2009; (3), p. 32-45

195 M. Guercio, *Archivistica informatica. I documenti in ambiente digitale*, Carocci, Roma 2000 p. 65

georeferenziare[196] i documenti in base allo spazio che occupano nell'archivio materico, o utilizzando strumenti che sfruttano la tecnologia RFID di cui abbiamo parlato nel primo capitolo, lo storico potrebbe recuperare tutte le funzioni di ricerca che l'archivio materico garantisce.
Ma senza arrivare a teorizzare la costruzione di archivi virtuali che georeferenziano le fonti dando così l'opportunità al ricercatore di agire in SRO anche se ci si trova in SLO, attraverso l'utilizzo di avatar o robot, l'analisi critica del concetto di presenza non può essere trascurata in un momento storico in cui la dematerializzazione porta un nuovo significato di presenza e interazione tra lo storico e la sua materia prima, ossia i documenti.
The International Society for Presence Research (ISPR) è una organizzazione no-profit fondata nel 2002 per supportare la ricerca accademica rispetto al concetto di telepresenza[197], scrive:

«La (tele)presenza è uno stato psicologico o una percezione soggettiva in cui, anche se parte o tutta l'esperienza attuale dell'individuo è generata o filtrata da una tecnologia adattata all'umano, parte o tutta la percezione dell'individuo non è in grado di riconoscere accuratamente il ruolo giocato dalla tecnologia nell'esperienza. A eccezione dei casi più estremi, l'individuo può indicare correttamente che sta utilizzando la tecnologia, ma "a qualche livello" e in "qualche misura" le sue percezioni non rilevano che conoscenza, oggetti e enti e ambienti sono percepiti come se la tecnologia non fosse coinvolta nell'esperienza. L'esperienza è definita come l'osservazione da parte di una persona di, e/o l'interazione con, oggetti enti o eventi nel suo ambiente; la percezione, il risultato del percepire, è definita come la significativa interpretazione dell'esperienza (…). La presenza si verifica quando l'esperienza parziale o totale di un individuo è mediata non solo dai sensi umani e dai processi percettivi, ma anche dalla tecnologia creata dall'uomo, mentre la persona percepisce l'esperienza come se fosse solo mediata dai sensi umani e dai processi percettivi. Pertanto, la presenza richiede ai ricercatori di confrontare le percezioni e le risposte umane nel contesto della tecnologia con le percezioni e le risposte umane in contesti che non coinvolgono la tecnologia (spesso indicati come contesti "faccia a faccia" o "interpersonali" o, in qualche modo confusamente, contesti "non mediati" o "reali")»[198]

Secondo la Società Internazionale per la ricerca sulla Presenza, l'interazione e l'esperienza di ricerca, se avviene grazie a una interazione tra umano e macchina, potrebbe mutare anche se al ricercatore potrebbe sembrare che nulla cambi, il nodo della questione risiede nel modo in cui si è abituati a reagire alle percezioni che si hanno nell'interazione con l'ambiente nel quale si agisce, senza il coinvolgimento delle tecnologie le percezioni dell'uomo sono veicolate unicamente

196 Con il termine georeferenziazione ci si riferisce al processo mediante il quale si assegnano delle coordinate del mondo reale a ciascun pixel del raster http://www.qgistutorials.com/it/docs/3/georeferencing_basics.html
197 https://ispr.info/ Ultima consultazione 28/07/2021
198 International Society for Presence Resource https://ispr.info/about-presence-2/about-presence/ Trad mia, Ultima consultazione 28/07/2021

attraverso i nostri sensi, mentre quando interviene una tecnologia che ci consente un accesso in tele-presenza o telepistemico all'ambiente con il quale entriamo in relazione, non sono più solo i nostri sensi a mutuare le nostre percezioni, c'è un nuovo elemento, quello tecnologico, che non sempre cogliamo come determinante nella percezione, ma che nella realtà dei fatti non solo la rende possibile ma ne stabilisce anche le modalità.

Prendere consapevolezza di questa nuova intermediazione dell'esperienza diviene necessario nel momento in cui le fonti non sono più fisiche, ma un insieme di dati collegati tra di loro, l'archivio come luogo fisico perde importanza e la ricerca viene mediata da più strumenti digitali che rispondono a processi informatici precostituiti capaci di rispondere alle nostre interrogazioni.

Quello che si richiede a questo punto è di provare a proporre una teoria capace di osservare i mutamenti che la digitalizzazione inevitabilmente impone alla ricerca e alla diffusione del contenuto storico.

Secondo Floridi, in base ad un processo che ha chiamato *design concettuale*[199], per riuscire a comprendere la rivoluzione che stiamo vivendo e dunque produrre pensiero in grado di descrivere questi cambiamenti, dobbiamo considerare noi stessi come degli inforganismi in un'esperienza che è descrivibile in termini di on-life experience. La rivoluzione tecnologica che stiamo vivendo ha mutato radicalmente l'ambiente in cui viviamo e la percezione di noi stessi, rendendo vano ogni tentativo di analisi che non tenga in considerazione la realtà digitale, anzi la proposta di Floridi è quella di considerare ormai irreversibilmente integrati l'aspetto fisico e quello digitale della realtà.

Se il discorso filosofico vuole essere analisi, oppure critica, o ancora design deve partire dal riconoscimento di questo stato di cose[200].

Secondo Floridi ogni sistema, ogni insieme di oggetti o soggetti osservati, può essere analizzato filosoficamente a un livello di astrazione (LdA), attraverso il quale è possibile generare un modello del sistema in grado di identificare delle proprietà che poi vengono attribuite al sistema stesso.

È chiaro che il punto fondamentale da comprendere è che l'individuazione di un LdA decide sull'impegno ontologico della nostra teoria[201].

Mantovani e Riva scrivono che il significato di presenza dipende dal concetto che abbiamo di realtà (dall'ontologia che in modo più o meno esplicito adottiamo) e che differenti posizioni ontologiche generano differenti definizioni di presenza, tele-presenza e presenza virtuale[202].

---

199 L. Floridi, *Pensare l'infosfera. La filosofia come design concettuale,* Raffaello Cortina editore 2020, Milano p. 97

200 O. Palmini, recensione a L. Floridi, *Pensare l'infosfera. La filosofia come design concettuale,* in «Pandora Rivista» 20-02-2020

201 https://www.pandorarivista.it/articoli/pensare-infosfera-di-luciano-floridi/ ultima consultazione 08/08/2021

202 G. Mantovani, G. Riva, *Buildinig a bridge between different scientific communities: On Sheridans eclectic ontology of presence,* in Presence: teleoperators and virtual Environments, 10 (5), pp. 537, 543

Con i processi di digitalizzazione delle fonti e degli archivi lo storico non entra più in un contatto fisico con le fonti e neanche con i luoghi di conservazione, ma accede a una interfaccia, ossia una modellizzazione, dove le variabili presenti nel modello corrispondono agli osservabili nella realtà. Vedremo come questo concetto di modellizzazione si attui nella costruzione degli archivi digitali ad un livello precedente a quello dell'interfaccia, dando forma a sistemi come EAD e EAC che sono la rappresentazione informatica degli standard ISAD(G) e ISAAR[203]; ma di questo, che ha più a che fare con i processi tecnici, ci occuperemo più avanti, adesso la mia intenzione è di provare a delineare un quadro all'interno del quale poter osservare i possibili funzionamenti della dematerializzazione nell'ambito della ricerca storica.

Per comprendere meglio quello che sto dicendo può essere utile tornare ad osservare la tabella 2.2, dove si mostra il sistema di organizzazione digitale delle fonti all'interno del portale ARCHOS, ciò che vediamo è un'interfaccia con la quale è possibile interagire grazie ad una modellizzazione dove le variabili presenti, l'ordinamento delle buste e dei fascicoli, corrispondono a quelle osservabili nella realtà. Se dunque ad un'interazione superficiale può sembrare che nulla cambi, nella realtà dei fati quei documenti, una volta digitalizzati non sono realmente in un relazione fisica uguale a come la vediamo riprodotta sullo schermo, quella è una modellizzazione che noi siamo in grado di interpretare; nella realtà dei fatti quei documenti sono in relazione tra di loro solo attraverso un processo logico e informatico, inoltre ARCHOS, in quanto una confederazione di archivi, rende la cosa ancora più affascinante, dal momento che non esiste nessun luogo fisico dove io mi possa recare e avere accesso a tutti i documenti a cui ho accesso tramite ARCHOS.

Possiamo dunque osservare un archivio digitale come un LdA, una interfaccia composta da un insieme di osservabili ciascuno con un proprio definito insieme di valori o risultati.

A questo punto rimane da chiederci cosa avviene quando il ricercatore entra in relazione con un archivio digitale.

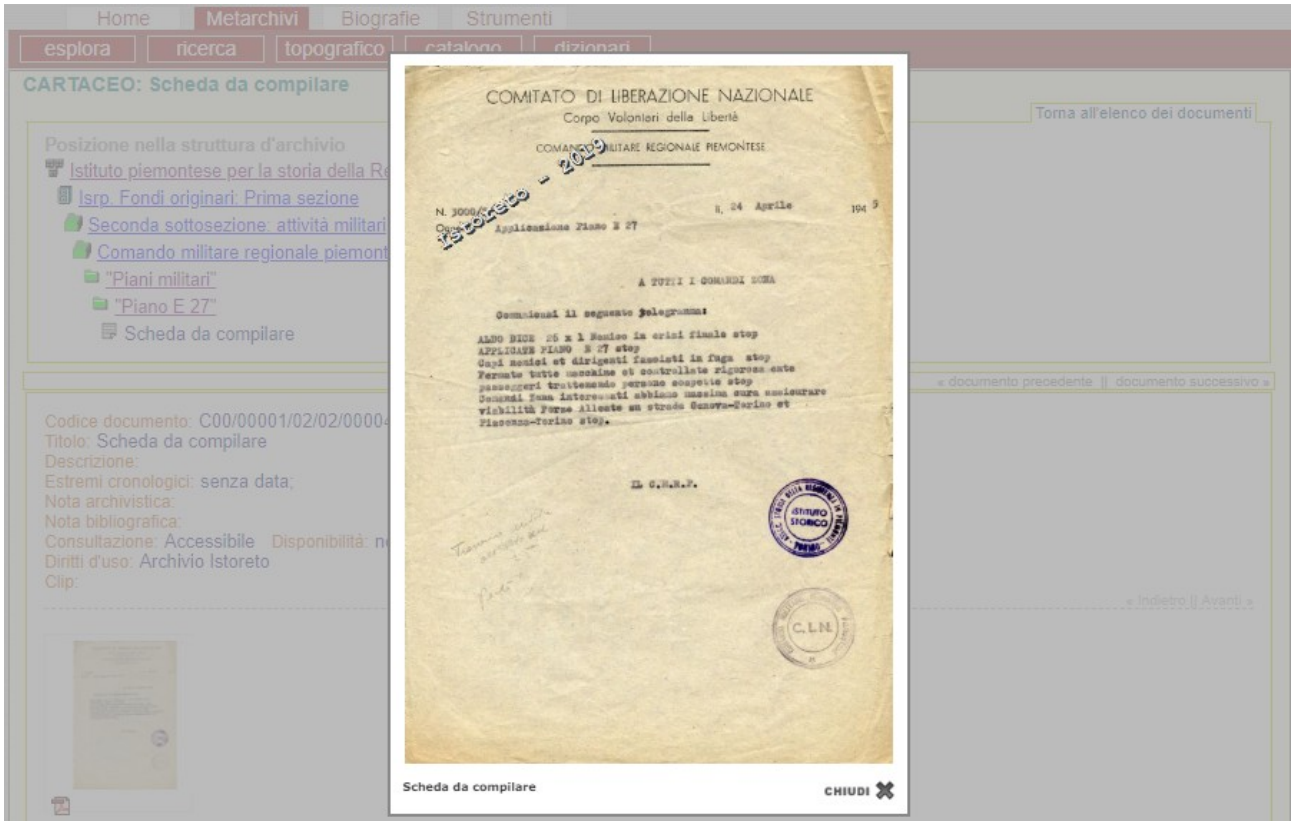
Attraverso un LdA l'osservatore ha accesso all'ambiente, un LdA potrebbe consistere per esempio nei cinque sensi, in un microscopio, in una telecamera, o nel nostro caso in un archivio digitale.

Negli archivi materici gli LdA rispondevano a delle variabili tipizzate specifiche, come per esempio l'iscrizione all'ingresso dell'archivio, il dialogo con l'archivista, la richiesta cartacea o informatica dei documenti, i tempi di attesa per riceverli, il numero massimo di prese giornaiere, la cura nel maneggiare i documenti e infine la restituzione con firma o assegnazione digitale o cartacea del deposito.

203 Su gli standard internazionali come Isad-G e ISAAE e la loro trasposizione digitale EAD e EAC-CPF, ci torneremo più avanti in un capitolo dedicato, in cui osserveremo la struttura e il funzinamento di questi standard

In un archivio digitale le variabili tipizzate che compongono LdA mutano in modo sensibile. L’accesso avviene attraverso una ricerca sul web, non ci sono orari di entrata e uscita, non si deve segnare la presenza in archivio, non si ha un’interazione con l’archivista. Si produce una richiesta su un motore di ricerca interno all’archivio, usando parole chiave o usando l’indirizzo di collocazione se già se ne è a conoscenza. La richiesta di visione del documento non viene formalizzata con una pratica e non ci sono tempi di attesa, in seguito alla ricerca il documento appare sullo schermo. Perché il documento appaia sullo schermo vi è stato però un lungo e complesso lavoro multidisciplinare composto dalla dematerializzazione del documento fisico originale, l'attribuzione di metadati e, in alcuni casi, la realizzazione di ontologie capaci di rendere interconnesse le fonti.

Torniamo sempre ad ARCHOS e una volta visualizzata l’interfaccia decidiamo di interagire con l’ambiente digitale cliccando sulla parola in rosso “mostra” che si trova accanto alle parole “documenti digitalizzati (1)”, quello che avviene è che il sistema mostra sullo schermo la riproduzione simulacrale del documento ri-mediato dell’ordine operativo E 27. Una volta che siamo difronte a questo documento, siamo noi forse dentro all’archivio? In realtà la nostra presenza risiede nel luogo in cui siamo, dunque il nostro Spazio Locale di Osservazione, il nostro corpo non si è spostato e non vi sono avatar o robot che ci consentono di interagire in tele presenza con lo Spazio Remoto di Osservazione dentro al quale risiede il documento. E’ come se nella stanza in cui siamo seduti fosse scomparsa una parete e dunque ci è possibile vedere ciò che c’è nella stanza accanto, dandoci compiutezza di ciò che contiene ma senza che noi ci si sia fisicamente spostati.



È necessario comprendere che non è rendendo lo spazio remoto (SRO) disponibile da un punto di vista epistemico a far sì che noi si sia presenti in SRO come enti.

Quando si naviga un archivio digitale e si procede con le nostre ricerche interagendo dunque con lo spazio virtuale e con la rappresentazione simulacrale dei documenti, non stiamo agendo in tele-presenza, ma in una presenza telepistemica in cui non è l'osservatore x (il ricercatore) in SLO, che accede all'ente y (il documento) in SRO a essere presente in SRO, ma al contrario è l'ente y in SRO a cui l'osservatore ha accesso telepistemicamente a essere presente anche nello SLO dell'osservatore[204].

«Possiamo paragonare questa situazione all'illusione di movimento generata dalla navigazione sul Web: abbiamo la sensazione che stiamo caricando la nostra presenza in spazi differenti, quando invece stiamo scaricando tali spazi nel nostro.»[205]

Seguendo le teorie sulla presenza all'interno dei processi che governano la filosofia dell'informazione e l' infosfera giungiamo a poter ipotizzare che nell'interazione uomo digitale non sia più lo storico a recarsi in un archivio, ma che sia l'archivio, o meglio le informazioni che contiene, a divenire presenti nello spazio di lavoro dello storico. Questo punto di vista ci fa

204 L. Floridi, *Pensare l'infosfera. La filosofia come design concettuale*, Raffaello Cortina editore 2020, Milano p. 83
205 ibidem

osservare le tecnologie e la digitalizzazione delle fonti come un processo al completo servizio del possibile fruitore. Saranno le informazioni, i documenti, a giungere a noi.
Il documento del comando E 27, una volta visualizzato sarà all'interno del nostro Spazio Locale di Osservazione nonostante risieda anche nello Spazio Remoto di Osservazione nel quale abbiamo un accesso telepistemico. Il documento dunque si moltiplica e moltiplica la propria presenza telepistemica, pensiamo infatti che lo stesso documento può essere osservato e scaricato contemporaneamente nello Spazio Locale di Osservazione di un numero potenzialmente infinito di fruitori.

2.7 Telepistemica e didattica

Prima ho affermato che il ruolo sociale dello storico può prendere forma attraverso due diverse pratiche, la prima, riguarda la produzione storiografica, che con l'avvento del digitale si potrà manifestare con l'elaborazione di prodotti multimediali o ancora più interessante direttamente all'interno di archivi digitali che sfruttino le possibilità del Web semantico, la seconda, che ho intenzione di approfondire con degli esempi pratici, riguarda l'insegnamento del metodo storico di ricerca e analisi critica delle fonti.
Durante l'anno scolastico 2017-2018, l'Istituto Luce Cinecittà ha sviluppato in collaborazione con il Liceo Tacito di Roma, all'interno di un programma di Alternanza Scuola Lavoro, un progetto dal titolo "Fascismo tra conformismo e repressione"[206], che aveva nei suoi propositi quello di far lavorare studenti e studentesse della V B, sull'analisi delle fonti storiche per arrivare a elaborare un prodotto multimediale.
Ciò che a noi interessa adesso non è tanto analizzare il prodotto finale, ma soffermarci sul processo di metodo che ha portato allo sviluppo dell'elaborato finale e in particolar modo come il tutto sia potuto avvenire grazie ad un accesso telepistemico alle fonti.
Durante una prima fase di confronto con la classe si è arrivati a decidere di lavorare sul fascismo concentrandosi su cinque macrotemi: scuola; sport; comunicazione; urbanistica; opposizione.
La scelta di non imporre un tema, ma di far decidere alla classe su cosa volesse lavorare, oltre ad avere un significato pedagogico mettendo studenti e studentesse al centro dei processi decisionali senza dover accettare passivamente una decisione dall'alto, ha fatto sì che la classe iniziasse da subito a impostare il lavoro come farebbe uno storico o una storica, interrogandosi sul tema da studiare e decidendo i diversi punti di vista con cui affrontare l'archivio.

206 https://www.camera.it/_dati/leg18/lavori/documentiparlamentari/indiceetesti/015/134/00000005.pdf ultima consultazione 03/08/2021

Dopo una prima revisione della letteratura storiografica riguardo i temi scelti, la classe è stata divisa in gruppi e a ogni gruppo è stato assegnato un tema sul quale iniziare a fare una ricerca all'interno dell'archivio audiovisivo dell'Istituto Luce.

Per procedere, studenti e studentesse hanno iniziato ad interfacciarsi con il sito dell'Archivio Storico dell'Istituto Luce[207], archivio tra i primi in Italia ad essere digitalizzato e a consentire la fruizione dei suoi documenti, se pur in bassa definizione, gratuitamente a chiunque acceda al portale.

Non essendo esperti di ricerca archivistica per interrogare l'archivio, senza confrontarsi con un archivista, gli studenti hanno iniziato la loro ricerca utilizzando parole chiave che secondo loro avevano attinenza con i singoli argomenti da approfondire. Così facendo, digitando per esempio la parola "scuola" il sistema gli ha proposto una serie di risultati che nella descrizione archivistica riportavano quella parola.

Invece di doversi recare fisicamente in un archivio, consultare l'indice, chiedere indicazioni all'archivista, fare richiesta dei documenti e aspettare di trovarsi davanti le buste, gli studenti e le studentesse hanno potuto agire utilizzando un mezzo con cui hanno enorme confidenza, interrogare l'archivio del luce infatti non è stato diverso dal fare una ricerca su YouTube e i risultati ottenuti hanno messo gli studenti immediatamente davanti alle fonti, a quel punto non dovevano far altro che premere play per poter iniziare a consultare delle fonti primarie.

La prima parte del lavoro di ricerca è stata svolta in completa autonomia, senza che docenti e formatori li guidassero nel processo, man mano che gli studenti andavano avanti nella navigazione dell'archivio riuscivano ad affinare sempre di più la ricerca capendo come muoversi anche grazie all'aiuto di metadati che consentivano di utilizzare nomi propri, date e luoghi.

Fino a qui possiamo osservare come il digitale abbia consentito una nuova modalità di fruire un archivio rendendolo inconfutabilmente maggiormente accessibile e consentendo una ricerca più semplice di come si sarebbe potuta svolgere se avesse riguardato ancora i documenti in forma materica.

Pensiamo solo ai diversi formati delle bobine e alla necessità di macchinari analogici per poterle fruire, ai tempi che avrebbe richiesto compiere un'azione del genere e al personale che si sarebbe dovuto impiegare per coadiuvare il ricercatore.

Mentre la classe proseguiva nella ricerca archivistica, per ogni documento che trovavano di interesse per il lavoro da svolgere dovevano registrare su un cartella condivisa su Google Drive i

207 https://www.archivioluce.com/ quando si è tenuto il progetto l'archivio digitale dell'istituto luce era organizzato in modo differente da oggi. Ora infatti già dalla home page appaiono dei pulsanti che danno l'opportunità di procedere nella ricerca attraverso percorsi tematici

link al quale era possibile trovare il documento, il titolo con il quale era stato registrato, la tipologia di filmato e la collocazione con cui il documento era stato archiviato.

Nella seconda parte del lavoro i gruppi hanno iniziato a selezionare il materiale trovato, concentrandosi su quei documenti che meglio di altri riuscivano a raccontare il tema su cui stavano facendo la ricerca. è in questa fase che si sono confrontati e confrontate con l'analisi critica delle fonti che partiva dall'analisi critica dell'archivio stesso, che come sappiamo nasce come organo di propaganda di un regime.

In questa fase il lavoro dei gruppi è stato supportato dalla professoressa di Storia e Filosofia e dal formatore incaricato dall'Istituto Luce, ed è da qui in poi che l'accesso telepestimico alle fonti digitali si manifesta in tutte le sue possibilità.

Come abbiamo detto nel paragrafo precedente con una presenza di natura telepistemica non è l'osservatore x (in questo caso lo studente o la studentessa) in SLO (in classe) che accede all'ente y (il documento) in SRO a essere presente in SRO (all'interno dell'archivio Storico dell'Istituto Luce), ma al contrario è l'ente y (il documento) in SRO (nell'Archivio Storico dell'Istituto Luce) a cui l'osservatore ha accesso telepistemicamente a essere presente anche nello SLO (in classe) dell'osservatore (lo studente o la studentessa). Questo vuol dire che agli studenti e alle studentesse è stato possibile portare all'interno del loro spazio di lavoro digitale i documenti.

Il fatto che il documento sia all'interno dello spazio di lavoro del ricercatore muta in modo deciso il rapporto che da sempre vi è stato tra ricercatore/archivio/documento. I documenti materici infatti in nessun modo possono uscire da un Archivio storico, sono vincolati al luogo e al resto della documentazione presente. Con l'entrata in vigore della Legge annuale per il mercato e la concorrenza, n. 124/2017[208], il 29 agosto è stato introdotto il regime di libera riproduzione con mezzo proprio nelle biblioteche e negli archivi pubblici italiani art. 1, c. 171. In precedenza negli archivi di Stato, l'uso della propria fotocamera o smartphone era normalmente interdetto in presenza di un concessionario esterno di fotoriproduzione, mentre nei casi, più frequenti, in cui la fotografia con mezzo proprio era permessa, rimaneva sempre subordinata a una specifica richiesta di autorizzazione preventiva e al pagamento di una tariffa di tre euro per ogni unità archivistica da riprodurre[209].

La nuova formulazione dell'art. 108 del Codice dei beni culturali[210] ha determinato la rimozione di alcuni divieti, tariffe e autorizzazioni, consentendo per la prima volta agli utenti di tutte le biblioteche e degli archivi pubblici italiani di riprodurre liberamente con dispositivi digitali a

208 https://www.gazzettaufficiale.it/eli/id/2017/08/14/17G00140/sg ultima consultazione 08/08/2021

209 http://www.ilmondodegliarchivi.org/component/content/article?id=546:la-liberalizzazione-delle-riproduzioni-digitali-con-mezzo-proprio-negli-archivi-di-stato-italiani ultima consultazione 08/08/2021

210 https://www.isontina.beniculturali.it/getFile.php?id=4034 ultima consultazione 08/08/2021

distanza sia gli stampati sia i documenti d'archivio non sottoposti a restrizioni di consultabilità per ragioni di riservatezza (ai sensi degli artt. 122-127 del codice dei beni culturali), nel rispetto in ogni caso delle norme poste a tutela del diritto d'autore e della privacy.

È bene precisare che riproduzioni "libere" non significa semplicemente "gratuite", ma anche esenti da qualsiasi richiesta scritta di autorizzazione. Dal 2017, per gli archivi afferenti allo Stato italiano, storia a sé gli archivi vaticani, è dunque possibile creare autonomamente copie digitali dei documenti che si sta consultando senza dover pagare o essere costretti ad utilizzare un servizio di scansione interno all'archivio.

Questa legge aveva già posto le basi per un cambiamento nella relazione tra ricercatore/documento/archivio, ma ancora prevedeva la presenza fisica del ricercatore all'interno dell'archivio e una fruizione che non fosse telepistemica, la digitalizzazione infatti interveniva solo in seguito all'azione del ricercatore o della ricercatrice, non era sistematizzata e avveniva in condizioni non ottimali per una precisa riproduzione simulacrale del documento, che comunque subiva un'azione di dematerializzazione e rimediazione, ma solo nello spazio di lavoro del singolo ricercatore che scattava la foto digitale.

La possibilità di scattare foto digitali al documento materico consente la riproduzione, non professionale, solo di documenti a stampa o manoscritti di libera consultazione, ma in nessun modo consente la riproduzioni di fonti audiovisive e, altro dato di interesse, l'archiviazione più o meno ordinata delle fonti a cui si è scattato una foto digitale rimane in uno spazio privato di fruizione.

Con la creazione di archivi digitali dove è possibile fruire la fonte digitalizzata e ordinata all'interno dell'archivio stesso, si consente un accesso pubblico alle informazioni e un utilizzo potenzialmente massificato del documento-medium.

Tornando al laboratorio sviluppato dall'Istituto Luce, mi voglio soffermare sulla seconda fase di lavoro che ha coinvolto studenti e studentesse, dopo la selezione dei documenti utili ai fini del lavoro da svolgere, ogni gruppo ha potuto operare dei tagli alla fonte audiovisiva in base ai frame di cui avevano bisogno per realizzare il montaggio con altri documenti audiovisivi[211].

Questo lavoro ci dimostra che grazie ai processi di digitalizzazione, alla costruzione di archivi digitali e ad un accesso telepistemico alla fonte, il ricercatore, oltre ad avere un accesso facilitato alla fonte, si può "appropriare" del documento digitale e una volta ottenuto può utilizzarlo per metterlo in scena e renderlo parte intelleggibile del discorso[212].

---

211 E. Becattini, M. Scanagatta, *La nuova grande guerra, la prima guerra mondiale nel cinema contemporaneo*, G. Gozzini, P. Masciullo (a cura di) *La guerra delle immagini nel XXI Secolo,* «Cinema e Storia», Rubbettino Editore, Catanzaro 2020, pp. 49-70 Sul tema della modificazione delle fonti, in particolar modo audiovisive, per la realizzazione di prodotti multimediali torneremo nel terzo capitolo.

212 M. Scanagatta, *Public Historian, tra ricerca e azione creativa* in, a cura di L. Bertucelli, P. Bertella Farneti, A. Botti, *Public History. Discussioni e pratiche*, Mimesis Edizioni, Milano Udine, 2017 p.p. 324-328

Con questo esempio ho voluto dimostrare che le riflessioni sulla telepistemica non sono fini a se stesse come potrebbero sembrare; è necessario interrogrsi su tutte le implicazioni della digitalizzazione così da poterne cogliere i vantaggi e chiaramente limiti e problemi che presenta.
Ora sappiamo che grazie ad un processo di digitalizzazione di un intero archivio e grazie alla costruzione di una interfaccia semplice da fruire, i documenti all'interno di uno degli archivi più importanti d' Italia, non solo sono fruibili per chiunque abbia una connessione ad internet, ma è possibile utilizzarli per fini didattici, consentendo agli studenti e alle studentesse di confrontarsi con i processi della ricerca storica. Poter lavorare direttamente con le fonti aiuta studenti e studentesse ad approcciarsi ala storia non come qualcosa da imparare a memoria, ma come una disciplina indispensabile per la costruzione di un pensiero critico che si sviluppa a partire dall'analisi delle fonti, comprendendo che queste sono l'unica rappresentazione della storia nel suo verificarsi e proprio per questo non appartengono al passato ma contribuiscono alla costruzione e alla comprensione del presente.[213]
L'Archivio Storico dell'Istituto Luce è l'unico tra gli archivi audiovisivi italiani ad essere entrato nel Registro Memory of the World dell'UNESCO con la seguente motivazione:

> […] la collezione costituisce un corpus documentario inimitabile per la comprensione del processo di formazione dei regimi totalitari, i meccanismi di creazione e sviluppo di materiale visivo e le condizioni di vita della società italiana. Si tratta di una fonte unica di informazioni sull'Italia negli anni del regime fascista, sul contesto internazionale del fascismo (tra cui l'Africa orientale e l'Albania, ma anche ben oltre le aree occupate dall'Italia durante il fascismo, soprattutto per quanto riguarda il periodo della Seconda Guerra Mondiale) e sulla società di massa negli anni Venti e Trenta del Novecento[214]

Per comprendere come e quanto potersi confrontare direttamente con le fonti primarie muti il rapporto con la storia, rendendola qualcosa da capire prima che da imparare, credo sia interessante fare un breve focus su uno dei gruppi del progetto appena descritto, ossia il gruppo dell'opposizione.
L'Istituto Luce, come è facile immaginare, essendo stato l'organo di propaganda del Partito Fascista e primo veicolo per la costruzione del mito della personalità di Mussolini, al suo interno non ha nessuna fonte primaria riguardante l'opposizione al fascismo. Gli studenti e le studentesse che dovevano lavorare su quel tema si sono trovati e trovate a confrontarsi con l'assenza, che da storici

213 https://luceperladidattica.com/2018/05/08/narrare-la-storia-ricerca-e-creativita-a-cura-di-manfredi-scanagatta-con-gli-studenti-della-vb-liceo-classico-cornelio-tacito-di-roma/ ultima consultazione 03/08/2021
214 http://www.unesco.org/new/en/communication-and-information/memory-of-the-world/register/full-list-of-registered-heritage/registered-heritage-page-6/newsreels-and-photographs-of-istituto-nazionale-luce/ ultima consultazione 09/08/2021

sappiamo essere uno dei motori più potenti per la produzione di domande. E' infatti nell'assenza di fonti riguardanti l'opposizione che si manifesta in modo assoluto la vera essenza di una dittatura totalitaria, è in quel vuoto che sono nate domande complesse che hanno portato tutta la classe a riflettere sulle implicazioni di una dittatura. In quell'assenza non vi poteva essere retorica, non vi era spiegazione o narrazione dello storico, semplicemente le fonti non c'erano, e nel loro non manifestarsi hanno consentito di comprendere la storia senza intermediazioni.

2.8 Archivi fisici e rivoluzione digitale

La digitalizzazione invita a una nuova e innovativa apertura degli archivi, da sempre considerati luoghi complessi che nascondono documenti e informazioni difficili da reperire, luoghi attraversabili e utilizzabili solo da esperti ricercatori capaci di muoversi tra percorsi burocratici e interlocuzioni complesse, grazie ai processi di digitalizzazione possono divenire luoghi non fisici capaci di offrire una nuova esperienza epistemica con le tracce del passato.
In entrambi i sistemi, sia quello fisico sia quello digitale, il ricercatore ha un'interazione con l'ambiente e modifica ciò che è circostante a lui, in base alle sue volontà e alle sue azioni. Se questo è perfettamente comprensibile in uno spazio fisico, più complesso è comprenderlo quando si tratta di una presenza in un ambiente digitale.
Come ci fa riflettere Lorenzo Pezzica, il processo trasformativo che sta coinvolgendo gli archivi e le fonti non è un processo indolore e non trova la comunità degli archivisti in completo accordo su quelli che possono essere i sistemi di riorganizzazione digitale degli archivi e i reali vantaggi di questo processo.

La produzione e la circolazione dell'informazione sul Web scivolano fuori dalle maglie della funzione archivistica e il mito della ricerca globale sulle banche dati e la potenzialità totalizzante delle tecnologie di information retrival ha fatto ritenere a qualcuno superata l'organizzazione archivistica. Per non parlare poi di tutti i problemi relativi al documento digitale e alla sua gestione amministrativa (con tutto ciò che essa comporta); in primo luogo, la compresenza di analogico e digitale e, poi, la firma digitale, la PEC, la validazione temporale, i protocollo informatico ecc. Tutti temi che diventeranno dirimenti tra poco tempo (se non lo sono già), quando ci si dovrà confrontare con l'archivio storico digitale propriamente detto.[215]

215 L. Pezzica, *L'archivio liberato. Guida teoricoì-pratica ai fodni storici del novecento*, Editrice Bibliografica, Milano 2020, p. 70

Per questo Pezzica, riprendendo il pensiero di Paola Cilandri[216] invita gli archivisti a essere resilienti, dunque capaci di modificare il loro modo di intendere l'archivio, osservando le innovazioni date dalla digitalizzazione come una possibilità di rimettere al centro del dibattito un rinnovato interesse per la questione degli archivi, oltre a non farsi limitare «dalla difesa di ciò che si fa da sempre.»[217]

Seguendo questa richiesta di resilienza è interessante analizzare il pensiero di Vitali che nel suo saggio *La ricerca archivistica sul Web* tra le altre cose si sofferma sull'importanza del metodo che deve guidare il ricercatore nel suo approccio all'archivio, dettagliando in cinque mosse le fasi della ricerca archivistica "classica"[218]:

1. individuazione dell'esigenza informativa da soddisfare e formulazione del proprio questionario di ricerca
2. traduzione del questionario storico in un questionario archivistico
3. identificazione dell'archivio o degli archivi che possono dare risposte positive e valutazione della loro effettiva pertinenza
4. individuazione e localizzazione degli specifici documenti da consultare
5. consultazione e interpretazione critica dei documenti

Aggiungendo che

> Oggi alcune di queste mosse e, in taluni casi, l'intero percorso di ricerca possono essere compiuti sul web. Ciò non altera la logica profonda del modello, muta semmai le modalità operative per raggiungere i risultati prefissi e richiede non solo la conoscenza dei meccanismi di funzionamento del web ma anche la capacità di destreggiarsi fra le possibili insidie che esso può nascondere, non lasciandosi troppo sedurre dal miraggio di rapide scorciatoie e di illusorie semplificazioni e non rinunciando ad applicare alla ricerca sul Web in tutti i suoi passaggi, il rigore e l'acribia filologica che lo studio critico del passato sempre richiede.[219]

E' indubbiamente vero che perché il lavoro di ricerca avvenga nel modo corretto le fasi descritte da Vitali debbano rimanere valide anche all'interno di un sistema digitale, ma la distanza e la neo intermediazione digitale che come vedremo è data dai metadati, così come l'assenza fisica del documento, mutano in senso profondo da un punto di vista operativo le azioni del ricercatore o della ricercatrice. Le domande di ricerca, essenziali per procedere, oggi possono essere ridotte a parole chiave da inserire nella barra di ricerca, che il sistema è in grado di leggere e conseguentemente di

216 P. Cilandri, *Lo zen e l'arte della manutenzione archivistica. Ibridi e resilienti*, jlis.it, 9, 2018, 3, pp. 171-179
217 L. Pezzica, *L'archivio liberato. Guida teoricoì-pratica ai fodni storici del novecento*, Editrice Bibliografica, Milano 2020, p. 71
218 S. Vitali, *La ricerca archivistica sul Web*, in a cura di R. Minuti, *Il Web e gli studi storici*, Carocci, Roma 2017, p. 65
219 Ivi, pp. 65-66

proporre un record di risultati molto vasto, ma che a differenza di un tempo potrà essere visionato senza i tempi dettati dal sistema di funzionamento archivistico classico, dunque sarà possibile spostarsi di busta in busta senza avere il limite delle prese giornaliere.
Questi due processi mutano in seguito a due fattori, uno relativo al linguaggio, l'altro ai processi consentiti da una presenza telepistemica.

> l'informatica ha innescato una serie di sviluppi che hanno inevitabilmente portato dei cambiamenti sia nelle pratiche sia nei linguaggi dell'archivistica. Ciò a partire dall'illusione iniziale - dimostratasi in breve tempo impossibile – di operare una semplice esposizione nei nuovi supporti informatici delle prassi e degli strumenti di corredo preesistenti (l'informatizzazione degli inventari storici e delle guide) senza confrontarsi con le esigenze dei nuovi linguaggi del Web.[220]

Se ancora siamo in una fase iniziale di ripensamento dei processi di archiviazione, grazie all'utilizzo di metadati che possono assumere valore semantico e dunque consentire al ricercatore o alla ricercatrice di dare avvio alla propria ricerca utilizzando parole chiave, è già in essere una trasformazione che modifica *de facto* i processi che Vitali individua nei punti tre e quattro: «identificazione dell'archivio o degli archivi che possono dare risposte positive e valutazione della loro effettiva pertinenza; individuazione e localizzazione degli specifici documenti da consultare».
Per individuare un archivio e le fonti che possano essere utili alla propria ricerca ad oggi è possibile agire in modo "semplificato" - fermo restando che si deve avere ben chiaro l'oggetto di ricerca e aver sviluppato in modo compiuto le domande di ricerca - grazie a sistemi a vocazione tematica come SAN[221], SIAS[222] e SIUSA[223], strumenti che si sono posti innanzitutto l'obiettivo di censire, conoscere e poi descrivere di nuovo e in maniera più approfondita il patrimonio documentario, diversificando, ma anche sovrapponendo la propria azione in ragione della tipologia giuridica dei soggetti produttori[224].
La dematerializzazione delle fonti, la trasformazione da artefatti tangibili a oggetti digitali, apre quindi un processo di revisione epistemica della fonte stessa e delle possibilità di fruizione, processo che non si misura solo in ambito teorico, ma che ha delle ricadute profonde nella pratica. La condivisione, la presenza telepistemica e la metadatazione, fanno sì che la ricerca, la fruizione e la comunicazione delle fonti primarie si manifestino oggi in modo assolutamente innovativo rispetto

220 L. Pezzica, *L'archivio liberato. Guida teoricoì-pratica ai fodni storici del novecento*, Editrice Bibliografica, Milano 2020, p. 72
221 http://san.beniculturali.it/web/san/home;jsessionid=64BCBB8C1B69429976A5BA7F84AB157C.sanapp01_portal ultima consultazione 30/07/2021
222 https://sias.archivi.beniculturali.it/cgi-bin/pagina.pl ultima consultazione 30/07/2021
223 https://siusa.archivi.beniculturali.it/cgi-bin/siusa/pagina.pl ultima consultazione 30/07/2021
224 Ivi, p. 73

al passato e ad archivisti e storici credo sia richiesto, in collaborazione con informatici e filosofi dell'informazione e del linguaggio, di collaborare alla creazione di un nuovo sistema di archiviazione e fruizione delle fonti primarie.

Trovo che non sia molto utile soffermarsi su quelli che sono i limiti ad oggi evidenziati dalla digitalizzazione e dalle possibilità di ricerca sul Web, ciò che ci è richiesto è di immaginare, unire le competenze tecniche e scientifiche a dei Livelli di Astrazione capaci di ridisegnare l'idea stessa di fruizione delle fonti, partendo da una attenta critica dell'essenza epistemica della fonte storiografica.

Ripensiamo al post su Facebook della pagina "Autonomia bianconera" il documento riguardante l'ordine di attuazione del Piano E 27, è stato già condiviso in rete attraverso canali informali, è stato già utilizzato per produrre discorso storico da chi storico non è, e a prescindere dalla volontà o meno di storici e archivisti, sempre di più le fonti primarie vengono condivise nel Web. Questi sono dati di fatto e se vogliamo essere in grado di costruire una critica a specifici processi, così come proporre innovativi sistemi di archiviazione e di fruizione, non possiamo che accettare la sfida, le informazioni continueranno a essere veicolate nel Web che noi lo si voglia o meno.

2.9 Archivi, una definizione?

Federico Valacchi nella premessa al suo libro *Gli archivi tra storia e uso futuro* scrive: «Gli archivi hanno molti perché, che arrivano dritti dall'ombra dei secoli e si affacciano ai molti futuri che gli attendono»[225].

Bastano queste poche parole per avere chiaro come confrontarsi con gli archivi in senso concettuale e pratico sia proposito non semplice, soprattutto se chi scrive, come me, non è un archivista, dunque non padroneggia il metodo di ordinamento e conservazione di un archivio e conseguentemente non ha l'esperienza e le conoscenze per poter comprendere a fondo il funzionamento e i significati di un concetto, quello di archivio, che come abbiamo già accennato ha un doppio significato, quello di insieme di documenti e quello di struttura fisica architettonica che i documenti li contiene.

Parlare di archivi è però parlare di storia, lo è perché gli archivi si costruiscono nello svolgersi della storia dando dunque senso concreto all'idea che la storia non sia divenuto, ma bensì divenire. Gli archivi hanno a che fare con la storia non perché gli storici li usano per svolgere le loro ricerche, ma perché è al loro interno che è possibile confrontarsi con ciò che del passato rimane a noi in quanto storia, ossia i documenti. Non solo, gli archivi hanno a che fare con la storiografia, di nuovo non perché aiutano lo storico a scrivere saggi, articoli e libri, ma perché al loro interno si costruiscono narrazioni storiografiche dove però, spiegazione e narrazione, cardini sui quali inevitabilmente si sviluppa i lavoro storiografico, non si mostrano al fruitore, che può essere erroneamente convinto che un archivio non sia niente di più che un deposito dove i documenti stanno, si sedimentano e delle volte ammuffiscono, un luogo di mera conservazione acritica.

Gli archivi non possono essere definiti come deposito perché senza documenti non esistono, non esistono come insieme e non esistono come strutture, una scatola *è* anche senza nulla al suo interno, un archivio *non è* senza documenti.

Sono i fatti a creare l'archivio, ancora prima che le istituzioni, le azioni a determinarne la fisionomia.

Gli archivi vivono delle storie degli enti produttori che hanno creato i documenti, documenti che a loro volta vivono delle storie che raccontano sempre in modo diverso per tipologia e scopo, storie che esistono solo perché altri esseri umani hanno compiuto azioni e sviluppato pensiero. Ma gli archivi vivono anche della storia degli archivisti che con loro si confrontano nella fase corrente, di

225 F. Valacchi, *Gli archivi tra storia e uso futuro*, Editrice Bibliografica, Milano 2020, p. 9

deposito e storica. Gli archivi sono fatti di vincoli, legami, selezioni, gli archivi sono la prova tangibile della complessità umana, tutto è vero, nulla è compiuto.
Come scrive sempre Valacchi nel saggio *La memoria dell'acqua*

> Il termine archivio è infatti intriso di ricordo, così come ogni suo derivato. Qualsiasi molecola di questi complessi organismi, indipendentemente dall'attimo evenemenziale nel quale si manifesta, porta memoria di se stessa e con le sue circonvoluzioni contribuisce a definire un sistema sociale nel tempo e nella sua interezza[226].

de Certeau scrive che lavorare con gli archivi è lavorare con la morte e contro la morte[227] - salvo che per gli archivi correnti - ci si deve immergere in ciò che è stato e che non è più, e lo si fa perché ciò possa essere ancora, in modo nuovo. Gli archivi sono l'unione tra passato presente e futuro perché non smettono mai di modificarsi e di modificare le possibilità del loro racconto e non può esistere tempo e volontà perché ciò possa smettere di essere.
Entrando in un archivio ci si confronta con la morte data dal passare del tempo e lo si fa per ridare la vita, un primo racconto di ciò che è stato perché possa continuare ad essere presente nel nostro presente, perché l'oblio non predomini, non solo sulla memoria, ma sui dati di fatto.
La produzione di documenti è il modo che si ha per affermare la nostra esistenza, per costruire oggetti sociali, per ordinare e organizzare la società, non smetteremo mai dunque di creare archivi perché non smetteremo mai di creare documenti utili da un punto di vista pratico prima e di conoscenza del passato e comprensione del presente poi.
Affascinante ciò che scrive Cristina Baldacci riferendosi all'interesse dell'arte contemporanea - già presente in particolar modo nelle avanguardie del XX secolo[228] - verso la necessità di collezionare, conservare, documentare e archiviare, tanto da chiedersi se l'archivio non possa essere considerato più che un tema o una metafora, un nuovo genere dell'arte contemporanea[229]. Il MAXXI, Museo Nazionale delle Arti del XXI secolo, al suo interno non solo ha i depositi con le opere d'arte contemporanea che fanno parte della sua collezione, ma ha un grande archivio relativo all'architettura del Novecento e, ancora più interessante, un Centro Archivi con tanto di sala di consultazione, un laboratorio sperimentale che, insieme alle tradizionali attività di conservazione, restauro e catalogazione, svolge funzioni di ricerca, promozione e formazione[230].

226 F. Valacchi, *La memoria dell'acqua,* in AA.VV., A cura di, L. Pezzica, F. Valacchi, *Dimensioni archivistiche. Una piramide rovesciata,* Editrice Bibliografica, Milano 2021, p. 173
227 M. de Certeau, *La scrittura della storia*, Il Pensiero Scientifico Editore, Roma 1977, p. 6
228 L. Pezzica, *L'archivio liberato*, Editrice Bibliografica, Milano 2020, p. 26
229 C. Baldacci, *Archivi impossibili. un'ossessione dell'arte contemporanea*, Johan e Levi, Monza 2016
230 https://www.maxxi.art/centro-archivi/ ultima consultazione 22/08/2021

Gli archivi dunque, ben lungi dall'essere solo luoghi di conservazione sono espressione più o meno esplicita di racconti, sono collezioni d'arte, sono portali d'acceso al passato, sono i luoghi fisici o virtuali dove riponiamo le tracce del nostro agire, sia questo fatto in modo ordinato, sia casuale.
Gli archivi possono esser messi in mostra come meravigliose strutture architettoniche, si pensi all'Archivio Centrale dello Stato di Roma o di Milano o come ci ricorda Zanni Rosiello, l'Archivio di Torino, di Venezia, di Genova, di Lucca di Siena[231], e si può mettere in mostra ciò che vive al loro interno, sono luoghi complessi e affascinanti, spaventosi e forieri di incredibili soddisfazioni. Contengono polvere, errori, decisioni, morti, soprusi, vittorie e sconfitte. Negli archivi risiedono sogni e fallimenti, sono il cuore delle mille storie che compongono la storia, quel singolare collettivo che Adelung descrive con i termini *Die Geschichte* riferendosi a «diversi cambiamenti, collegati gli uni con gli altri, che messi in insieme costituiscono un certo intero»[232], gli archivi sono parti di racconti a più voci che in vario modo si incastrano in una sorta di gioco di scatole cinesi[233]. Ci mostrano lettere d'amore e comandi di esecuzione, ci condividono scambi epistolari tra artisti e intellettuali intrisi di emozioni e passioni, portano a noi i più alti sentimenti dell'animo umano così come ce ne mostrano i lati più profondi e oscuri. Gli archivi contengono ciò che nel loro passaggio gli esseri umani lasciano sul mondo terreno, non importa vi sia consapevolezza, ciò avviene, è inevitabile, non si può vivere senza lasciar traccia, basta nascere perché l'universo degli archivi ci accolga al suo interno.
Portando in se una mole incredibilmente eterogenea di informazioni che a loro volta possono essere lette e criticate in modo diverso in base agli occhi che vi si pongono sopra, l'archivio come idea è così accogliente, che è possibile utilizzarne la definizione per chiarire o rendere più esplicite le idee che si intende sostenere[234].
Nel descrivere l'idea di archivio ho parlato di vita, di complessità, di arte, di narrazioni che si sovrappongono, di errori, di processi narrativi, ho messo l'uomo al centro del mio ragionamento sottolineando la difficoltà di coglierne l'essenza in senso assoluto, palesando così se pur sottotraccia, la convinzione che il lavoro dello storico non possa avvenire senza la necessità della spiegazione e della narrazione, ammettendo di conseguenza che il fatto puro a noi non può giungere se non mutuato attraverso un processo di iscrizione[235] e descrizione[236]. Ho parlato di singolari collettivi facendo riecheggiare in lontananza l'idea di una storia dal volto umano, composta inevitabilmente dalle azioni, dalle volontà e dalle emozioni di noi esseri umani, avvicinandomi non

231 I. Zanni Rosiello, *Gli archivi nella società contemporanea*, Il Mulino, Bologna 2009, p. 27
232 R. Koselleck, *Storia la formazione del concetto moderno*, CLUEB, Bologna 2009, p. 29
233 I. Zanni Rosiello, *Gli archivi nella società contemporanea*, Il Mulino, Bologna 2009, p. 30
234 Ivi, p. 23
235 M. Ferraris, *Documentalità. Perché è necessario lasciare tracce*, Laterza, Bari Roma 2019
236 J. Topolski, *Narrare la storia. Nuovi principi di metodologia storica*, Bruno Mondadori, Milano, 1997

senza circospezione alle affascinati e in alcuni frangenti meravigliose suggestioni che Spengler riporta nel suo *Tramonto dell'occidente,* quando parla dell'*animità* della storia[237]. Ho fatto vibrare il diapason che armonizza il pensiero storicista, ma non determinista[238], di autori quali Croce[239] e Dilthey[240] quando affrontano se pur in modo differente l'idea di scienze dello spirito. Quell'idea di singolare collettivo così ben espressa da Koselleck, quando ci spiega il significato di un termine complesso come Geschichte e che inevitabilmente mi riporta alla mente l'idea di come il tutto sia nell'uno, visione tanto cara alla controcultura che si sviluppa negli Stati Uniti nei primi anni sessanta del Novecento[241].

Ho poi, sempre parlando di archivi, accennato alla possibilità di mettere in mostra ciò che al loro interno è contenuto, andando così a toccare uno degli argomenti più cari alla Public History, ma che trova non pochi detrattori, soprattutto tra gli archivisti.

In poche righe e girando sempre e solo introno allo stesso concetto, non ho potuto fare a meno di far emergere molti dei punti di vista che maggiormente mi attraggono quando l'intenzione è di argomentare sui concetti di storia e storiografia.

Ora però non possiamo non tener di conto del fato che gli archivi per esistere sono anche e soprattutto pratica, esistono grazie ad una tecnica, Minuti definisce gli archivi come un «complesso di scritture che, legate da un vincolo naturale, sono prodotte da entità pubbliche o private nell'espletamento delle loro attività, per il raggiungimento di finalità contingenti e per la conservazione della propria memoria»[242].

Luciana Duranti, dopo aver esaminato le definizioni di archivio elaborate da autori diversi che hanno operato in epoche, luoghi e istituzioni differenti, giunge alla conclusione che «da tutte le enunciazioni […] si evincono le medesime caratteristiche: organicità, struttura, sedimentazione naturale, associazione a funzioni e attività e il requisito della tenuta e custodia da parte del produttore», definendo l'archivio come «l'insieme di documenti redatti e ricevuti da una persona fisica o giuridica nel corso delle sue attività come loro strumento e residuo, e conservati per proprio riferimento da quella stesa persona o da un suo successore legittimo»[243].

237 O. Spengler, *Il tramonto dell'occidente*, Longanesi, Milano, 2012, p. 454
238 K. R. Popper, *Miseria dello storicismo*, Giangiacomo Feltrinelli Editore, Milano 2019
239 B. Croce, *Teoria e storia della storiografia,* Adelphi, Milano 2001
240 W. Dilthey, *Esperienza vissuta e poesia*, Il nuovo melangolo, Genova 1999
241 M. Scanagatta, *E l'America creò gli hippie. Storia di una avanguardia*, Mimesis, Milano Udine 2013
242 A. Romiti, *Archivistica generale. Primi elementi*, Torre del Lago, Civita editoriale 2011, p. 128
243 L. Duranti, *Il documento archivistico,* in a cura di, L. Giuva, M. Guercio, *Archivistica. Teorie, metodi, pratiche,* Carocci, Roma 2014, p. 21

La Direzione Generale Archivi[244], DGA, riprendendo in parte la definizione data dal glossario multilingua della terminologia archivistica pubblicato dall'International Council on Archives nel 1988, definisce il termine archivio come:

1. Complesso dei documenti prodotti o comunque acquisiti da un ente (magistrature, organi e uffici centrali e periferici dello Stato; enti pubblici territoriali e non territoriali; istituzioni private, famiglie e persone) durante lo svolgimento della propria attività. I documenti che compongono l'archivio sono pertanto collegati tra loro da un nesso logico e necessario detto vincolo archivistico. In questa accezione si usa spesso la parola fondo come sinonimo di archivio.
2. Locale in cui un ente conserva il proprio archivio.
3. Istituto nel quale vengono concentrati archivi di varia provenienza che ha per fine istituzionale la conservazione permanente dei documenti destinati alla pubblica consultazione.

La cosa affascinante è vedere come nella versione digitale del glossario multilingua della terminologia archivistica pubblicata dall'ICA, si riporti una definizione di archivio differente sia dalla pubblicazione del 1988 sempre dell'ICA sia dalla definizione data dalla DGA.
Se nel motore di ricerca del portale dell'ICA si immette la parola "archivio" vengono prodotti differenti risultati, se si seleziona quello relativo al termine si ottiene questo:
«[Fondo archivistico] L'insieme dei documenti formati o ricevuti e conservati da una persona fisica o da una persona giuridica nello svolgimento delle proprie attività»[245]
Nonostante le differenze, ciò che deve cogliere la nostra attenzione è l'utilizzo dei termini "complesso" e "insieme", è in queste parole che si ritrova uno dei significati più profondi dell'archivio, significati che però hanno un valore differente per lo storico. Per un'archivista è metodologicamente scorretto analizzare un documento a prescindere dal complesso o dall'insieme documentario a cui è legato, è in quella relazione che il documento assume maggior valore; per uno storico e ancor di più per un public historian il documento ha o può avere valore anche singolarmente, nel suo portato informazionale come complessità narrativa. È proprio in questa differenza che si annida molta della diffidenza che gli archivisti hanno nel procedere nella messa in scena degli archivi attraverso l'esposizione di una più fonti prelevate da un fascicolo, rompendo dunque il vincolo archivistico che le lega alle altre; ancor con più diffidenza il mondo degli archivisti guarda alla costruzione di archivi inventati o percorsi tematici in cui le fonti sono riportate senza nessun tipo di vincolo strutturale.

244 http://www.archivi.beniculturali.it/index.php/abc-degli-archivi/glossario ultima consultazione 20/08/2021
245 http://www.ciscra.org/mat/mat/term/1854 ultima consultazione 20/08/2021

Questi temi che stanno in stretta correlazione all'utilizzo di standard e metadati nella costruzione di archivi digitali, verranno affrontati nei paragrafi successivi, ora vorrei riportare quella che forse è la riflessione più sincera sul perché è necessario interrogarsi sugli archivi:
«Gli archivi, perché sono belli...»[246]

2.10 Archivi in trasformazione

Nel capitolo precedente abbiamo osservato in che termini si possa definire il cambiamento che sta coinvolgendo la nostra società attraverso i processi di digitalizzazione.
Ci siamo soffermati sull'analisi dei concetti di dato e informazione per sottolineare come oggi ogni azione che compiamo in un ambito digitale – ambiti che permeano sempre di più il nostro agire – generi un dato che se contestualizzato e unito ad altri dati produce informazione.
Ora sappiamo che nelle società iperstoriche gli esseri umani possono essere definiti inforganismi, il che non li contraddistingue per questo da un passato analogico nel quale comunque si era portati a produrre informazione e lasciare tracce documentarie del nostro passaggio, ciò che muta rispetto al passato è che attraverso l'utilizzo di tecnologie ICT lo scambio informazionale tra inforg. è ormai costante e genera un flusso di dati enorme, ricordiamo che nel 2020 nel mondo sono stati prodotti 64,2 zettabite di dati.
Alla base del processo di archiviazione non vi possono che essere i fatti che a loro volta vengono rielaborati in informazioni alle quali abbiamo accesso grazie a documenti, che appunto generano e compongono archivi. Se questo ragionamento è vero all'interno di sistemi di elaborazione analogica, lo è altrettanto in un ambito digitale, la dematerializzazione del documento però, così come abbiamo già visto, modifica in senso ontologico l'esistenza stessa del documento che non avendo più un struttura tangibile, ma manifestandosi in bit, richiede necessariamente un processo nuovo di archiviazione e degli strumenti digitali per essere fruito. Quando parlo di documenti dematerializzati alludo sia alle fonti nate digitali sia alle fonti materiche che subiscono una dematerializzazione e rimediazione.
Non è difficile comprendere che archiviare dei bit è molto differente rispetto ad archiviare dei documenti materici. La diversità non si afferma in una facilitazione o una complicazione del processo di archiviazione, ma nella necessità di immaginare nuove strutture architettoniche non fisiche ma digitali, capaci di organizzare l'archiviazione dei dati fin tanto che sono in una fase

246 F. Valacchi, *Gli archivi tra storia e uso futuro*, Editrice Bibliografica, Milano 2020, p. 11

corrente dell'archivio, quando quini vengono ancora utilizzati dal soggetto produttore e quando saranno poi trasferiti in nuovi archivi storici, dunque pubblici e fruibili a chiunque.

L'ordinamento di un archivio materico prende forma grazie ad un ordine logico, ma anche fisico, in cui i documenti si distribuiscono e vengono organizzati nello spazio. Anche nell'ambito materico ai documenti vengono assegnati dei codici di protocallazione, e fascicoli e buste vengono indicizzati tramite parole e numeri che descrivono il contenuto e gli anni di riferimento, in qualche modo anche i documenti materici perché possano essere archiviati hanno bisogno di passare per un processo di sovraiscrizione, dove ai complessi documentali, già carichi delle proprie informazioni, ne vengono aggiunte altre, non estranee al documento stesso, ma contestualizzanti.

La rivoluzione digitale che sta coinvolgendo la società e dunque inevitabilmente gli archivi in quanto riflesso documentario del nostro agire, dalla metà degli anni settanta del Novecento ha iniziato a sollevare questioni di natura metodologica e pratica sia nell'ambito della storiografia, con l'avvento delle storie quantitative, sia nell'ambito dell'archivistica che inevitabilmente ha dovuto iniziare a porsi delle domande su come riuscire a costruire nuovi percorsi che tenessero insieme le pratiche ormai consolidate dei processi di archiviazione, con le nuove possibilità date dallo sviluppo di strumenti informatici.

Dell'ambito storico e storiografico ce ne occuperemo nel prossimo capitolo, mia intenzione adesso è provare a far emergere le problematiche e le soluzioni che maggiormente impattano il mondo degli archivi.

Gli archivi, quando istituzionali sono una rappresentazione del potere, l'ordine con sui si cerca di organizzare la società e molto del sapere che essa ha prodotto, l'archivio è dunque uno strumento di controllo e di liberazione e non si può non tenere conto del loro portato politico, per questo nell'ambito dell'archivistica sempre più spesso si parla di *archivistica attiva*, «con questa iniziativa si vuole vedere riconosciuto il ruolo sociale e civile degli archivi, andando oltre la dimensione squisitamente storica. Perché gli archivi sono strumenti di civiltà e l'archivistica è passione civile»[247]. L'intenzione dunque è di portare gli archivi fuori dal concetto unico di bene culturale per dare loro uno spazio nell'azione politica e sociale.

«Niente di rivoluzionario, peraltro. Se, guardiamo bene dentro alle peripezie storiche degli archivi ci accorgiamo di come essi siano sempre stati semplicemente e soprattutto questo, politica, esercizio attivo di diritti e doveri che cristallizzandosi hanno contribuito a fare la storia»[248].

La prima e più semplice questione da cogliere nei processi trasformativi che si stanno generando attraverso la digitalizzazione del dato e dell'informazione è il processo di dematerializzazione, che

247 http://blog.unimc.it/archivisticattiva/pagina-di-esempio/ ultima consultazione 25/08/2021
248 F. Valacchi, *Gli archivi tra storia e uso futuro*, Editrice Bibliografica, Milano 2020, p. 15

impatta inevitabilmente anche sull'altra definizione che si ha di archivio, dunque sulla struttura muraria, architettonica che contiene l'insieme dei documenti.

La storia degli enti e degli ambienti preposti alla conservazione dei documenti, dunque alla strutturazione di archivi, è una storia lunga e complessa, della quale però è necessario fare brevi accenni per meglio comprendere quale sia oggi l'impatto della dematerializzazione, che rendendo libero il documento da una struttura fisica pone nuove domande su come e dove questi documenti "liquidi" possano essere archiviati.

E' necessari tornare agli albori dello Stato post unitario, e alla Commissione Cibreo[249] che nel 1870 si interrogava su come organizzare la gestione degli archivi, la prima domanda alla quale si diede una risposta fu se dare il controllo degli archivi al Ministero degli interni o al Ministero della pubblica istruzione, si decise per il Ministero dell'interno, nonostante alcune perplessità che si radicavano prevalentemente nel timore che se la gestione degli archivi fosse stata data al Viminale si sarebbe perso il valore storico dei documenti a vantaggio di quello amministrativo. I fatti avrebbero dimostrato nel corso dei decenni successivi che questi timori erano infondati e che si potevano governare gli archivi senza tralasciare nessuna finalità.

È nel 1975 che la gestione degli archivi passò dal Ministero degli Interni al neo nato Ministero dei beni culturali, che ancora oggi sovraintende gli archivi, penalizzandone però la dimensione corrente[250].

La prima conseguenza che ebbe questo passaggio di poteri fu l'esclusione dal controllo dello Stato degli archivi prodotti da soggetti diversi da quelli statali pre o post unitari che fossero. Nel progettare il proprio sistema di gestione documentale lo Stato si preoccupò prima di tutto di sé e dei suoi progenitori, per questo molti archivi di notevole importanza a partire dagli archivi comunali e privati, rimasero sullo sfondo di questa nuova trasformazione organizzativa.

Questa azione portò a strutturare un policentrismo determinato dalla gestione di Archivi di Stato disseminati su tutto il territorio nazionale, e sulle ali del metodo storico questo permise una rigorosa applicazione del principio di provenienza che rendeva inscindibile il legame tra i documenti, i contesti politici e territoriali in cui questi si realizzavano. Questo processo legò in modo profondo gli archivi al territorio in cui risiedevano fisicamente.

«È ormai indubitato che le carte per essere meglio intese vanno lette là dove furono scritte. La carta che illustra un monumento è resa più intelligibile dal monumento medesimo; i fatti narrati dove accaddero si fanno come visibili»[251].

249 Commissione Cibreo, *Sul riordinamento degli archivi di stato,* 15 marzo 1870, https://www.icar.beniculturali.it/biblio/pdf/Studi/cibrario.pdf ultima consultazione 20/08/2021

250 F. Valacchi, *Gli archivi tra storia e uso futuro,* Editrice Bibliografica, Milano 2020, p. 37

251 C. Guasti, Commissione Cibreo, *Sul riordinamento degli archivi di stato,* 15 marzo 1870, p. 4 https://www.icar.beniculturali.it/biblio/pdf/Studi/cibrario.pdf ultima consultazione 20/08/2021

E come poter conciliare questa esigenza di stretto legame tra documento, soggetto produttore e territorio, in uno spazio virtuale in cui il “dove” in qualche modo perde centralità per aprirsi ad un “ovunque” all’apparenza così inclusivo?

Il primo tema su cui Valacchi, ma come lui altre come Rossi o Di Marcantonio pongono l'attenzione è di non cadere nella rete di un’euforia tecnolgico digitale, «Talkin’ about revolution… it sound like a whisper»[252] scrive Annalisa Rossi parafrasando Tracy Chapman, così come Giorgia di Marcantonio ci mette in guardia sul non abusare del termine dematerializzazione senza comprenderne a fondo le implicazioni[253].

Queste due affermazioni ci devono far riflettere in prima istanza sul fatto che questa non è la prima rivoluzione che il mondo dell’archivistica vive nella sua storia, ma sicuramente è la più travolgente perché richiede di mutare sia il punto di vista epistemologico sia il punto di vista tecnico. Il rischio è che si inizi ad osservare gli archivi come meri depositi di informazioni slegate non solo dai loro vincoli archivistici, ma anche dal tempo e dallo spazio in cui queste informazioni sono state prodotte, non avere più la necessità di entrare fisicamente in un archivio implica non dover più viaggiare e conoscere il luogo in cui l’archivio si è formato, vuol dire non tenere più il territorio come elemento centrale dell’analisi delle fonti e rischia, secondo alcuni, di offuscare il rapporto di interdipendenza che vi può essere tra un fondo archivistico e l’arte, la politica e la cultura nel quale questo si è formato.

In un mondo globalizzato, dove informazioni, merci e persone, si spostano in modo sempre più rapido e semplice l’idea che gli archivi e le fonti debbano essere necessariamente collegate fisicamente al luogo di appartenenza e non possano uscire dai loro confini, risulta in controtendenza rispetto alla direzione in cui a livello globale, ci sta portando la rivoluzione digitale.

Se si pensa al valore culturale, storico, artistico e informazionale che risiede all’interno di un archivio e se riprendo il discorso fatto nei paragrafi precedenti riguardo alla condivisione, pensare che un ricercatore o una ricercatrice, così come uno studente o una studentessa, per accedere ad un certo tipo di bene culturale debbano essere obbligati a prendere treni aerei e macchine, ecco credo che al di là delle spinte tecnologiche, possiamo immaginare come un mondo affascinante quello in cui è possibile accedere alla conoscenza da ovunque noi ci si trovi, perché questo rende la conoscenza libera dai confini, non appartiene ad uno Stato, un comune o una persona, ma al mondo, ed è confortante pensare che ciò possa essere e che la tecnica oggi ci consenta questo incredibile passo in avanti.

252 A. Rossi, *La palla di vetro. Immaginare la tutela e la comunicazione archivistica diffusa*, in AA.VV., a cura di, L. Pezica, F. Valacchi, *Dimensioni archivistiche. Una piramide rovesciata,* Editrice bibliografica, Milano 2021, p. 56

253 G. Di Marcantonio, *Le “semplificazioni tecnologiche”. Il rischio di processi immaturi*, in AA.VV., a cura di, L. Pezica, F. Valacchi, *Dimensioni archivistiche. Una piramide rovesciata,* Editrice bibliografica, Milano 2021, p. 125

Ma d’altra parte simao tenuti a tenere in considerazioni visioni critiche rispetto a questo processo di internazionalizzazione digitale.

Con buona pace di quanti vedono il processo di internazionalizzazione come ineludibile indicatore di qualità della ricerca scientifica, l’archivistica, fatte salve alcune eccezioni metodologiche di fondo, è destinata da sempre a navigare all’interno di contesti fortemente localizzati, che non si possono né si devono snaturare, pena la loro decontestualizzazione e la conseguente condanna a una genericità anticamera dell’oblio[254].

Credo che questa affermazione vada letta attraverso un doppio livello di analisi, il primo ci impone di chiedersi come l’ente produttore possa mantenere un legame e un vincolo con la propria produzione documentaria anche se questa non risiede nello stesso territorio dell’ente, il secondo invece ci invita ad interrogarci su come organizzare un sistema che consenta l’accesso a tale documentazione, se non nella fase corrente di archiviazione sicuramente in quella sedimentata nella fase storica.
La questione credo la si possa affrontare non interrogandoci sul *dove*, ma sul *chi* e sul *come*.
Un ente produttore sia esso un’ istituzione, un’azienda o una persona, è inevitabilmente legato ad un luogo e ad un territorio, dunque ad un dove, i comuni come istituzioni acquisiscono senso solo in relazione al territorio, le aziende e le società, salvo fatto per i casi in cui si cerca di eludere la pressione fiscale dello stato in cui si risiede, per legge sono legate a Camere di commercio e registri che assegnano all’attività una residenza fiscale e una residenza operativa, così come le persone sono legalmente legate ai luoghi attraverso una residenza o una domiciliazione data da un contratto di affitto o di proprietà. Il *dove* dunque è implicito nell’esistenza dell’ente stesso, è sul *come* e sul *chi* che è necessario porre l’attenzione al fine che non diventino proposizioni eccessivamente “liquide”.
Nella revisione della letteratura archivistica attenta ai mutamenti che il digitale sta imponendo, emerge come problema centrale non tanto una questione di natura tecnica, molte sono infatti le pubblicazioni che mettono in luce come realizzare, organizzare e custodire un processo digitale di archiviazione[255]. Ciò sul quale mi pare ci siano delle difficoltà non taciute, ma espresse, risiede maggiormente nella diffidenza che si genera verso lo sviluppo di algoritmi sempre più in grado di gestire specifici processi, tanto da far parlare alcuni autori di algocrazia, e sull’utilizzo dei metadati come motore computazionale e potenzialmente semantico di processi di archiviazione, vincolo e fruizione digitale del documento.

254 F. Valacchi, *Gli archivi tra storia e uso futuro*, Editrice Bibliografica, Milano 2020, p. 39
255 Sul tema di indubbio interesse vi sono tra gli altri i lavori di S. Vitali, *Passato digitale;* di S. Pigliapoco, *Progetto archivio digitale;* di M. Guerrini, T. Possemato, *Linked data per biblioteche, archivi museiI;* G. Bonfiglio Dosio, S. Pigliapoco, a cura di, *Formazione, gestione conservazione degli archivi digitali;* M. Guercio, S. Pigliapoco, F. Valacchi, *Archivi e informatica*

Prima di andare avanti, anche se non entra a pieno titolo all'interno degli intenti della mia ricerca, vorrei soffermarmi sul termine algocrazia.

2.11 Chi ha paura dell'algocrazia?

La definizione più semplice di algoritmo è quella di un procedimento per risolvere un determinato problema attraverso un numero finito di istruzioni elementari. Il problema può essere più o meno complesso, ma la mente umana ha sempre lavorato suddividendo obiettivi complessi in sotto-passi elementari finiti, per raggiungere lo scopo, anche se magari molto numerosi come per atterrare sulla luna.
Noah Harari sostiene che tutta la biologia si possa descrivere come un insieme di algoritmi prodotti dall'evoluzione delle specie. Le nostre abilità fisiche, la capacità di un virus di adattarsi ad un nuovo ospite, anche le nostre abilità sociali e persino le nostre emozioni sarebbero algoritmi prodotti dalla selezione naturale[256].
Che si voglia accettare o meno la teoria di Harari, nel lavoro di Salvatore Spina troviamo una posizione non troppo differente, quando parlando di archivi digitali si inoltra in delle riflessioni rispetto alla genetica.

> Così mentre gli archivi rappresentano le «banche dati» del passato, e sono testimonianza dell'azione dell'uomo, il futuro sta nel patrimonio genetico, il quale si costituisce come struttura a priori dell'espletamento delle azioni dell'uomo, e nel contempo, con la stessa importanza degli archivi, come dimensione conservatrice del passato[257].

Negli ultimi anni c'è stata una vera e propria fioritura di algoritmi artificiali. Che ne si sia consapevoli o meno, la nostra vita all'interno di società iperstoriche è condizionata e pervasa dagli algoritmi. Gli algoritmi di Spotify, YouTube Music, iTunes, ci propongono quale musica ascoltare, quelli di Netflix, Amazon Prime video, Disney Puls ci suggeriscono quali film vedere, cosa acquistare ci è suggerito dagli algoritmi di Amazon, Zalando, Ebay. I vari servizi di intermediazione delle informazioni, Facebook, Linkedin, Google News, Twitter, come abbiamo già visto, funzionano in base ad algoritmi che ci mettono in evidenza quelle che in base alle nostre interazioni e ai nostri contatti potrebbero essere le notizie più interessanti per noi.

256 N. Harari, *Homo deus. Breve storia del futuro*, Bompiani, Milano2018
257 S. Spina, *Archivi nell'era delle digital humanities dei big data e della genetica,* Algra Editore, Catania 2020, p. 107

L’aspetto più rilevante degli ultimi due o tre anni è che gli algoritmi stanno diventando incredibilmente bravi in quello che fanno. Del resto è incontestabile che, incrociando le nostre scelte passate con le scelte di persone in qualche modo “simili” a noi, gli algoritmi di profilazione riescono a proporci il libro, il film o la canzone giusta spesso meglio dei nostri famigliari e amici. Questo avviene non per magia, ma perché siamo noi a condividere con sistemi informatizzati quelli che sono i nostri interessi. Gli algoritmi non ci conoscono meglio di quanto conosciamo noi stessi, semplicemente hanno contezza computazionale di ciò che ci possono offrire e indagando ciò che noi normalmente cerchiamo o i temi con i quali interagiamo *matchano* le variabili fino ad arrivare ad una offerta discretamente precisa, costruita su un lavoro di profilazione che noi abbiamo acconsentito che loro facessero.

Con gli algoritmi di profilazione abbiamo fatto un salto tecnologico, dall’algoritmo presentato nella definizione iniziale come procedura strutturata e immobile ad un algoritmo di tipo nuovo, che riesce ad imparare. Il machine learning e il deep learning hanno generato una classe totalmente nuova di strumenti, capaci di modificarsi e di evolversi imparando dai dati. Sfruttando i nostri dati e di altre persone simili a noi l’algoritmo impara e si modifica continuamente.

Gli algoritmi possono essere di natura gestionale, dunque utilizzati per sviluppare sistemi complessi, che possono anche essere utilizzati per sviluppare forme di controllo da parte dei governi sui cittadini. La pandemia ci ha messo davanti alla reale compenetrazione di questi sistemi, non tanto in Italia dove *Immuni* o qualsiasi altro strumento di tracciamento si è risolto in un fallimento, ma guardando piuttosto alla Cina e al suo progetto SkyNet possiamo avere un’idea più chiara di come funziona questa tipologia di algoritmi; vi sono poi gli algoritmi “preventivi” e anche qui la Cina ci mostra il loro potenziale di controllo.

Attraverso il riconoscimento facciale che ogni cinese è tenuto a fare ogni volta che acquista una SIM, il governo, che è *de facto* proprietario delle aziende di telecomunicazione, raccoglie una mole enorme di dati biometrici dei propri cittadini e dotando le forze dell’ordine di specifiche tecnologie è in grado di sviluppare un controllo e potenzialmente azioni di arresto preventivo di cittadini che secondo le informazioni utilizzate per realizzare gli algoritmi, possono essere considerati pericolosi. Il caso del controllo e della repressione degli Uiguri, minoranza cinese musulmana, tramite sistemi di intelligenza artificiale e riconoscimento facciale è divenuto un caso internazionale[258].

Dopo questo brevissimo excursus su un tema in realtà incredibilmente complesso che tiene insieme, tecnologia, società, controllo e democrazia, per abbassare anche il livello di distopia che questi argomenti possono sviluppare, credo sia essenziale ricordarsi che gli algoritmi, così come ogni

258 https://www.limesonline.com/rubrica/riconoscimento-facciale-cina-huawei ultima consultazione 25/07/2021

sistema di intelligenza artificiale, anche quando capace di apprendere in autonomia, non hanno nessun potere se tolti dalle mani dell'uomo.

Le intelligenze artificiale così come gli algoritmi non hanno in sé il concetto di razza o di religione, non distinguono tra bene e male, non hanno interesse a profilare noi esseri umani di nascosto, sono governi e imprese a svilupparli, come già detto il loro è un incredibile potere computazionale, ma siamo noi esseri umani a continuare a dire, proprio attraverso la scrittura di algoritmi, come le macchine devono usare il loro potere.

Durante un Workshop[259] tenuto da Krino alla conferenza nazionale dell'AIUCD del 2021, si arrivò a sintetizzare il rapporto tra umano e intelligenza artificiale con una semplice frase:

«Intelligenza artificiale, bias umani».

Con questo voglio dire che è giusto e importante continuare a monitorare gli sviluppi di algoritmi sempre più potenti e capaci di gestire la nostra vita quotidiana, ma è necessario comprendere che ancora una volta questo non è che uno strumento, per quanto molto potente, in mano ad esseri umani e che se usciamo da sistemi dittatoriali come quello Cinese, in cui è il Partito a governare ogni azione del popolo, gli algoritmi e i sistemi di intelligenza artificiale si nutrono delle informazioni che noi stessi li consegniamo, anche se spesso in modo non consapevole o come nel caso di Cambridge Analityca[260], credendo che Facebook non avrebbe venduto a terzi i nostri dati, dati che comunque noi avevamo pubblicato e condiviso.

Lo stesso Vaticano vede nell'intelligenza artificiale un'opportunità incredibilmente affascinante, ma pone l'attenzione sulla necessità di sviluppare un'etica riguardo gli algoritmi[261] e l'etica sicuramente non è di competenza delle macchine.

Nei prossimi paragrafi approfondiremo meglio le procedure tecniche informatiche che sottendono alla realizzazione di archivi digitali dove standard e metadati consento l'accesso e l'interoperabilità alle fonti digitali, ciò che adesso mi interessa è osservare come gli algoritmi che rendono e sempre più renderanno possibile l'archiviazione e l'accesso ad archivi digitali sono algoritmi di tipo gestionale, hanno dunque lo scopo di sviluppare sistemi complessi e di renderne la fruizione semplice per gli utenti. Già così osserviamo il doppio volto dell'algoritmo che da una parte è utilizzato per affrontare questioni organizzative che si stratificano su più livelli di complessità e dall'altra parte deve riuscire a fornire agli utenti una fruizione user friendly, è chiaro che questa parte del lavoro coinvolge anche le interfacce, le quali però non possono che rispondere a necessità di tipo architettonico, che predono forma nel momento in cui i programmatori, insieme ad

259 https://aiucd2021.labcd.unipi.it/programma/ ultima consultazione 25/08/2021
260 https://www.theguardian.com/news/series/cambridge-analytica-files ultima consultazione 27/078/2021
261 https://www.vaticannews.va/it/papa/news/2020-02/papa-francesco-intelligenza-artificiale-algor-etica-discorso.html ultima consultazione 23/08/2021

archivisti, storici e perché no filosofi dell'informazione e scienziati della comunicazione, decidono di realizzare la struttura di archiviazione digitale.

Gli algoritmi che consentono l'accesso e la fruizione di archivi digitali non hanno una funzione differente dai regolamenti e dalle procedure che da sempre è obbligatorio rispettare per una fruizione ordinata e controllata degli archivi materici. Senza dunque inoltrarci nella complessità sociale ed etica che riguarda l'utilizzo degli algoritmi da parte dello Stato per ordinare e controllare specifiche azioni dei cittadini, potremmo pensare che la paura di un'algocrazia che governi il funzionamento di un archivio, non dovrebbe essere maggiore di quella di una burocrazia che già ora lo governa, anzi come accennato gli algoritmi sono pensati anche per facilitare la fruizione da parte degli utenti, cosa che forse non si può sempre dire della burocrazia, che il più delle volte rende le procedure macchinose e complesse, andando a rendere più complicato il lavoro di ricerca.

Il timore espresso da alcuni archivisti riguardo il potere egli algoritmi credo sia riferito in particolar modo alle possibilità espresse dal Web semantico, che grazie alla costruzione di triple, composte da soggetto, predicato e verbo, costruisce una serie di relazioni precostituite dall'azione umana, e che *de facto* son in grado di guidare l'azione di ricerca di documenti che, identificati come oggetti informativi univoci, sono in relazione con altri oggetti informativi, non più all'interno di un sistema strutturato di archiviazione, ma liberi nel Web come dati autoconsistenti.

2.12 Archivi correnti digitali

Per uscire da un piano generale di riflessione e tornare in un ambito maggiormente specifico e attinente al nostro lavoro è da notare che uno dei temi su cui gli archivisti pongono maggiore attenzione è quello relativo alle gestione e organizzazione di archivi correnti digitali.
Ad una prima riflessione si potrebbe dire che la gestione degli archivi correnti non sia un tema di interesse degli storici e delle storiche, questi infatti lavorano su documenti storicizzati che hanno quindi chiuso il ciclo di conservazione e vengono depositatati all'interno di archivi storici.
Per quanto sono convito che sia superfluo è importante specificare che un approccio di questo tipo alla conservazione documentaria è sbagliato se non dannoso.
In precedenza, osservando le diverse definizioni di archivio, abbiamo affrontato, se pur in modo molto rapido, una delle coppie dicotomiche che sottendono al pensiero di Spengler e in realtà più in generarle al pensiero storicista riguardo alla contrapposizione storia-natura o meglio scienze naturali e scienze dello spirito.
In quelle poche righe abbiamo ipotizzato che gli archivi, come la storia, non siano divenuto, ma divenire, ed è proprio nella relazione tra archivio corrente e archivio storico che prende forza questa riflessione, non solo perché l'archivio muta nel tempo, ma perché lo fa coinvolgendo le vite di altri esseri umani oltre a quelle che già porta nei documenti al suo interno.

L'archivistica, come la letteratura, recupera i significati profondi e primitivi della vita dell'uomo e, pertanto, non può escludere il rapporto con la politica e con la storia. […] Riprendendo (e ovviamente sostituendo "archivistica" alla parola "letteratura") quanto scriveva nel 1939 Bo[262]sulla letteratura, va rifiutata l'idea di un'archivistica come illustrazione di consuetudini e costumi comuni, aggiogati dal tempo, quando sappiamo che è una strada, e forse la strada più completa, per la conoscenza di noi stessi, per la vita della nostra conoscenza[263].

Secondo Spengler il divenire si manifesta proprio nella vitalità, che invece si cristallizza immobile nel divenuto, è quel processo inarrestabile che coinvolge noi esseri umani come esseri viventi e che si snoda attraverso decisioni e azioni; ed è su queste azioni che lo storico costruisce l'impianto della propria ricerca, ricerca che coinvolge inevitabilmente anche le azioni degli archivisti che hanno generato l'archivio corrente e degli archivisti che hanno selezionato i documenti per l'archivio storico.

262 C. Bo, *Letteratura come vita,* Venezia, Marsilio 2003
263 G. Locatelli, *Diventare visibili. Quattro tempi per una maieutica archivistica,* in AA.VV., A cura di, L. Pezzica, F. Valacchi, *Dimensioni archivistiche. Una piramide rovesciata,* Editrice Bibliografica, Milano 2021, pp. 75,76

«Per spigare l'agire di un individuo e le sue conseguenze, lo storico ricostruisce il calcolo che l'agente ha compiuto in una data situazione, e dei mezzi necessari per realizzare lo scopo che si prefiggeva»[264]. Queste sono le parole con cui Dray inizia un processo di analisi critica delle fonti ed è riflettendo su queste parole che probabilmente è necessario considerare il lavoro dell'archivista, dunque di colui che insieme all'agente e all'ente produttore consente allo storico di potersi confrontare con i documenti.

«Come al solito gli archivi si trasformano per rispondere ad esigenze reali della società di cui sono prodotto e sostegno»[265]

E' dunque necessario confrontarci anche a livello teorico con le modificazioni che la rivoluzione digitale impone ai processi di costruzione di archivi correnti, prima di procedere però potrebbe essere utile ricordare in che modo è suddiviso il ciclo di vita di un documento, questa la definizione della DGA (Direzione Generale Archivi).

La vita di un archivio si articola in tre fasi: archivio corrente, archivio di deposito e archivio storico.

Per "archivio corrente" s'intende l'archivio che è attualmente in uso ed è in continuo accrescimento; per necessità pratiche, i fascicoli che compongono l'archivio corrente vengono conservati in locali facilmente accessibili o nella stessa stanza degli impiegati che li utilizzano.

I fascicoli relativi alle pratiche ormai concluse non servono più alle attività quotidiane e si possono quindi spostare in locali di minor facile accesso (cantine, soffitte, magazzini decentrati). In questa fase, gli archivi vengono definiti "di deposito" o "intermedi": si tratta infatti di una fase intermedia del ciclo di vita degli archivi, tra quella dell'archivio corrente e quella dell'archivio storico. In questo stadio si conservano documenti che un domani potranno magari essere scartati, ma che per il momento occorre conservare perché possono ancora essere utili (ad esempio per accertamenti fiscali).

Infine, dopo un certo numero di anni – che per gli archivi dello Stato è definito dalla legge in 30 anni – si selezionano i fascicoli che vale la pena di conservare per sempre, e si scartano gli altri. Si costituisce così l'archivio storico, che ha un interesse prevalentemente storico-culturale, ma che comunque può ancora servire anche per fini pratici. Si pensi, ad esempio, che per risolvere dispute sui confini, bisogna a volte ricorrere a documenti vecchi di secoli.

Durante il loro intero ciclo di vita, gli archivi hanno sia un valore pratico-giuridico, che un valore culturale: nella fase iniziale, il primo è senz'altro prevalente, poi mano a mano si affievolisce, mentre si fa sempre più forte il secondo; la valenza giuridica degli archivi e quindi la loro utilità pratica però in molti casi non scompare, anche se passano molti anni[266]

264 P. Rossi, a cura di, *La teoria della storiografia oggi*, Il Saggiatore, Milano 1983, p. x

265 F. Valacchi, *Gli archivi tra storia e uso futuro*, Editrice Bibliografica, Milano 2020, p.

266 http://www.archivi.beniculturali.it/index.php/abc-degli-archivi/conosciamo-gli-archivi/item/196-cosa-s%E2%80%99intende-per-%E2%80%9Carchivio-corrente%E2%80%9D-e-per-%E2%80%9Carchivio-storico%E2%80%9D ultima consultazione 23/08/2021

Come possiamo notare in questa definizione il *dove* è centrale in ogni fase di descrizione, gli archivi correnti devono stare in un locale di facile accesso, quelli di deposito possono essere spostati in cantine e soffitte, mentre quelli storici entrano all'interno di strutture pubbliche, monumenti all'interno di monumenti. Oltre a questo è interessante porre l'attenzione sul processo di selezione che avviene nel passaggio tra l'archivio di deposito e l'archivio storico, questa selezione come vedremo è il primo atto della costruzione di una narrazione, in questa fase è l'archivista che con le sue competenze da avvio ad un lavoro di analisi delle fonti e struttura i vincoli archivistici, così da offrire allo storico e alla storica del materiale già organizzato.

Tornando al *dove* è facile intuire che in un ambiente digitale non ha senso pensare ai documenti che si spostano tra stanze, cantine e monumenti, a ben pensare la definizione data dalla DGA, con una connotazione così legata ad uno spazio fisico, pare davvero non avere nulla a che fare con il concetto stesso di Web, all'interno del quale poi si ritroveranno i documenti digitali e digitalizzati. Si deve invece guardare con la massima attenzione agli archivi digitali che produciamo giorno dopo giorno per molteplici motivi, per l'importanza pratica, come già ci aveva fatto riflettere Zanni Rosiello, ricoperta per i singoli individui che in essi vi vedono raccolti i dati sulle loro vite e i documenti attestanti i loro diritti; per l'impatto che la loro corretta tenuta ha su tematiche, fino a pochi anni fa appena prese in considerazione, quali il rispetto della privacy, il diritto all'accesso ai dati e documenti prodotti, alla loro trasparenza e apertura; infine in qualità di imprescindibile testimonianza/cinghia di trasmissione di valori e ideali a beneficio delle generazioni che verranno dopo di noi.

Nel corso degli anni Novanta si aprì un importante dibattito riguardo a chi spettasse la gestione degli archivi digitali e a quali metodi implementare per renderla strutturata, come ricorda Maria Guercio

> un confronto serrato era allora in corso nella comunità archivistica anglosassone (nordamericana e australiana soprattutto) proprio in relazione alla responsabilità per la conservazione del patrimonio documentario digitale e al modo adeguato (e sostenibile) di esercitarla: da un lato i sostenitori della continuità anche nel mondo digitale della funzione di custodia affidata alle istituzioni archivistiche dedicate (agli archivi nazionali innanzi tutto e in ogni caso a servizi d'archivio identificati con certezza e con compiti e competenze riconosciute all'interno degli enti); dall'altro i fautori di una politica cosiddetta post-custodial che affidava a responsabilità distribuite non specialistiche la sostenibilità della conservazione delle nuove fonti[267].

267 M. Guercio, *Gli archivi come depositi di memorie digitali*, «DigItalia. Rivista del digitale nei beni culturali», III/2 (2008), p. 37, http://digitalia.sbn.it/article/view/280/181, ultima consultazione 25/08/2021

Come scrive Simone Vettore

Tale confronto, anche alla luce dei risultati raggiunti da progetti di ricerca come InterPARES18, si è infine risolto a favore dei primi: la necessità di realizzare specifici sistemi (non necessariamente unici) destinati ad accogliere i documenti digitali lungo il loro intero ciclo-vita, assolvendo dunque a compiti che spaziano dalla gestione documentale alla conservazione nel lungo periodo, ha portato infatti all'individuazione di una serie di requisiti (penso in particolare alla terzietà del soggetto conservatore e alla necessità di seguire procedure affidabili e "tracciabili" rispettose dei principali standard internazionali nel frattempo elaborati) che non potevano essere garantiti con estemporanee soluzioni *in house*[268].

Le amministrazioni pubbliche, salvo rare eccezioni, non si sono dotate di adeguate infrastrutture informatiche destinate a ospitare i propri archivi correnti digitali, motivo per cui assistiamo oggi ad una ricerca di accordi con provider privati. Se aggiungiamo che altrettanto stanno facendo aziende private, professionisti e semplici cittadini, appare evidente come il quadro sia completamente differente rispetto a quello ipotizzato fino a pochi anni fa.
A non farsi trovare impreparati sono i *cloud service provider* privati, i quali stanno implementando soluzioni, definite *cold storage*, per l'archiviazione di lungo periodo; esse, grazie al ricorso a dotazioni tecnologiche robuste e affidabili, permettono di proporre ai potenziali clienti un prodotto dal rapporto spazio di archiviazione/prezzo altamente competitivo[269].
Un esempio di servizio di cold storage è *Glacier* di Amazon e basta leggere la breve descrizione generica presente sul sito per ritrovare tutti gli elementi che Vettore sottolinea nel suo saggio del 2014

Amazon S3 Glacier e S3 Glacier Deep Archive costituiscono classi di storage nel cloud Amazon S3 sicure, durature ed estremamente economiche per l'archiviazione dei dati e il backup a lungo termine. Sono progettate per garantire una durabilità del 99,999999999%, funzionalità di sicurezza e conformità complete in grado di soddisfare anche le normative e i requisiti più severi. I clienti possono archiviare dati a soli 1 USD per terabyte al mese: un risparmio notevole rispetto alle soluzioni in locale. Per mantenere i costi bassi adeguati alle necessità di recupero variabili, Amazon S3 Glacier offre tre opzioni per accedere agli archivi, con tempi che vanno da pochi minuti a diverse ore e S3 Glacier Deep Archive fornisce due opzioni di accesso che vanno dalle 12 alle 48 ore[270].

268 S. Vettore, *Archivi digitali contemporanei. Arsenali di autorità o democrazia?*, in Archivi, 9 (2014), 1, pp. 32, 33
269 Ivi, p. 34
270 https://aws.amazon.com/it/s3/glacier/ ultima consultazione 26/08/2021

Posto che finora sono stati gli utenti a decidere volontariamente di aderire ai servizi erogati in modalità cloud da aziende private, tutto lascia supporre che nel prossimo futuro non ci saranno molti cambiamenti, dobbiamo comunque considerare che i dati di centinaia di milioni, se non di miliardi persone, sono all'interno di server di cui i proprietari sono gigantesche multinazionali alle quali stiamo appaltando la gestione di quantità impressionanti di dati. È il potere computazionale, ancora una volta, così come già descritto nel primo capitolo a manifestarsi come il vero nuovo centro di forza politica ed economica.

I data center di realtà come Amazon, Google, Facebook, Apple, ma anche nomi meno noti come Carpathia[271], Cogent[272], OVH[273], Digital Realty[274], hanno capacità di archiviazione veramente impressionanti e finiscono con il contenere le "esistenze digitali" di centinaia di milioni di persone,

su questo tema è di grandissimo interesse un articolo pubblicato sulla pagina engineering di Facebook, di cui riporto un estratto.

«Due miliardi di foto vengono condivise ogni giorno sui servizi di Facebook. Molte di queste foto sono ricordi importanti per le persone su Facebook e la nostra sfida è quella di assicurarci di poter conservare questi ricordi per tutto il tempo che le persone vogliono in un modo che sia il più sostenibile ed efficiente possibile. Poiché il numero di foto ha continuato a crescere ogni mese, abbiamo visto un'opportunità per ottenere efficienze significative nel modo in cui conserviamo e serviamo questi contenuti e abbiamo deciso di correre con esso. L'obiettivo era quello di fare in modo che le foto degli anni passati fossero accessibili tanto quanto l'ultimo popolare meme sui gatti, ma che occupassero meno spazio di archiviazione e usassero meno energia. Le foto più vecchie, e quindi meno popolari, potevano essere memorizzate con un fattore di replicazione inferiore, ma solo se eravamo in grado di mantenere una copia aggiuntiva, altamente durevole, da qualche altra parte.

Invece di cercare di utilizzare una soluzione esistente - come le enormi librerie a nastro - per adattarla al nostro caso d'uso, ci siamo sfidati a rivedere l'intero stack da cima a fondo. Siamo fortunati in Facebook - siamo autorizzati a ripensare i sistemi esistenti e a creare nuove soluzioni ai problemi tecnologici. Con la libertà di costruire un sistema end-to-end interamente ottimizzato per noi, abbiamo deciso di reimmaginare l'edificio stesso del data center convenzionale, così come l'hardware e il software al suo interno. Il risultato è stato un nuovo data center basato sullo storage costruito letteralmente da zero, con server che si accendono a seconda delle necessità, gestito da un software intelligente che verifica e riequilibra costantemente i dati per

271 https://www.crunchbase.com/organization/carpathia-hosting ultima consultazione 28/08/2021

272 https://www.cogentco.com/en/products-and-services/success-stories/717-cloudhelix ultima consultazione 28/08/2021

273 https://www.ovhcloud.com/it/bare-metal/?xtor=SEC-13-GOO-[it_int_2020_ovh_baremetal_undefinite_sale_acquisition_srch_defensive()]-[266448374853]-S-[ovhcloud]&gclid=Cj0KCQjwvaeJBhCvARIsABgTDM5hHOmhbYYn5UQrEXwpSI2xYmbmi0kcuvm22HMkTbXdEMky1iPlgG8aAlqyEALw_wcB ultima consultazione 28/08/2021

274 https://www.digitalrealty.com/data-centers/london/cloud-house-1-muirfield-crescenthttps://www.digitalrealty.com/data-centers/london/cloud-house-1-muirfield-crescent ultima consultazione 28/08/2021

ottimizzare la durata. Abbiamo costruito una nuova struttura che utilizza una quantità relativamente bassa di energia ma ha un sacco di spazio. I data center sono dotati di meno di un sesto della potenza disponibile per i nostri data center tradizionali e, a pieno carico, possono supportare fino a un exabyte (1.000 PB) per sala dati. Dato che la densità di archiviazione generalmente aumenta con il progredire della tecnologia, questa è stata la linea di base da cui siamo partiti. In altre parole, c'è molto spazio per crescere»[275].

L'articolo è del 2015 e gli ingegneri parlano in ordine di exabyte e non di zettabite, le parole finali di questo estratto paiono in questo senso assolutamente profetiche, sapiamo infatti che la produzione di dati è aumentata esponenzialmente negli ultimi anni.

Ma alla fine cosa sono questi server che si accendono solo nel momento della necessità, che consumano meno energia ma che sono capaci di archiviare al loro interno fonti, documenti, che non hanno più un utilizzo corrente da parte del soggetto produttore, ma che non si può pensare di cancellare perché comunque "ricordi" che per gli utenti è piacevole mantenere? Ciò che gli ingegneri di Facebook hanno realizzato non è niente di diverso da un archivio di deposito, dunque un inseme di documenti che non hanno più un utilizzo o un'interesse quotidiano per l'utente/soggetto produttore, ma che mantengono un valore documentale. Questi server a basso consumo di energia sono le vecchie cantine e soffitte, sono i nuovi luoghi del deposito prima che si arrivi ad una selezione che conduca i documenti dentro ad archivi storici, ma attenzione, in questo specifico caso, per accedere a quella cantina l'unica chiave che ogni utente di Facebook ha bisogno di avere è di essere tra i contatti del "proprietario" di quella cantina. Ciò che emerge e che abbiamo già potuto osservare in relazione al portale del Parlamento Europeo, è che con la digitalizzazione dell'informazione, le fonti vivono un'accessibilità nuova anche quando risiedono in archivi correnti e di deposito e ancora più incredibile è pensare che ognuno di noi, se iscritto a WhatsApp, Facebook, Messanger o Instagram, per citare solo i social network e i programmi di messaggistica di proprietà di Zuckemberg, ha i suoi dati archiviati in archivi correnti e di deposito, perché e come questi dati possano entrare in archivi storici è una domanda alla quale temo ancora non si possa dare risposta.

Quando pensiamo all'archiviazione di dati, quando ci domandiamo come si possano gestire archivi correnti e di deposito digitali, dobbiamo considerare che questa è la realtà del nostro contemporaneo e che difficilmente vedremo una inversione di trend, è chiaro che è sbagliato fare di tutta un'erba un fascio, le istituzioni pubbliche, così come gran parte delle aziende, non producono la quantità di dati degli utenti di Facebook, ma lo storico del futuro potrà, o più probabilmente dovrà, confrontarsi anche con questa tipologia di archivi.

275 https://engineering.fb.com/2015/05/04/core-data/under-the-hood-facebook-s-cold-storage-system/ ultima consultazione 23/08/2021, Tradotto con www.DeepL.com/Translator (versione gratuita)

Come abbiamo già scritto nel primo capitolo, il problema è quantitativo, ma la risposta se davvero crediamo che gli archivi, così come la storia, siano composti dalla complessità dell'umano, non potrà che essere qualitativa, perché come scrive Ferraris «un tempo l'oblio era la regola e la traccia l'eccezione, oggi vale l'opposto»[276].
L'importanza di sviluppare degli standard di gestione e di metadatazione che consentano l'interoperabilità tra fonti digitali conservate in data center differenti, questione che approfondiremo nei prossimi paragrafi, vibra di intensità se all'articolo di Facebook noi affianchiamo le parole di Vettore che nel 2014, se pur riferendosi alla gestione da parte delle istituzioni italiane dei dati digitali prodotti, dunque abbandonando la visione globale, si poneva delle domande fondamentali per garantire un accesso e una gestione libera e trasparente dei dati presenti nei diversi data center

Limitando qui il ragionamento agli archivi delle sole pubbliche amministrazioni le problematiche che insorgono riguardano perlomeno due aspetti: il primo è relativo all'interoperabilità tra i sistemi di storage/archiviazione del provider e quelli dell'istituto di conservazione che dovrebbe ricevere il versamento: ipotizzando una ripartizione dei compiti in cui i fornitori privati gestiscono gli archivi nella loro fase attiva/corrente, mentre gli archivi pubblici si fanno carico di creare repository per la conservazione di lungo periodo, sarebbe infatti opportuno studiare per tempo protocolli che consentano il "dialogo" tra i vari data center, nonché idonee modalità e procedure di versamento[277].

La gestione degli archivi correnti si mostra dunque come una questione dirimente per la strutturazione di archivi storici digitali e conseguentemente per il lavoro degli storici e delle storiche del futuro. Credo sia necessario chiedersi come si possano gestire i dati digitali oggi prodotti da istituzioni e privati; come sia possibile costruire infrastrutture che garantiscano un sistema di archiviazione uniforme; quali siano i professionisti che si devono incaricare di "disegnare" queste infrastrutture e per ultima una domanda rivolta agli storici e alle storiche, vista l'accessibilità, avremo noi la pazienza di far storicizzare queste fonti?
Dopo aver compreso che probabilmente non sia utile problematizzare un *dove* si archivino fisicamente le fonti digitali in un' ottica corrente di archivio, visto che l'essenza digitale delle fonti consente una fruizione di queste da qualsiasi luogo ove vi sia una connessione ad internet, rimangono ancora aperte, se ben già problematizzate, le questioni del *chi* e del *come.*
Rispondere a queste domande, o meglio provare a proporre delle indicazioni, richiede di approfondire in modo attento differenti passaggi tecnici che afferiscono direttamente a ciò che oggi viene definita archivistica informatica e che ha una diretta correlazione con l'ingegneria

276 M. Ferraris, G. Paini, *Scienza nuova.,* Op cit. p.22
277 S. Vettore, *Archivi digitali contemporanei. Arsenali di autorità o democrazia?*, in Archivi, 9 (2014), 1, p. 33

informatica, ma prima di fare un passo nella tecnica, per avere un quadro completo della situazione dobbiamo confrontarci con il tema delle fonti materiche presenti in archivi fisici, che se non digitalizzate rimarranno esterne ai funzionamenti e ai processi di scambio informativo delle società iperstoriche, rischiando così di creare un gap profondo tra il contemporaneo, la storia futura e la storia passata.

Se gli archivi, come abbiamo più volte ribadito non sono solo conservazione di documenti e di beni culturali, se hanno un ruolo sociale e politico, se il poterli consultare può modificare il livello di consapevolezza sociale riguardo la storia e dunque il presente, è folle immaginare che il patrimonio documentario giunto a noi dal passato, non ci segua in un futuro in cui accesso come concetto sarà molto più legato all'utilizzo di un device digitale che ad un portone.

2.13 Medium-documento

La costruzione di archivi digitali si presenta come fenomeno capace di mutare l'approccio alla ricerca storiografica, di favorire la narrativizzazione della storia attraverso la messa in scena delle fonti, in un' ottica di Public History[278], e di aprire il mondo degli archivi e delle informazioni in essi contenute a chiunque abbia una connessione internet.

Il processo di digitalizzazione delle fonti storiche d'archivio comporta diversi livelli di riflessione, il primo, più immediato da comprendere come abbiamo visto riguarda la dematerializzazione della fonte, che per divenire digitale però può essere sottoposta a due processi non equivalenti.

Uno è di riscrittura e ricodificazione del testo attraverso l'utilizzo di metalinguaggi di marcatura come XML o SGML[279], l'altro, più interessante e sicuro da un punto di vista storico scientifico, è di ricodificazione e rimediazione della fonte attraverso strumenti di scansione digitale e l'utilizzo di metadati. In entrambi i casi si assiste ad un processo di riscrittura o sovrascrittura attraverso linguaggi differenti da quello originariamente presente nella fonte, non sempre intellegibili al fruitore, ma necessari per rendere eseguibile ad ogni livello l'interazione uomo macchina[280].

La rimediazione del documento attraverso scansione ha come limite l'impossibilità di indagare corpora linguistici attraverso l'analisi che consentono i metalinguaggi di marcatura[281], ha dunque un

278 M. Scanagatta, *Publci historian, ricerca e azione creativa*, in L. Bertucelli, P. Bertella Farnetti, A. Botti, Publci History. Discussioni e Pratiche, Mimesi Edizioni, 2015, p. 315, 331

279 S. Vitali, *Passato digitale. Le fonti dello storico nell'era del computer,* Bruno Mondadori, Milano, 2004, p. 54

280 Rispetto all'esplosione della scrittura e riscrittura legati alla iscrizione e alla conservazione della memoria in ambito digitale, Maurizio Ferraris, Germano Paini, Op Cit., p. 20

281 F. Meschini, *Frammenti e connessioni, il WorldWidWeb come ambiente informativo e il ruolo delle Digital Humanities*, in Offcine della storia, n° 21, Maggio 2020,

limite nella portata d'analisi quantitativa e linguistica, ma rende il documento immodificabile se non nella fase di messa in scena, dove l'autore può volontariamente modificare l'immagine, che rimane però identica all'originale nel deposito archivistico digitale.

Se la documentazione di cui lo studioso è in cerca, è stata oggetto di riproduzione digitale, la consultazione e l'interpretazione critica dei documenti, possono svolgersi anche sul Web. Ciò, può essere fatto proficuamente a condizione che non si ignori la specifica natura degli oggetti che si mostrano sullo schermo del computer.

I documenti digitali e i sistemi che li rendono disponibili sono infatti prodotti complessi e stratificati che anche quando riproducono oggetti materici non ne costituiscono mai un semplice doppio.[282]

La loro migrazione su un diverso formato e la loro consultazione in un diverso ambiente modificano infatti le modalità con le quali essi possono essere percepiti e interpretati, nonostante si possa pensare il contrario, il digitale non riduce il livello delle mediazioni che si interpongono fra ricercatore e i documenti, ne cambia semmai la natura e gli effetti.

Per non accostarsi con sguardo ingenuo alle fonti archivistiche digitalizzate occorre essere consapevoli delle stratificazioni di cui esse sono intessute e nel farne uso soppesarne alcuni caratteri che possono costituire altrettanti elementi di valutazione sull'affidabilità dei progetti e dei sistemi che le rendono accessibili.

La digitalizzazione apre alla possibilità di un accesso molto più rapido alle fonti, permettendo anche a chi non si è mai confrontato con un archivio classico di procedere con una ricerca utilizzando il Web come strumento, permette allo storico di poter accedere da remoto alle fonti e di poterne avere una fruizione più attenta dal momento che le tecnologie attuali di scansione consentono di avere una risoluzione ad altissimo livello del documento originario, il che permette per esempio, di zoomare l'immagine così da poterla osservare e leggere in un modo che la fonte originale probabilmente non ti permetterebbe. Non si può però non considerare che la distanza con la quale è possibile fruire un documento digitale, se da un parte apre ad una serie di opportunità di grande interesse, dall'altra mette il ricercatore e la ricercatrice in un rapporto completamente novo con la fonte in prima istanza perché non si instaura più una relazione fisica.

Sul valore della fonte di archivio come testo portatore di più significati e sul conseguente rapporto che si instaura tra la fonte e chi la sta studiando è interessante una riflessione di Topolski, secondo cui «le fonti non sono soltanto una raccolta di informazioni, ma giocano nello stesso tempo un ruolo di ispirazione non verbalizzata, di visualizzazione, offrendosi anche come una base di esperienza

https://www.officinadellastoria.eu/it/2020/05/07/frammenti-e-connessioni-il-worldwideweb-come-ambiente-informativo-e-il-ruolo-delle-digital-humanities/

282 S. Vitali, *La ricerca archivistica sul web*, in, a cura di Rolando Minuti, *Il web e gli studi storici*, Carocci, Roma, 2015

diretta del passato, di contatto diretto con il passato a livello di sensazione e vissuto».[283] Vitali sostiene che c'è una dimensione del rapporto con le fonti che non si esaurisce nella quantità di informazioni che da essa è possibile ricavare, ma che investe la sua qualità profonda, il complesso delle sensazioni dei meccanismi cognitivi - compresi quelli percettivi estetici - che esse sono in grado di evocare e mettere in moto e per attivare i quali l'esperienza del contatto diretto con la mentalità dei documenti continua a rimanere insostituibile[284].
Riguardo l'esperienza e la relazione tra storico e fonte non di minore interesse sono le riflessioni portate da Dilthey, sostiene che

> lo storico giunge a conoscere il passato iniziando da documenti e dati che in se stessi no lo rivelano. Questi dati, egli replica, gli offrono solo l'occasione per rivivere nella propria mente l'attività spirituale che originariamente li produsse. È in virtù della propria vita spirituale e in in proporzione all'intrinseca ricchezza di quella vita, che egli può così infondere vita a materiali morti ai quali si trova di fronte. Così la genuina conoscenza storica è un'esperienza interiore (erlebnis) del proprio oggetto, mentre la conoscenza scientifica è il tentativo di comprendere (begreifen) fenomeni presentatigli come spettacoli esterni.[285]

E' chiaro che il pensiero di Topolski e Vitali non possa essere applicato ad una società in cui i processi informativi sono ormai quasi completamente digitali e dunque lo storico del futuro che voglia o meno non si confronterà mai, se non in rarissimi casi, con una documentazione di tipo materico. Il loro pensiero va dunque analizzato criticamente per coglierne spunti utili nell'immaginare i processi di fruizione di fonti digitali a partire da quelle digitalizzate.
Ciò su cui possiamo porre l'attenzione è il valore dei documenti non solo per le informazioni che portano, ma per come le portano. Se è vero che ogni documento veicola delle informazioni è altrettanto vero che ogni documento lo fa in modo differente e anche in queste differenze è possibile coglierne il significato storico e narrativo.
Solo per fare alcuni esempi, ma la lista potrebbe essere lunghissima, un codice medievale è differente da un comando militare, così come questo è differente da un telegramma, che a sua volta è differente da un documento di nascita, ancora diverso da un documento di schedatura di un antifascista o dall'invito Reale ad una cena di gala, per non parlare di quanto questo possa essere differente da un documento amministrativo.

283 J. Topolski, *Narrare la storia. Nuovi principi di metodologia storica*, Bruno Mondadori, Milano, 1997 p. 57
284 S. Vitali, *La ricerca archivistica sul web*, in, a cura di Rolando Minuti, *Il web e gli studi storici*, Carocci, Roma, 2015 p. 95
285 R. G. Collingwood, *il concetto della storia*, Fabbri Editori, Milano 1966, ed. or. The idea of history, clarendon press, oxford 1946, p. 186

Questi documenti non si differenziano solo nel contenuto che portano, ma per la forma in cui sono redatti, dal fatto che siano scritti a mano o con una macchina da scrivere, dal colore della carta, dalla presenza o meno di miniature o di appunti manoscritti e non in ultimo dal rapporto cognitivo e emotivo, esperenziale direbbe Dilthey, che si può sviluppare con la ricercatrice o il ricercatore. Quando ci troviamo in mano un documento materico, prima ancora che dalle informazioni in esso riportate la nostra attenzione viene catturata dalla forma del documento, lo osserviamo nella sua interezza cercando di cogliere tutti gli aspetti che lo denotano. La rimediazione tramite scansione digitale ci assicura che il documento digitalizzato, se ben ontologicamente differente da quello materico, ne riproduca in modo esatto ogni proprietà, tranne peso e odore.

Se ritrascrivere il contenuto di un documento a stampa in formato digitale implica la riscrittura "umana" dell'intero testo con l'aggiunta di metalinguaggi di marcatura, privando così il documento del suo contesto originario[286] e prestando il fianco ad errori umani di trascrittura fino alla possibile modificazione del testo se il file non è protetto in maniera adeguata, la digitalizzazione per scansione del documento consente la rimediazione di ciò che vorrei definire come medium-documento, cercando con questo termine di affermare il valore dell'oggetto documento come medium che veicola un messaggio e delle informazioni.

Le tecnologie già oggi a nostra disposizione consentono delle riproduzioni digitali dell'artefatto materico di altissima definizione, così da poter effettuare degli zoom in, che ci consentono di osservare i documenti digitalizzati da prospettive fino ad oggi neanche immaginabili, non solo, i metadati e i linguaggi di marcatura aggiungono informazioni al documento digitalizzato e ci consentono di navigarlo come fosse un testo multimediale.

E' chiaro che questi processi richiedano una sovrascrittura di dati, il documento, che ai nostri occhi appare in modo identico rispetto all'originale è in realtà composto da una stratificazione di livelli informativi.

Da storico e da public historian che si interessa del valore epistemologico delle fonti e della possibilità di metterle in scena attraverso processi di narrativizzazione in cui la fonte diviene parte intelleggibile del discorso, ritengo, in accordo con gran parte della comunità scientifica, che la rimediazione tramite scansione digitale sia la pratica in assoluto più interessante, ma conodivido anche l'interesse di filologi, storici e linguisti nel considerare la riscrittura con linguaggi di marcatura dell'intero testo.

286 E. Pierazzo, *Edizioni digitali di documenti complessi,* in ciclo di seminari di informatica umnistica, *La tecnologia informatica applicata alle scienze filologiche e librarie,* Scuola di Alta Formazione A. Varvaro, Unina, https://www.youtube.com/watch?v=IbIySwYOsgM&feature=share&fbclid=IwAR1nDi5RiaNDFeI4BlFNRfvpfiVmgd_D69oCvXga4EzGG0QC4jOWOWI5r_4

Questo procedimento consente infatti di produrre delle edizioni critiche messe in relazione diretta con il documento stesso e di procedere con un'analisi di tipo linguistico sui corpora.

Sempre più spesso è possibile confrontarsi con progetti che sviluppano entrambi i livelli di digitalizzazione, da un parte la scansione digitale del documento materico, dall'altra la trascrizione con linguaggi di marcatura.

Molti sono i progetti dedicati alla pubblicazione di determinate tipologie documentarie, come pergamene, mappe, piante, fotografie, audiovisivi, che richiedono spesso anche specifici tracciati descrittivi. Su tutti vorrei parlare del progetto del Codice Pelavicino, sviluppato dal Laboratorio di Cultura Digitale dell'Università di Pisa, su iniziativa di Enrica Salvatori, già da tempo membro di un gruppo ricerca dedicato all'edizione critica del Codice Pelavicino e ora responsabile scientifico del progetto e di una corposa equipe. Grazie al software EVT, open source, sviluppato da Roberto Rosselli Del Turco, è stato possibile proporre all'Archivio Capitolare Lunense, presso cui il Codice è conservato, un'edizione digitale che da un lato consentisse lo studio del manoscritto senza la sua consultazione fisica e dall'altro presentasse tutte le caratteristiche – potenziate – di una tradizionale edizione critica cartacea. L'edizione digitale veicolata dal Web consente agli utenti di interagire con coloro che ne stanno curando l'edizione, suggerendo modifiche e correzioni o chiedendo chiarimenti[287].

Come si legge in un articolo pubblicato su «Umanistica Digitale»[288]

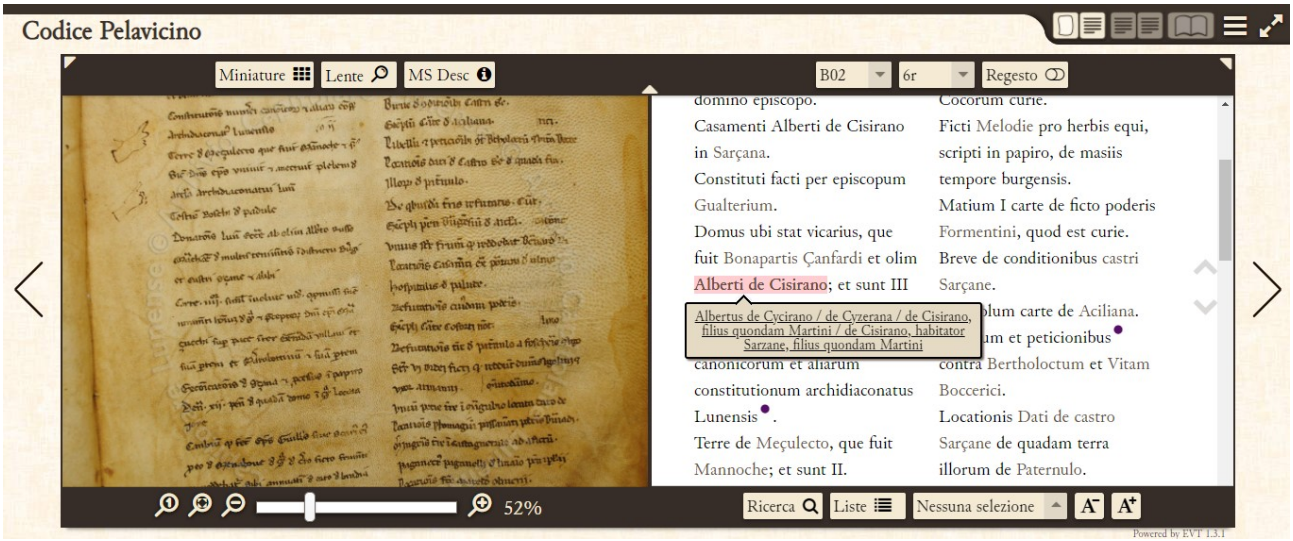
> Le edizioni digitali di manoscritti possono essere di diversa natura; in questo caso è stato scelto per il Pelavicino di realizzare un'edizione basata su immagini.[289] L'idea di una digitalizzazione sistematica dei manoscritti antichi, con il duplice scopo di tutelare i supporti originali e fornire l'accesso a documenti non altrimenti consultabili, costituisce ormai una priorità condivisa.[290] La semplice riproduzione dell'immagine di un manoscritto non costituisce, ovviamente, una vera edizione critica: perché questo avvenga l'utente deve essere messo in grado di accedere all'immagine, alla trascrizione e a tutto l'apparato critico considerato necessario dalla comunità scientifica. Nel caso specifico del Codice Pelavicino, ogni pagina del testo trascritto e annotato viene collegata alla sua rispettiva immagine, in modo da fornire anche la consultazione del volume nella sua forma originaria. Tramite una idonea codifica del testo tutte le caratteristiche e gli strumenti delle edizioni digitali tradizionali sono state poi integrate per garantire un prodotto che intende superare i limiti intrinseci di una pubblicazione cartacea senza perdere di validità scientifica.[291]

287 http://www.labcd.unipi.it/progetti/codice-pelavicino-digitale/ ultima consultazione 27/08/2021
288 https://umanisticadigitale.unibo.it/article/view/7232/6991 ultima consultazione 27/08/2021
289 U. Mazzini, *Correzioni critiche di alcune date del Regesto del Codice Pelavicino per il socio Ubaldo Mazzini.* Genova: Società ligure di storia patria, 1914
290 S. Noiret, *Public History' e 'storia pubblica' nella rete*, Ricerche storiche, 39, num. 2-3 (2009): 275–327
291 E. Pierazzo, *Digital Scholarly Editing: Theories, Models and Methods*. Farnham, Surrey, England: Ashgate, 2015

L’equipe di ricercatori non si è dunque limitata alla rimediazione per scansione del documento, ma ha attuato una codifica dei testi in XML, un noto linguaggio di *markup* di tipo descrittivo, considerato dall’attuale e vasta comunità scientifica impegnata nell’edizione digitale di testi storici il miglior metodo attualmente a disposizione a tale scopo. Basandosi su questo linguaggio la Text Encoding Initiative (TEI)[292] ha sviluppato e continua a implementare una serie di linee guida per la codifica di testi in ambito umanistico. Le caratteristiche specifiche dei testi da codificare hanno reso necessario la preparazione di uno schema di codifica che, pur basato sullo standard TEI e pienamente compatibile con gli schemi globali, permettesse di includere solo gli elementi XML effettivamente necessari. Questa scelta pone l’edizione del Codice Pelavicino all’interno di una nutrita e attiva comunità di filologi e storici digitali e favorisce, proprio grazie all’uso di uno standard condiviso, ma continuamente rielaborato, che vi siano apporti ulteriori al miglioramento delle stesse linee guida o alla enucleazione di problematiche prima trascurate.
Il lavoro svolto dall’equipe coordinata dalla Professoressa Salvatori è veramente di altissimo livello e consente molte e diverse interazioni con il documento:

- è possibile evidenziare per ogni documento, con colori diversi, gli elementi più rilevanti, quali le date, i nomi di persona e di luogo, gli enti religiosi, i mestieri/ruoli, le monete; questa caratteristica (ingl. *named entities*) ha comportato importanti modifiche alle funzionalità esistenti in EVT, tra cui l’estrazione automatica di tutte le entità codificate nel testo e la gestione all’interno dell’interfaccia principale di un filtro per selezionare quali tipologie di entità evidenziare sul testo
- si può effettuare una ricerca per testo libero oppure si possono consultare le liste e gli indici predisposti
- si possono evidenziare tutti i *signa tabellionum* del codice, ovvero le riproduzioni dei timbri notarili, sia nel sistema di visualizzazione, sia nel catalogo appositamente predisposto sul sito web[293]

292 https://tei-c.org/ ultima consultazione 27/08/2021
293 E. Salvatori, R. R. Del Turco Et. al, *Il codice Pelavicino tra edizione digitale e Public History,* https://umanisticadigitale.unibo.it/article/view/7232/6991 DOI: https://doi.org/10.6092/issn.2532-8816/7232,ultima consultazione 27/08/2021



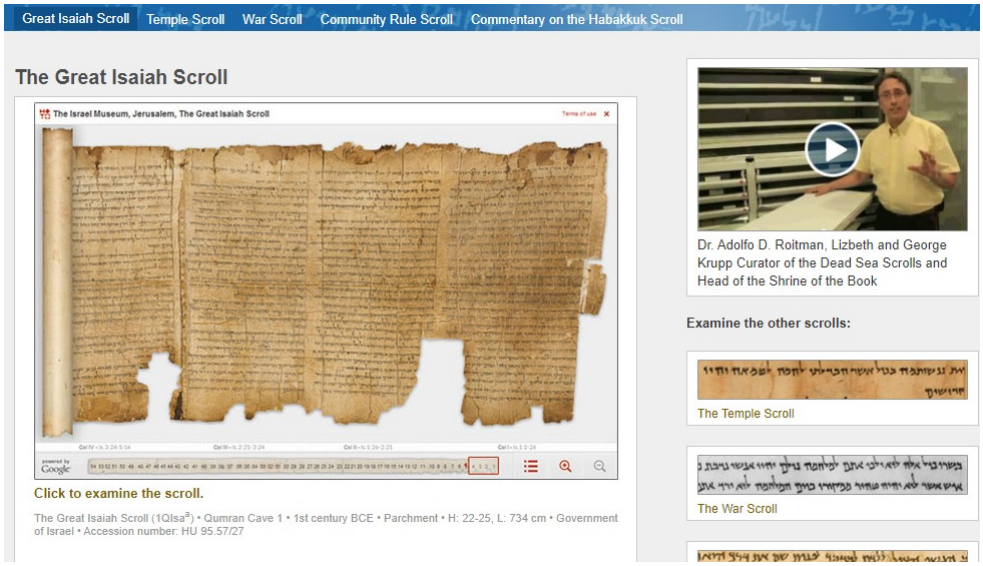
Questo lavoro ci mostra le possibilità che la digitalizzazione offre nello studio di un documento, non si tratta più solo di accessibilità. L'integrazione di differenti sistemi tecnologici, in questo caso la scansione e la codifica con linguaggi di marcatura, consente una fruizione del documento in modo più consapevole anche per chi non avrebbe le competenze per poterlo analizzare criticamente senza un supporto esterno. La neo-interemediazione digitale si manifesta qui in modo chiaro, il documento non è abbandonato ad un disordinato flusso informativo, ma anzi è iper contestualizzato, valorizzandone il portato contenutistico e il valore materico.

Un altro progetto che ritengo meriti di essere analizzato è l'edizione digitale dei Rotoli del Mar Morto, i circa 900 manoscritti, rinvenuti in undici grotte a Qumran, una località vicina a Gerico in Medio Oriente, e oggi conservati in diverse strutture tra cui il Museo di Israele. È questo l'ente, che in collaborazione con Google ha dato avvio ad un incredibile processo di digitalizzazione dei documenti che oggi sono fruibili sul Web, come si legge sul sito

Il Museo d'Israele ti dà il benvenuto al Dead Sea Scrolls Digital Project, che permette agli utenti di esaminare ed esplorare questi antichissimi manoscritti del Secondo Tempio ad un livello di dettaglio mai raggiunto prima. Sviluppato in collaborazione con Google, il nuovo sito web offre agli utenti l'accesso a immagini ricercabili, a caricamento rapido e ad alta risoluzione dei rotoli, oltre a brevi video esplicativi e informazioni di base sui testi e la loro storia. I Rotoli del Mar Morto, che comprendono i più antichi manoscritti biblici conosciuti, offrono una visione critica della società ebraica nella Terra d'Israele durante il periodo del Secondo Tempio, il tempo della nascita del cristianesimo e del giudaismo rabbinico. Cinque rotoli completi del Museo d'Israele sono stati digitalizzati per il progetto in questa fase e sono ora accessibili online […] I cinque rotoli del Mar Morto che sono stati digitalizzati finora includono il Grande Rotolo di Isaia, il Rotolo della Regola della Comunità, il Commento a Habakkuk, il Rotolo del Tempio e il Rotolo della Guerra, con query di ricerca su Google.com che inviano gli utenti direttamente ai rotoli online. Tutti e

cinque i rotoli possono essere ingranditi in modo che gli utenti possano esaminare i testi nei minimi dettagli. Dettagli invisibili a occhio nudo sono resi visibili attraverso la fotografia digitale ad altissima risoluzione del fotografo Ardon Bar-Hama - a 1.200 mega pixel ciascuno, queste immagini hanno una risoluzione quasi duecento volte superiore a quelle prodotte da una macchina fotografica standard. Ogni immagine ha utilizzato tubi flash protetti dai raggi UV con un'esposizione di 1/4000 di secondo per minimizzare i danni ai fragili manoscritti. Inoltre, il Grande Rotolo di Isaia può essere ricercato per colonna, capitolo e verso, ed è accompagnato da uno strumento di traduzione inglese e da un'opzione per gli utenti di presentare traduzioni di versi nelle loro lingue[294]

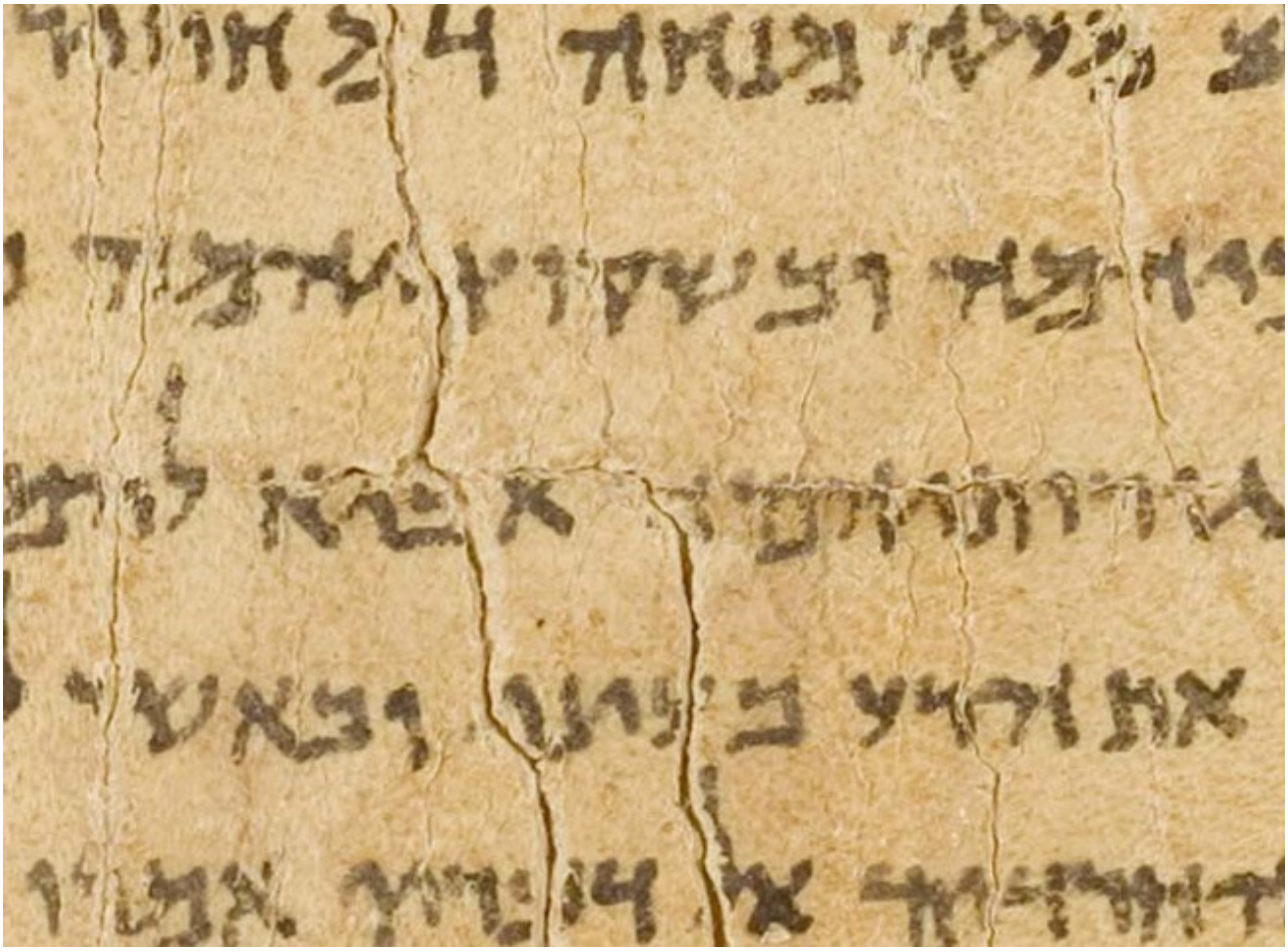
Il Grande Rotolo di Isaia (1QIsaa) è uno dei sette rotoli originali del Mar Morto scoperti a Qumran nel 1947. È il più grande (734 cm) e meglio conservato di tutti i rotoli biblici, e l'unico quasi completo. Le 54 colonne contengono tutti i 66 capitoli della versione ebraica del libro biblico di Isaia. Risalente al 125 a.C. circa, è anche uno dei più antichi rotoli del Mar Morto, circa mille anni più vecchio dei più antichi manoscritti della Bibbia ebraica conosciuti prima della scoperta dei rotoli[295].



Questo è lo screenshot di parte della pagina relativa al Grande Rotolo di Isaia (1QIsaa), oltre ad un’anteprima della pergamena è possibile vedere sulla colonna destra un breve video del Dr. Adolfo D. Roitman, che “ci porta” nel luogo dove il manoscritto originale è conservato e mentre ce lo mostra ricorda brevemente la storia di questo documento. Sotto al video troviamo i link per accedere alle riproduzioni digitali delle altre pergamene, ma ciò che più mi interessa portare

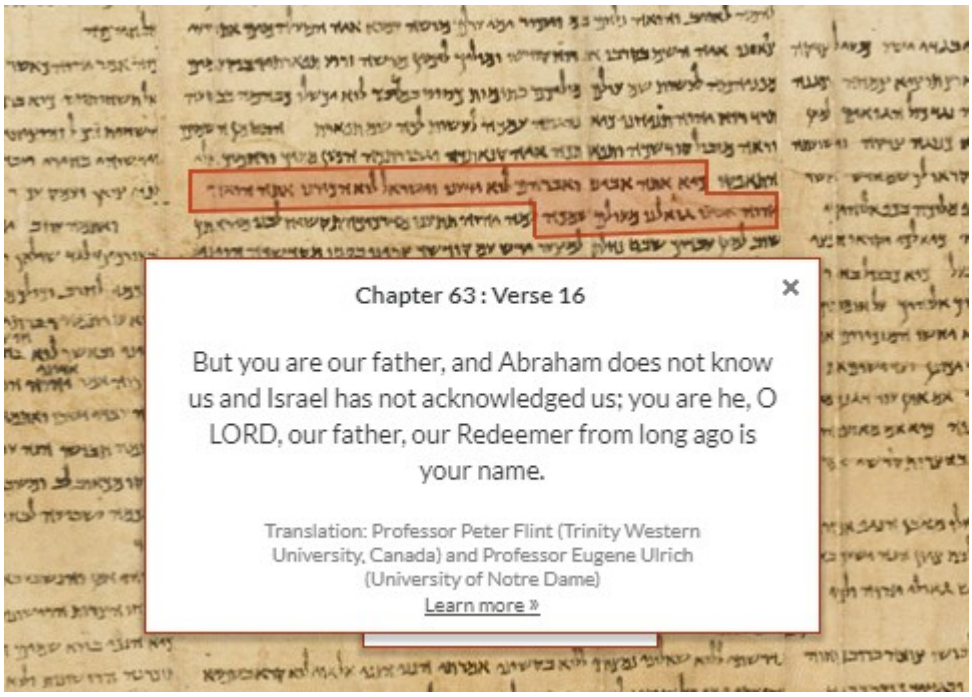
294 http://dss.collections.imj.org.il/project ultima consultazione 25/08/2021, Tradotto con www.DeepL.com/Translator (versione gratuita)
295 http://dss.collections.imj.org.il/isaiah#1:27 ultima consultazione 25/08/2021

l’attenzione è quella barra numerata che si trova nella parte bassa dell’anteprima del documento. Scorrendo la barra infatti è possibile scorrere il documento stesso, così da poterlo vedere nella sua interezza o poter andare nelle colonne che ci interessa osservare.
Cliccando sull’anteprima il potere della digitalizzazione si manifesta in modo ancora più chiaro, ecco cosa è possibile vedere zoomando sull’immagine digitale



La definizione dell'immagine è sbalorditiva, l’inchiostro è perfettamente leggibile e l’inquadratura che ho scelto consente di vedere le pieghe che il tempo ha lasciato sulla pelle, consegnandoci un punto di vista sul documento che in nessun modo avremmo potuto avere utilizzando solo la nostra vista.
Questo è un esempio che spiega molto bene ciò che definisco come medium-documento, ciò che abbiamo difronte non è solo il contenuto della pergamena, ma l’oggetto capace di trasmettere informazioni che non si manifestano solo nel contenuto ma nell’*essere* del documento.

Anche in questo caso però il processo di digitalizzazione non si limita ad un riproduzione del documento materico, gli sviluppatori infatti hanno reso possibile la lettura della pergamena anche per chi non sa leggere l'aramaico.



Cliccando su uno dei versetti il programma seleziona in rosso la parte che abbiamo segnalato e ci offre una traduzione in inglese, essendo un testo complesso che oltre alla lingua per essere compreso va letto attraverso le interpretazioni, selezionando "learn more" abbiamo un'analisi comparata del testo:

«Mentre utilizzate lo strumento traduttore nel visualizzatore di rotoli, vorremmo richiamare la vostra attenzione sulle complessità della traduzione delle parole del profeta Isaia di circa 2.800 anni fa, come si riflette nelle diverse varianti ebraiche e nelle successive traduzioni inglesi. La missione del museo qui è di fornirvi le informazioni di base necessarie per raggiungere la vostra prospettiva obiettiva quando leggete questa traduzione inglese del testo biblico»[296].

I due esempi che ho riportato sono progetti che si sono concentrati sulla digitalizzazione di un unico corpo documentario, manifestando dunque un interesse specifico verso l'oggetto materico, questi documenti digitali non sono in una relazione con altri documenti, non esistono infatti dentro ad archivi digitali, ma in portali costruiti ad hoc per l'interazione da parte degli utenti con quello specifico documento. La loro funzione è dunque molto vicina allo spirito della Public History, non

296 http://dss.collections.imj.org.il/version ultima consultazione 27/08/2021

solo perché il prodotto digitale è pensato in modo specifico per la fruizione di non addetti ai lavori, ma perché in entrambi i progetti è prevista un’interazione tra i portali e gli utenti, così da creare una maggior collaborazione. In un’ottica di Digital Public History, credo si possa riflettere sulla possibilità resa dal digitale, di portare le fonti fuori dagli archivi. Se da una parte è vero che la digitalizzazione allontana lo storico dal prodotto materico, quindi dalla fonte originale, dall'altra parte, paradossalmente, avvicina alla fonte originale chi storico non è.

3 ISAD (G) e ISAAR-CPF. Standard per una prospettiva internazionale degli archivi

Come abbiamo potuto vedere la digitalizzazione dei documenti e la creazione di documenti digitali ha cambiato in modo sensibile la possibilità di fruire le informazioni. La rivoluzione digitale consente innovative possibilità di produzione, accesso, fruizione e condivisione di quello che abbiamo chiamato medium-documento. Questo mutamento che ha implicazioni anche di natura epistemologica, implicazioni che trovano delle convergenze con il pensiero di autori tra i quali Topolski, White o Dray, ci richiede lo sforzo di essere analizzato anche da un punto di vista tecnico.
Il caos informativo che è proprio del Web configge in modo sensibile con le basi metodologiche dell'archivistica, come abbiamo visto ogni documento oltre a presentarsi come una complessità narrativa non esiste mai in senso isolato è infatti sempre inserito all'interno di un flusso informazionale composto da una serie di documenti attinenti al soggetto produttore o all'evento.
Perché le informazioni non si perdano nel Web è necessario immaginare e strutturare dei processi atti a garantire la reperibilità dei dati e delle informazioni. Dal momento che gli storici e le storiche del futuro lavoreranno quasi unicamente su fonti nate digitali ci dobbiamo chiedere come sarà possibile assicurare un contesto di archiviazione che possa garantire la validità delle fonti.
Già da alcuni anni i processi di archiviazione digitale si basano su strumenti che, da un punto di vita concettuale, derivano dai sistemi di archiviazione analogica. Certo le loro funzionalità sono enormemente potenziate dal digitale, vedremo infatti come i metadati e gli standard consentono oltre alla già citata interoperabilità delle fonti, anche una contestualizzazione sia dell'informazione sia dell'ente produttore davvero notevole.
Ma la vera rivoluzione è appena iniziata, ed è forse questa che davvero fa emergere alcune perplessità in parte degli archivisti. Il Web Semantico infatti prevede una decostruzione dell'informazione fino alla sua particella elementare, il dato, sul quale però si andrà a costruire nuova informazione. Questo porterà ad una atomizzazione dell'informazione che renderà i processi informativi digitali davvero liquidi e svincolati da contesti di archiviazione organizzati, non di meno però l'idea di dato autoconsistente ci mostra la possibilità, grazie ai metadati, di incapsulare sul dato tutte le informazioni necessarie ad una sua contestualizzazione anche di natura archivistica. Di tutto ciò ci occuperemo in modo più approfondito nel prossimo capitolo.
Prima di poter comprendere cosa sia e come funzioni il Web Semantico, dobbiamo necessariamente comprendere cosa siano e come funzioni gli standard e i metadati, questo perché, senza di essi il Web Semantico, per quanto rivoluzionario, non si sarebbe potuto mostrare.

La necessità di sviluppare standard condivisi di descrizione dei documenti non è figlia della digitalizzazione, anche se è indubbio che la rivoluzione digitale abbia imposto un'accelerazione a questo processo, come scrive Zanni Rosiello «Si trattava del resto di una tematica molto viva a livello internazionale e strettamente connessa a possibili progetti di informatizzazione e di costruzione di sistemi automatizzati, di cui anche nel nostro paese si avvertiva sempre di più la necessità»[297]

Anche se con un occhio volto al futuro già in ambito materico la comunità internazionale degli archivisti, che trova una sua strutturazione formale nell'ICA, l'Internationa Council on Archive, aveva iniziato ad elaborare degli standard internazionali per omologare i processi di descrizione così da consentire una fruizione dei documenti di archivio e delle informazioni a loro connesse, metodologicamente sempre più strutturata e standardizzata a livello internazionale. Come si legge sul sito

ICA believes that effective records and archives management is an essential precondition for good governance, the rule of law, administrative transparency, the preservation of mankind's collective memory, and access to information by citizens.

The International Council on Archives (ICA) is dedicated to the effective management of records and the preservation, care and use of the world's archival heritage through its representation of records and archive professionals across the globe[298]

Per poter comprendere lo sviluppo di standard internazionali per la descrizione e metadatazione dei documenti d'archivio in un ambiente digitale è necessario ripercorrere le tracce dei processi che hanno portato alla costituzione di standard internazionali pensati per la catalogazione e descrizione di archivi anche in ambito materico.

Ciò che maggiormente deve cogliere il nostro interesse è la riflessione riguardo la strutturazione di standard di descrizione archivistica, sia dell'inventario che di soggetti produttori, che avessero un valore internazionale, così da consentire un'omologazione necessaria per coadiuvare il lavoro di ricercatori e ricercatrici.

L'omologazione di standard è alla base dei processi di confederazione di archivi e di interoperabilità tra le fonti, processi questi, che oggi si mostrano tra gli elementi maggiormente innovativi generati dalle pratiche di digitalizzazione. Non solo, è grazie a sistemi come EAD o EAC-CPF che è

297 I. Zanni Rosiello, *Gli archivi nella società contemporanea*, Il Mulino, Bologna 2009, p. 156
298 https://www.ica.org/en ultima consultazione 28/10/2021

possibile consultare online indici archivistici, operazione che come abbiamo già detto risulta essere basilare per procedere nella ricerca attraverso il Web.

Pensiamo a sistemi informativi come il SAN[299] il SIAS o il SIUSA che rispetto alle guide tradizionali danno la possibilità di generare sistemi di relazioni che collegano reciprocamente e dinamicamente le entità informative, riuscendo a restituire con estrema immediatezza il complesso panorama del percorso di produzione, uso e conservazione e ancorando i complessi documentari a tale contesto, che è al tempo stesso chiave primaria ai fini della reperibilità e garanzia di corretta utilizzazione di dati attendibili[300].

Prima di poter comprendere cosa siano EAD Encode Archival Description e EAC-CPF Encoded Archival Context for Corporate Bodies, Persons, and Families, standard in XML rispettivamente per la descrizione dell'inventario d'archivio e per la descrizione dei soggetti produttori, standard su cui si basano i successivi processi di metadatazione degli oggetti digitali, dobbiamo comprendere cosa sono gli standard ISAD(G) e ISAAR(CPF).

Nell'ottobre del 1988, su iniziativa di Harold Naugler, funzionario del National Archives of Canada e, in seguito, Segretario dell'ICA Automation Committee, il National Archives of Canada e il Consiglio internazionale degli archivi, organizzarono congiuntamente, ad Ottawa, un incontro di esperti nella descrizione archivistica, in vista del XII Congresso internazionale degli Archivi che doveva tenersi a Montréal nel settembre 1992.

All'incontro presero parte archivisti inviati da quindici paesi diversi: D.R. Castonguay, V. Chabot, J.E. Dryden, J. Grimard, K.M. Haworth, H. MacNeil, H. Naugler, H.L.P. Stibbe J. Thacker D.R. Webwe, (Canada), Deng Song Dian (Cina), G. Brunel (Francia), M. Cook, M. Procter, D.L. Thomas (Gran Bretagna), Gurmeet Singh (India), P. Carucci (Italia), H.Z. Yahaya (Malesia), V. San Vicente (Messico), P. Horsman (Olanda), W. Werner (Repubblica federale tedesca), Marane Fall (Senegal), J. Dahlin (Svezia), D. Bearman, S.G. Thibodeau, B.R. Tucker (USA), S. Njovana (Zimbawe)[301].

Dal confronto tra le diverse realtà nazionali, emerse la necessità di una attenta riflessione sul complesso rapporto tra standard descrittivi e automazione, e sullo scambio di informazioni a livello nazionale ed internazionale, tra istituti e banche dati diversi.

299 Il Sistema Archivistico Nazionale SAN ha sviluppato dei suoi standard per l'interoperabilità delle fonti METS-SAN e CAT-SAN, ma per i complessi archivistici e i soggetti produttori si appoggia su EAD e EAC-CPF

300 F. Valacchi, *Software di descrizione e sistemi informativi archivistici: aspetti teorici*, https://docenti.unimc.it/federico.valacchi/teaching/2019/19962/files/archivi-e-ict, ultima consultazione 26/10/2021

301 Gli interventi presentati sono raccolti nel volume *Toward International Descriptive Standards for Archives. Papers presented at the ICA Invitational Meeting of Experts on Descriptive Standards*. National Archives of Canada, Ottawa, 4-7 October 1988, K.G. Saur, Munchen, New Providence, London, Paris, 1993

Risultato del lavoro del Working Group era stata la pubblicazione, nel 1985, del rapporto intitolato *Toward Descriptive Standards*[302] che, analizzando la poca sistematicità e coerenza delle descrizioni archivistiche correnti, aveva proposto una metodologia per lo sviluppo e l'applicazione di standard di descrizione, per tutti i livelli di ordinamento e tutti i supporti, ed indicato tre obiettivi prioritari per il lavoro del neo-costituito Planning Committee on Descriptive Standards.

Partendo da queste premesse, negli anni successivi furono elaborate le *Rules for Archival Description*[303] che cercarono un punto di unione tra la tradizione archivistica europea e quella americana.

Difronte alla generale incertezza sulle caratteristiche, i contenuti, le finalità e l'utilizzo degli standard di descrizione, gli esperti riuniti ad Ottawa analizzarono le diverse pratiche descrittive nazionali esistenti, misero a confronto la specificità della descrizione archivistica con le forme di automazione già in atto per la catalogazione bibliografica e formularono alcuni punti di accordo sul valore delle norme per la creazione di sistemi informativi unificati, comprendenti biblioteche, archivi, musei e centri di documentazione in genere. Le risoluzioni finali sottolinearono come l'utilizzo degli standard dovesse favorire l'accesso ai documenti, la gestione e lo scambio di informazioni tra e negli istituti archivistici.

Sulla base delle conclusioni dell'incontro di Ottawa, nel dicembre del 1989 l'ICA organizzò a Parigi con il sostegno dell'UNESCO una riunione consultiva ristretta, con la partecipazione di alcuni degli esperti di Ottawa, al fine di elaborare un piano di azione a livello internazionale.

Durante questo incontro, il gruppo di lavoro individuò tre obiettivi prioritari[304]: la redazione di una dichiarazione contenente i principi generali per la descrizione archivistica, l'elaborazione di regole generali per la descrizione dei documenti di archivio a livello di fondo, basate su quei principi, e la stesura di regole specifiche per documenti su supporti speciali e particolari livelli di ordinamento. Stabilirono, inoltre, che la stesura delle regole dovesse essere basata sulle indicazioni condivise dai principali manuali di descrizione archivistica nazionali esistenti quali quello statunitense *Archives, Personal Papers, and Manuscripts: a Cataloguing Manual for Archival Repositories, Historical Societies, and Manuscripts Libraries* (APPM)[305], il britannico *A Manual of Archival Description*

302 *Toward Descriptive Standards. Report and Recommendations of the Canadian Working Group on Archival Descriptive Standards*, Ottawa, Bureau of Canadian Archivists, 1985.

303 *Planning committee on descriptive standards, Comité de planification sur les normes de description, Rules for Archival Description, Regles pour la description des documents d'archives*, Ottawa (Canada), Bureau of Canadian Archivists, 1990.

304 *Consultation restreinte de l'Unesco sur l'etablissement d'un plan d'action international à long terme pour l'élaboration de normes de description pour les archives*, Paris (UNESCO), France, 1-2 décembre 1989, in «Bulletin du Conseil international des Archives», 34, juin 1990, p. 13

305 S.L. Hensen, *Archives, Personal Papers, and Manuscripts: a Cataloguing Manual for Archival Repositories, Historical Societies, and Manuscripts Libraries*, Chicago, Society of American Archivist, 1989

(MAD)[306] e le già citate *Rules for Archival Description/Regles pour la description des documents d'archives* (RAD/RDDA) canadesi.

Lo *Statement of Principles Regarding Archival Description* si propose come un primo indispensabile passo nella formulazione di una teoria generale sulla descrizione archivistica accettata a livello internazionale, ancorandosi alle pratiche e alle teorie archivistiche tradizionali. Fin dalla premessa vengono enunciati alcuni principi fondamentali condivisi a livello internazionale quali la centralità del contesto e della storia istituzionale in cui le carte sono state prodotte e conservate, e il riconoscimento del ruolo della descrizione archivistica come strumento di controllo intellettuale e di accesso ai documenti. Viene anche ribadita la necessità di una terminologia condivisa, valida a livello internazionale.

Lo *Statement of Principles Regarding Archival Description* enucleò alcuni punti fermi su cui, in seguito, si sarebbe basata ISAD(G): la validità del concetto di fondo come unità di descrizione, la corrispondenza tra i livelli di ordinamento e i livelli di descrizione in un modello gerarchico, e la descrizione plurilivellare dal generale al particolare, che consente di produrre descrizioni pertinenti al livello considerato, ma tra loro correlate gerarchicamente.

In entrambe le stesure è presente un glossario della terminologia archivistica utilizzata quali i termini: descrizione archivistica, fondo, punto di accesso, serie, file e item. Nell'ultima stesura è stato introdotto anche il concetto di *authority control* che porterà alla formulazione dell'International Standard Archival Authority Record for Corporate Bodies, Persons and Families ISAAR(CPF)[307].

La stesura di ISAD(G) fu soltanto una tappa di un vasto progetto di normalizzazione; il piano di lavoro della Commissione *ad hoc*, infatti, prevedeva anche l'elaborazione di norme per la redazione di liste d'autorità dei soggetti produttori d'archivio, di criteri per la scelta dei punti di accesso, di norme per gli archivi su supporti speciali (fotografici, cartografici, audiovisivi, informatici) e di indicazioni sulla scelta dei formati di scambio dei dati su supporto informatico. Fin dalla prima stesura dello *Statement of Principles Regarding Archival Description*, risultò evidente la necessità di puntualizzare e definire norme precise per regolare l'individuazione e la descrizione dei soggetti produttori e della loro storia istituzionale, amministrativa o della loro biografia, come potenziali punti di accesso per un'identificazione univoca e un reperimento efficace delle descrizioni archivistiche relative ai fondi da essi prodotti.

306 M. Cook, K. Grant, *A Manual of Archival Description*, Liverpool, Society of Archivists, 1985

307 F. Ricci, *Gli standard internazionali di descrizione archivistica: dalle origini alla seconda versione di Isad(G), in* «Rassegna degli Archivi di Stato», anno LXIII - n. 1, Roma, 2003, p. 22 https://www.icar.beniculturali.it/fileadmin/risorse/docu_standard/RAS_2003_1.pdf ultima consultazione 28/10/2021

Nel gennaio del 1993, a Stoccolma, durante la terza sessione plenaria della Commissione *ad hoc* sulle norme di descrizione (ICA/CDS), venne costituita una sotto-commissione, composta da W. Duff, J. Dahlin, C. Nougaret, M. Cook e H.L.P. Stibbe, col compito di elaborare norme internazionali per il controllo del contenuto e della forma dei punti di accesso ai sistemi informativi archivistici.

La bozza di questo nuovo standard, ideato e discusso dalla sotto-commissione nel novembre del 1993 a Liverpool, venne presentata per la prima volta durante la seduta plenaria della Commissione a l'Aja nell'ottobre del 1994, come International Standard for Archival Description (Authority Record)[308]. In quella occasione la prima versione del documento fu ampliata, modificata e proposta alla discussione della comunità archivistica internazionale, sotto il titolo di International Standard Archival Authority Record for Corporate Bodies, Persons and Families, ISAAR(CPF). Ai commenti ricevuti fu dedicata la sessione plenaria della Commissione *ad hoc* sugli standard descrittivi tenutasi a Parigi nel novembre 1995, durante la quale fu approvata la versione finale di ISAAR(CPF)[309].

Il fatto che si sia scelto di creare un modello di descrizione *ad hoc* per i soggetti produttori e dunque distinto dal modello di descrizione del documento, rende il lavoro svolto per la realizzazione dello standard ISAAR(CPF) particolarmente lungimirante in relazione ai processi di produzione informazionale che possiamo osservare all'interno dell'infosfera, dove inforganismi producono grandi quantità di documenti e dove dunque risulta sempre più importante poter approfondire e organizzare al meglio il maggior numero di informazioni relative ai soggetti produttori.

Come scrisse Cook già nel 1996, senza poter prevedere l'avvento del Web 2.0 « The appearance of ISAAR(CPF) is an indication that the future world in which significant numbers of users will get their initial contact information from the Internet, is seen as a place where standard indexes of creators will be the main access point»[310]

Questo standard fornisce una guida per l'elaborazione di record di autorità archivistici che offrano descrizioni di entità quali enti, persone e famiglie coinvolte nella produzione e conservazione degli archivi.

I record di autorità archivistici possono essere utilizzati per descrivere questi enti, come elementi di un sistema di descrizione archivistico; per sottoporre a controllo d'autorità l'elaborazione e l'uso di

308 H.L.P. Stibbe, *Commission Ad Hoc sur les norms de description (ICA/DDS)*, in « Bulletin du Conseil international des Archives », 46, 1996, pp. 38-39

309 International Council On Archives, ISAAR(CPF): International Standard Archival Authority Record for Corporate Bodies, Persons and Families, Prepared by the Ad Hoc Commission on Descriptive Standards, Paris, France, 15-20 November 1995, Ottawa, 1996. Il testo originale inglese, affiancato dalla traduzione francese, è stato pubblicato in Italia in RAS, LVIII (1998), 2-3, pp. 443-477 e la traduzione italiana, curata da StefanoVitali, in RAS, LIX (1999), 1- 2-3, pp. 223-252

310 M. Cook, *The International Description Standards: new departures*, in « Archivi e Computer », VI (1996), 3-4, p. 263

chiavi d'accesso alle descrizioni archivistiche; per documentare le relazioni fra differenti soggetti produttori e fra questi e la documentazione da essi prodotta e altre risorse ad essi relative o ascrivibili.

La descrizione dei soggetti produttori d'archivio è una attività fondamentale degli archivisti, a prescindere dal fatto che tali descrizioni siano gestite mediante sistemi manuali o automatizzati. Ciò richiede di documentare esaurientemente e di aggiornare continuamente le informazioni sul contesto di produzione e di uso degli archivi, soprattutto in riferimento alla provenienza di archivi e documenti[311].

L'idea di un modello di descrizione separata e correlata trae effettivamente la sua origine dalla constatazione di come la descrizione archivistica coinvolgesse non una sola, ma due entità in relazione tra loro: archivi e soggetti produttori.

Una proposta in tal senso era stata avanzata nel 1986 da M.J. Evans[312], che suggeriva di sostituire al tradizionale sistema descrittivo, basato su un modello gerarchico, con al livello più alto l'ente creatore dell'archivio per scendere via via fino alle serie e alle singole unità, un sistema fondato sulla descrizione separata delle serie archivistiche, da una parte *record series*, relative ai documenti e dall'altra *authority file*, relativi agli enti produttori[313].

Ci sono numerose ragioni per le quali la gestione separata di questo tipo di informazioni sul contesto è una componente essenziale della descrizione archivistica. Una prassi di questo genere permette di collegare le informazioni alla descrizione della documentazione archivistica dello stesso soggetto produttore che sia conservata in più istituzioni archivistiche, o alla descrizione di altre risorse quali materiale bibliografico e museale, riconducibili allo stesso soggetto produttore. Tali connessioni migliorano le pratiche di gestione degli archivi e facilitano la ricerca.

Questo breve *excursus* storico sula formazione di ISAD(G) e ISAAR(CPF), standard che consentono la costruzione di modelli descrittivi dei fondi d'archivio e degli enti produttori, ci mostra lo sforzo delle società di archivistica nazionali e internazionali al fine di garantire un primo accesso alla conoscenza dei documenti contenuti all'interno degli archivi, consentendo così alla comunità internazionale di ricercatori e ricercatrici di poter avere una panoramica standardizzata di quali documenti si trovino in quali istituzioni e ottenere alcune informazioni base riguardo al

311 H.L.P. Stibbe, K. Hawoth, *ISAAR (CPF) Standard internazionale per i record d'autorità archivistici di enti, persone, famiglie*, International Council on Archive, II edizione, 2004, https://www.ica.org/sites/default/files/CBPS_Guidelines_ISAAR_Second-edition_IT.pdf ultima consultazione 24/10/2021

312 M.J. Evans, *Authority Control: an Alternative to the Record Group Concept*, in « American Archivist », XLIX (1986), 3, pp. 249-261,

313 M. Savoja, *Descrizioni di archivi e soggetti produttori: ISAD(G) e ISAAR(CPF), in Gli standard internazionali per la descrizione archivistica: le regole ISAD(G).* Incontro seminariale sulle proposte di revisione elaborate dal gruppo di lavoro ANAI. Bologna, 11 febbraio 1998, Bologna, Archivio di Stato di Bologna - ANAI Sezione Emilia-Romagna, 1998, p 19

contesto dei documenti conservati e agli enti produttori, in modo da poter orientare al meglio i propri percorsi di ricerca e poter così implementare le proprie domande di ricerca.
Se la nostra riflessione fosse volta all'analisi dei processi di ricerca in un ambiente materico dove i documenti sono organizzati sia logicamente sia fisicamente attraverso il metodo storico di ordinamento, e dove le prime interlocuzioni con l'archivio avvengo anche grazie all'intermediazione di un archivista che aiuta ricercatrici e ricercatori ad orientarsi nelle complessità archivistiche, si potrebbe pensare che per storici e storiche non sia necessario conoscere gli schemi e i campi sviluppati e utilizzati per ISAD(G) e ISAAR(CPF), ma dal momento che in queste pagine ci stiamo occupando dei processi di digitalizzazione delle fonti e degli archivi, potrebbe mostrarsi come un limite non interrogarsi su quali campi e di conseguenza quali parole chiave è possibile utilizzare nell'interrogazione di un sistema informativo. Per semplificare: se accedendo al portale SAN siamo consapevoli di quali schemi abbiamo a disposizione per interrogare il portale, la nostra ricerca potrà essere meno serendipitosa e potrebbe fin da subito svilupparsi su processi di interrogazione ben più concreti.
I sistemi informativi archivistici (SIA), come è il SAN, rappresentano una delle ricadute più significative della riflessione sulle applicazioni tecnologiche agli archivi storici. Genericamente sono percepiti come un insieme non meglio specificato di risorse intorno agli archivi, ma nella realtà i SIA devono essere considerati strumenti di orientamento per la ricerca, e un primo "luogo" di incontro tra i contenuti della storia e fruitori non specializzati. Pensiamo quali potrebbero essere le implicazioni se compiendo una ricerca sul Web riguardo un evento o un personaggio legati al passato, partendo da un motore generalista ricevessimo come primi risultati non solo la pagina Wikipedia dedicata, ma la possibilità di accedere ad ad un SIA, l'utente potrebbe usufruire di informazioni attendibili con la possibilità di accedere alle fonti primarie.
Al sistema informativo spetta il compito di svolgere un ruolo che in ambiente cartaceo compete sostanzialmente alle guide, con la differenza sostanziale della maggiore potenza di combinazione e veicolazione delle informazioni.

La maggiore efficacia del SIA rispetto alla guida tradizionale si manifesta in particolare nella possibilità di generare sistemi di relazioni che collegano reciprocamente e dinamicamente le entità informative, riuscendo a restituire con estrema immediatezza il complesso panorama del percorso di produzione, uso e conservazione e ancorando i complessi documentari a tale contesto, che è al tempo stesso chiave primaria ai fini della reperibilità e garanzia di corretta utilizzazione di dati attendibili.[314]

314 F. Valacchi, *Software di descrizione e sistemi informativi archivistici: aspetti teorici*, 2019 https://docenti.unimc.it/federico.valacchi/teaching/2019/19962/files/archivi-e-ict ultima consultazione 28/10/2021

L’utilizzo degli standard, dunque, oltre a garantire un’omologazione nei processi descrittivi, omologazione che va a rafforzare la scientificità del metodo di ricerca, costruisce intorno alla fonte primaria un sistema di informazioni molto articolato; queste informazioni vanno dalle descrizioni orizzontali, relative cioè ai cosiddetti “livelli alti” della descrizione, alle descrizioni verticali, cioè descrizioni che possono entrare nel merito e nel dettaglio dei contenuti dei complessi archivistici e delle unità che li costituiscono.

È chiaro che per ricercatrici e ricercatori le informazioni di contesto presenti negli standard descrittivi possano risultare abbastanza superficiali, anche se non di meno offrono indicazioni bibliografiche e archivistiche collegate alle fonti, ma se noi immaginiamo uno studente delle scuole secondarie di secondo grado che per motivi didattici o anche solo per interesse, dopo una ricerca su motori di ricerca generalisti, “atterra” su un portale di un sistema informativo archivistico, sappiamo che con delle semplicissime ricerche potrà confrontarsi con una grande mole di informazioni già organizzate e della cui attendibilità si può essere sicuri, così da limitare i problemi della *misinformation* di cui abbiamo già parlato.

Ciò che è successo a Ottawa nel 1988 ha reso possibile iniziare a porre le basi, fosse anche solo da un punto di vista logico, di come organizzare l’informazione all'interno di sistemi complessi, già allora con la consapevolezza che la produzione informazionale sarebbe esponenzialmente aumentata, non solo, consapevoli anche che sempre più spesso si sarebbe manifestata la necessità di reperire fonti in modo ordinato da un target di fruitori meno specializzato di come si era abituati. È nella costituzione di ISAD e di ISAAR che dobbiamo cercare i prodromi per la costruzione di un sistema capace di dare accesso in modo ordinato alle informazione riguardo al passato non solo di una persona, di un ente o di uno stato, ma dell’umanità. Certo nel 1988 forse non si poteva immaginare ciò che sarebbe davvero successo con la diffusione di internet e del Web 2.0, così come sicuramente Licklider nel 1948 - come abbiamo visto nel primo paragrafo del primo capitolo - non poteva sapere quanto queste sue parole sarebbero potute essere profetiche:

> C’è un certo numero di esseri umani connessi da link di comunicazione in modo da formare una rete. Hanno accesso a determinate informazioni e un compito da svolgere, che richiede un intervento basato sull’informazione. La domanda è: in quale modo la loro performance nello svolgimento del compito dipende dalla configurazione della rete?[315]

La cosa incredibile è che se anche ponessimo che nessuno degli archivisti presenti ad Ottawa fosse a conoscenza della domanda posta da Licklider è anche grazie al loro lavoro che si è provato a

315 R.M. Fano, *Joseph Carl Robnett Licklider,* in National Research Council Biographical memories V.75, https://www.nap.edu/read/9649/chapter/12#191 ultima consultazione 15/1/2022

proporre una soluzione. Oggi sappiamo che la configurazione della rete è centrale per assicurare l'accesso a informazioni affidabili.

Nei sistemi informativi siamo in grado di costruire dei sistemi ai quali si può accedere "esplodendo" i diversi nodi a partire da uno dei punti di accesso e visualizzando quindi, sulla base delle relazioni generate nel modulo di produzione, le relazioni e i contenuti descrittivi delle diverse entità informative, ovvero interrogando il sistema a partire da una maschera di ricerca che "punta" direttamente alle informazioni desiderate. A prescindere dalle strategie di ricerca, proprio grazie alle relazioni correttamente stabilite tra le entità, da qualsiasi punto di accesso è possibile ricostruire nella sua interezza il panorama informativo relativo al sistema di fonti oggetto della ricerca.

L'accessibilità data dalla dematerializzazione è da considerarsi incredibilmente rilevante per la diffusione sociale dei contenuti storiografici e della storia. Grazie a questi standard di archiviazione e a strumenti come EAD e EAC-CPF gli archivi e i sistemi informativi ad essi connessi possono diventare dei potenti mezzi di comunicazione scientifica.

Ad avvalorare questo ragionamento, come ci ricorda Valacchi è la consapevolezza che EAD e EAC-CPF, gli standard di descrizione in XML che consentono la navigazione e l'interoperabilità tra descrizioni di diversi archivi, sono formati per la restituzione e non per la descrizione[316], in altre parole non sono altro che strumenti atti a consentire la comunicazione della descrizione organizzata con gli standard ISAD(G) e ISAAR(CPF).

Come ci aiuta a riflettere Gino Frezza[317], i processi di digitalizzazione delle fonti e la conseguente nascita degli archivi digitali, rendono gli archivi sempre più strumenti atti non solo alla conservazione, ma anche alla comunicazione degli enunciati che hanno al proprio interno, ma se è vero come è vero che la comunicazione esiste solo se ci sono un mittente, un messaggio e un ricevente, dobbiamo assicurarci che mittente e ricevente siano in grado di comunicare tra di loro attraverso un linguaggio condiviso. Come abbiamo visto i codici ASCII rendono possibile l'interlocuzione tra un linguaggio macchina e un linguaggio umano, ma se interrogando un archivio non si conoscono i campi di indagine di cui è composto, avere delle riposte corrette sarà molto complicato. Se invece si comprendono gli schemi che ci consentono di interrogare i multilivelli descrittivi di un archivio, sarà più semplice avviare un processo di comunicazione funzionale.

316 F. Valacchi, *L'applicazione degli standard agli strumenti archivistici nel Web*, Università i Macerata, 2020 https://docenti.unimc.it/federico.valacchi/teaching/2020/22027/files/standard-di-descrizione-archivistica ultima consultazione 28/10/2021

317 G. Frezza, a cura di, AA.VV., *L'arca futura. Archivi mediali digitai, audiovisivi, web*, Meltemi Editore, Roma 2008

3.1 ISAD(G) il padre degli standard

Partendo da ISAD(G) che come abbiamo visto è lo standard internazionale primigenio, la prima cosa da sapere è che il suo funzionamento si basa prima di tutto su descrizioni multilivellari che devono essere in correlazione tra di loro in un ordine gerarchico, dunque dal generale al particolare.

A livello di fondo fornire le informazioni relative al fondo nel suo complesso. Al livello seguente e ai successivi dare le informazioni relative a ciascuna delle parti che viene descritta. Disporre le descrizioni che ne risultano secondo uno schema di relazioni gerarchiche che metta in rapporto la singola parte con l'insieme e che proceda dal generale (il fondo) al particolare[318]

A partire dalla comprensione di questo passaggio, fondamentale per il funzionamento generale di ISAD(G), di seguito riporto in modo schematizzato i campi di cui è composto lo standard, così che si possa avere compiutezza dei diversi livelli di descrizione[319]. I livelli di descrizione non codificano solo l'ossatura dello standard ma costituiscono anche la struttura comunicativa dei fondi e delle risorse è dunque fondamentale comprenderne la struttura se si immagina che nel futuro digitale nel quale siamo già inseriti, potranno essere proprio i sistemi informativi ad essere primo veicolo per la diffusione della storia ad un pubblico generalista. Gli archivi non più solo come strumenti di conservazione e primi ambiti di ricerca per esperti ricercatori, ma strumenti di comunicazione.
È all'interno di ognuno di questi livelli che ritroviamo gli elementi descrittivi che poi, se pur in un orizzonte di atomizzazione, andranno a definire l'autoconsistenza del dato, che permette al Web Semantico di mostrarsi come una tecnologia e un riflessione adottabile anche per la conservazione delle informazioni in una prospettiva di lunga durata.
Ma procediamo con ordine, e con pazienza iniziamo il nostro viaggio all'interno di quella che possiamo definire come la sintassi del discorso archivistico all'interno del Web, ricordando le parole di Annalisa Rossi «I racconti si costituiscono non delle storie attingibili dagli archivi, ma delle trame narrative che le relazioni di cui un archivio si compone intessono: l'ordito che ne risulta è un tessuto possibile»[320]

AREA DELL'IDENTIFICAZIONE

318 F. Ricci, *Gli standard internazinali di descrizione archivistica: dalle origini alla seconda versione di Isad(G), in* «Rassegna degli Archivi di Stato», anno LXIII - n. 1, Roma, 2003, p. 85 https://www.icar.beniculturali.it/fileadmin/risorse/docu_standard/RAS_2003_1.pdf ultima consultazione 28/10/2021

319

320 A. Rossi, *La palla di vetro. Immaginare la tutela e la comunicazione archivistica diffusa*, in AA.VV., a cura di, L. Pezica, F. Valacchi, *Dimensioni archivistiche. Una piramide rovesciata,* Editrice bibliografica, Milano 2021, p.62

- Denominazione o titolo: Serve a denominare l’unità di descrizione. Richiede di riportare il titolo originale del documento oppure di attribuire un titolo conciso in conformità alle regole della descrizione in più livelli e delle convenzioni nazionali.
- Data: Serve a identificare e segnalare le date dell’unità di descrizione. È necessario indicare almeno uno dei seguenti tipi di data relativa all’unità di descrizione: 1. Le date nelle quali i documenti archivistici sono stati accumulati dal soggetto produttore nell’esercizio dei propri affari o nella condotta delle proprie attività, in riferimento dunque all'archivio di deposito; 2. Le date nelle quali i documenti sono stati formati, quando quindi l’archivio era nella fase corrente.
- Livello di descrizione: Serve a identificare il livello dell’unità di descrizione all’interno dell’ordinamento del fondo. È necessario indicare il livello dell’unità di descrizione in questione: Fonds (fondo) Sub-fonds (sub-fondo) Series (serie) Sub-series (sottoserie) File (unità archivistica) Item (unità documentaria)
- Consistenza e supporto dell’unità di descrizione (quantità, volume, dimensione fisica): Serve per identificare e descrivere: la consistenza materiale o logica e il supporto dell’unità di descrizione. Segnalare la consistenza dell’unità di descrizione, dando, in cifre, il numero totale delle unità materiali o logiche e l’unità di misura e a indicare i supporti dell’unità di descrizione

AREA DELLE INFORMAZIONI SUL CONTESTO

- Denominazione del soggetto produttore: Serve per identificare il soggetto (o i soggetti) produttori dell’unità di descrizione. Indicare il nome dell’organismo (o degli organismi) o delle persone responsabili della produzione, accumulazione e conservazione della documentazione dell’unita di descrizione. Il nome dovrebbe essere formulato in una forma standardizzata come prescritto dalle convenzioni nazionali in conformità dei principi di ISAAR (CPF)
- Storia istituzionale/amministrativa, nota biografica: Fornisce una storia istituzionale/amministrativa o un profilo biografico del soggetto produttore (o dei soggetti produttori), per collocare la documentazione nel proprio contesto e facilitarne la comprensione. Richiede di indicare in forma concisa tutte le informazioni significative sull’origine, l’evoluzione, le vicende e l’attività dell’organismo (o degli organismi) o sulla vita e l’attività della persona (o delle persone), responsabili della produzione dell’unità di descrizione e se ulteriori informazioni possono essere reperite in una fonte edita, e a citare questa fonte.

- Storia archivistica: Fornisce le informazioni sulla storia dell'unità di descrizione che sono significative ai fini dell'autenticità, integrità e interpretazione della documentazione. Serve per segnalare i passaggi di proprietà, responsabilità e custodia dell'unità di descrizione ed indicare quegli interventi che hanno contribuito alla sua struttura e ordinamento, quali la storia degli ordinamenti, l'elaborazione di strumenti di ricerca coevi, il riutilizzo della documentazione per finalità diverse o le migrazioni in ambienti software diversi. Serve a fornire le date di questi interventi, nella misura in cui possono essere accertate e a segnalare se la storia archivistica non è conosciuta.
- Modalità di acquisizione o versamento: Identifica il soggetto dal quale è stata acquisita o che ha versato la documentazione, indicando oltre al soggetto la data e il titolo di acquisizione, se tali informazioni non sono del tutto o in parte riservate e a segnalare se il soggetto non è conosciuto.

AREA DELLE INFORMAZIONI RELATIVE AL CONTENUTO E ALLA STRUTTURA

- Ambiti e contenuto: Permettono agli utenti di valutare la potenziale rilevanza dell'unità di descrizione illustrando sinteticamente gli ambiti (cronologici e geografici) e il contenuto (come la tipologia documentaria, le materie trattate, le procedure amministrative) dell'unità di descrizione pertinenti al livello di descrizione.
- Procedure, tempi e criteri di valutazione e scarto: Fornire le informazioni relative a tutte le operazioni di valutazione e di scarto e alle relative epoche. Indicare le operazioni di valutazione e di scarto, effettuate o pianificate sull'unità di descrizione nel corso del tempo, specialmente se esse sono rilevanti per l'interpretazione della documentazione.
- Incrementi previsti: Serve per informare l'utente su previsti accrescimenti dell'unità di descrizione e se opportuno, dare una stima della loro quantità e frequenza.

AREA DELLE INFORMAZIONI RELATIVE ALLE CONDIZIONI DI ACCESSO ED UTILIZZAZIONE

- Condizioni che regolano l'accesso: Fornisce le informazioni sulla condizione giuridica e su altre forme di regolamentazione che limitano o condizionano l'accesso all'unità di descrizione.
- Condizioni che regolano la riproduzione: Serve a segnalare tutte le limitazioni alla riproduzione dell'unità descrizione.
- Lingua/scrittura della documentazione: Segnala la/le lingua/e, la/le scrittura/e o i sistemi di simboli utilizzati nell'unità descrizione.

- Caratteristiche materiali e requisiti tecnici: Fornisce informazioni su tutte le caratteristiche materiali e i requisiti tecnici che sono rilevanti ai fini dell'utilizzazione dell'unità di descrizione
- Strumenti di ricerca: Segnala tutti gli strumenti di ricerca relativi all'unità di descrizione.

AREA DELLE INFORMAZIONI RELATIVE A DOCUMENTAZIONE COLLEGATA

- Esistenza e localizzazione degli originali: Indica l'esistenza, la localizzazione, la disponibilità e la distruzione degli originali nel caso in cui l'unità di descrizione consista di copie.
- Esistenza e localizzazione di copie: Indica l'esistenza, la localizzazione e la disponibilità di copie dell'unità di descrizione.
- Unità di descrizione collegate: Segnala l'esistenza di unità di descrizione collegate.
- Bibliografia: Identifica e indica tutte le pubblicazioni che si riferiscano all'unità di descrizione o siano basate sullo studio, l'analisi od una qualche altra forma di utilizzazione di essa.

AREA DELLE NOTE

- Fornisce informazioni particolari ed informazioni che non possono essere inserite in nessuna delle altre aree.

AREA DI CONTROLLO DELLA DESCRIZIONE

- Nota dell'archivista: Serve a spiegare come e da chi è stata elaborata la descrizione
- Norme o convenzioni: Identifica le regole e le convenzioni sulle quali è basata la descrizione.

Dopo aver letto le diverse aree di cui è composto lo standard è facile osservare come alcune di queste abbiano una pertinenza puramente tecnica e di interesse prevalentemente per gli archivisti, altre aree, a partire da quella di "informazioni sul contesto", forniscono utili informazioni anche per il lavoro del ricercatore e alcune di queste in modo specifico per il lavoro del public historian, come per esempio le informazioni relative al diritto di utilizzo o meno di specifiche fonti, informazione necessaria per immaginare una messa in scena delle fonti stesse o un loro riuso creativo.
Quello che abbiamo osservato è lo standard utilizzato per organizzare le informazioni dei fondi archivistici in relazione alle fonti stesse, questo vuol dire che ognuno di quei campi è e deve essere applicato ad ogni singola unità informazionale.

Le informazioni presenti nello schema ISAD descrivono la fonte e la sua storia, il contesto in cui è inserita, il lavoro di selezione degli archivisti, le correlazioni all'interno del processo di ordinamento, le relazioni con le altre fonti a partire da uno specifico punto di accesso.
Già con queste informazioni è possibile tracciare un filo che rappresenta i prodromi di un racconto, ancora privo della spiegazione, ma di per sé sufficiente a creare un sistema informazionale strutturato e capillare.
Pensiamo se, attraverso e grazie a questo reticolo di informazioni, l'utente arrivasse a confrontasi anche con la rappresentazione simulacrale della fonte primaria e di conseguenza con tutte le informazioni che abbiamo descritto quando abbiamo parlato del medium-documento. Anche se in proporzione ancora sono pochi i documenti materici digitalizzati rispetto alla mole documentale presente negli archivi di tutto il mondo, oggi è già possibile entrare in archivi digitali e avere la possibilità di confrontarsi con la rappresentazione simulacrale del documento e come vedremo, con l'utilizzo di metadati descrittivi, possiamo arrivare a teorizzare anche una nuova forma di spiegazione, così da rendere il racconto storiografico ancora più completo e complesso.
Perché ciò possa avvenire in modo ancor più compiuto abbiamo a disposizione un altro strumento davvero sorprendente, ossia lo standard che ordina, identifica e mette in correlazione tutte le informazioni relative al soggetto produttore.

3.2 ISAAR (CPF), l'importanza dell'agente.

Questo standard fornisce e organizza le informazioni rispetto ai soggetti produttori e consente un quadro di contesto ancora maggiore riguardo alla fonte archiviata, ordinando le informazioni relative all'agente che ha prodotto la fonte, dunque, per riprendere il pensiero di Topolski, colui che generando la fonte genera un nuovo evento.
«Ogni fonte che viene creata è essa stesso avvenimento storico»[321]
Per comprendere al meglio il portato di ISAAR(CPF) useremo l'esempio riportato nel testo finale di elaborazione dello standard[322], e andremo ad osservare un record dell'*authority file* relativo all'artista Fortunato Depero e presente nel Catalogo Unico del Museo di arte moderna e contemporanea di Trento e Rovereto (CUM), che raccoglie descrizioni di archivi e documenti, di

321 J. Topolski, *Narrare la storia. Nuovi principi di metodologia storica*, Bruno Mondadori, Milano, 1997 p. 220
322 H.L.P. Stibbe, K. Hawoth, *ISAAR (CPF) Standard internazionale per i record d'autorità archivistici di enti, persone, famiglie*, International Council on Archive, II edizione, 2004, https://www.ica.org/sites/default/files/CBPS_Guidelines_ISAAR_Second-edition_IT.pdf ultima consultazione 24/10/2021

opere d'arte e di pubblicazioni, collegate attraverso i record d'autorità dei rispettivi soggetti produttori o autori.
Il CUM può essere consultato all'indirizzo: http://www.mart.tn.it/[323]

Area dell'identificazione

- Tipologia del soggetto produttore: Persona
- Forma/e autorizzata/e del nome: Depero, Fortunato 1892-1960
- Forme del nome normalizzate secondo altre regole: Depero, Fortunato (RICA = Regole italiane di catalogazione per autore)

AREA DELLA DESCRIZIONE

- Date di esistenza: 1892 marzo 30 - 1960 novembre 29
- Storia: Nato a Fondo (Tn) nel 1892 si trasferisce giovanissimo con la famiglia a Rovereto (Tn) dove frequenta la Scuola reale Elisabettina ad indirizzo tecnico-artistico. Ritiratosi al quinto anno, nel 1908 tenta, senza successo, l'esame di ammissione all'Accademia di Belle Arti di Vienna […]
- Luoghi: Nato a Fondo, in Valle di Non (Tn), morto a Rovereto (Tn). Nel 1892, giovanissimo si trasferisce con la famiglia a Rovereto (Tn), dove risiede principalmente fino al 1913. Dalla fine del 1913 e per gli anni di guerra vive a Roma, ad eccezione del breve soggiorno a Capri nel 1917, ospite del poeta G. Clavel, e di quello a Viareggio nel 1918. […]
- Funzioni, occupazioni e attività: Pittore, scenografo, grafico pubblicitario, designer, scrittore, poeta.
- Contesto generale: Il Trentino dei primi decenni del XX secolo fa parte dell'Impero austroungarico e costituisce la propaggine meridionale della Contea principesca del Tirolo. L'Impero è caratterizzato, tra le altre cose, dalla multietnicità e da forti spinte autonomiste, contenute da una linea politica prevalentemente reazionaria e tradizionalista. […]

AREA DELLE RELAZIONI

Prima relazione

- Nomi/codici identificativi degli enti, persone e famiglie correlate: Mazzoni, Angiolo, 1894 – 1979; codice identificativo: IT MART FP 9
- Classificazione della relazione: Associativa
- Descrizione della relazione: collega (architetto) di ambito futurista.

323 Oggi il Catalogo Archivistico del MART è un catalogo integrato da qui l'acronimo CIM e non più CUM

A. Mazzoni aderisce nel 1933 al movimento futurista e nell'anno successivo sottoscrive il Manifesto futurista dell'architettura aerea accanto a F.T. Marinetti e a M. Somenzi, con il quale è condirettore fino al 1935 della rivista d'architettura futurista Sant'Elia.

- Date della relazione: 1932-1936 ca

Seconda relazione

- Nomi/codici identificativi dell'entità correlata: Galleria Museo Depero, 1957 – 1987; codice identificativo: IT MART FE 10
- Classificazione della relazione: Associativa
- Descrizione della relazione: Fondatore
  Fortunato Depero, in collaborazione con il Comune di Rovereto, istituisce nella stessa città la Galleria Museo Depero, ente destinato alla conservazione e valorizzazione della sua opera artistica e documentaria in relazione alla sua adesione al movimento futurista.
- Date della relazione: 1957 -1960

AREA Di CONTROLLO

- Codice identificativo del record d'autorità: IT MART FP 7
- Codici identificativi dell'istituzione responsabile: Museo di arte moderna e contemporanea di Trento e Rovereto (Tn), Italia; Codice identificativo IT MART
- Norme e/o convenzioni: ISAAR (CPF) – International Standard Archival Authority Record For Corporate Bodies, Persons and Families, Second Edition, Canberra, Australia, 27-30 October 2003; - ISO 8601 - Data elements and interchange formats –Information interchange—Representation of dates and times, 2nd ed., Geneva: International Standards Organization, 2000: - RICA (Regole italiane di catalogazione per autore); - Norme per i collaboratori [delle pubblicazioni degli archivi di Stato], Roma, Ministero per i beni culturali e ambientali. Ufficio centrale per i beni archivistici, 1992
- Grado di elaborazione: Versione finale
- Livello di completezza: Completo
- Data di redazione, revisione o cancellazione: 2002/08/30 (prima redazione) 2003/12/20 (ultima revisione)
- Lingua/e e scrittura: Italiano
- Fonti:
  Bibliografia: G. Belli - G. Marzari, Artisti del '900. Protagonisti di Rovereto, Rovereto, Longo Editore, 1991, pp. 40-41 […]

Fonti Archivistiche: Museo di arte moderna contemporanea di Trento e Rovereto, Fondo Fortunato Depero (1894-1960)

- Note sulla compilazione del record: Nome del compilatore del record – Stefania Donati

AREA DI COLLEGAMENTO DEGLI ENTI, PERSONE E FAMIGLIE CON LA DOCUMENTAZIONE ARCHIVISTICA E CON ALTRE RISORSE

Prima risorsa collegata

- Codici identificativi e denominazioni o titoli delle risorse collegate: Denominazione Fondo Fortunato Depero; codice identificativo: IT MART Dep
- Tipologia delle risorse collegate: Fondo archivistico
- Natura delle relazioni: soggetto produttore
- Data delle risorse collegate e/o delle relazioni: 1894-1960

Seconda risorsa collegata

- Codici identificativi e denominazioni o titoli delle risorse collegate: Denominazione Fondo Biblioteca Museo Depero; Codice identificativo: IT MART q – MD
- Tipologia delle risorse collegate: Raccolta libraria
- Natura delle relazioni: Creatore della raccolta
- Data delle risorse collegate e/o delle relazioni: 1910 – 1960

Terza risorsa collegata

- Codici identificativi e denominazioni o titoli delle risorse collegate: Titolo Cartellone per “balli plastici”; codice identificativo: IT MART, Coll Dep MD 0180-B
- Tipologia delle risorse collegate: Dipinto (olio)
- Natura delle relazioni: autore
- Data delle risorse collegate e/o delle relazioni: 1918

[…]

Settima risorsa collegata

- Codici identificativi e denominazioni o titoli delle risorse collegate: Titolo La Casa del Mago. Le arti applicate nell’opera di Fortunato Depero 1920 – 1942. Catalogo della mostra, Rovereto, Archivio del ‘900, 12 dicembre 1992- 30 maggio 1993, a cura di G Belli, Milano, Charta, 1992; Codice identificativo: AR IV MART 92/5a
- Tipologia delle risorse collegate: Catalogo
- Natura delle relazioni: Soggetto

- Data delle risorse collegate e/o delle relazioni: 1992

Ottava risorsa collegata

- Codici identificativi e denominazioni o titoli delle risorse collegate: Titolo B. Passamani, Fortunato Depero, Comune di Rovereto. Musei civici. Galleria Museo Depero, 1981; Codice identificativo q – AR IV MART 1981/5
- Tipologia delle risorse collegate: Monografia
- Natura delle relazioni: Soggetto
- Data delle risorse collegate e/o delle relazioni: 1981

Come possiamo notare lo standard ISAAR(CPF), concentrandosi sul soggetto produttore è costruito per fornire una serie di informazioni di contesto molto utili anche per l'utente non specializzato, che oltre ad ottenere informazioni relative all'archivio e alle fonti in esso conservate, accede ad una serie di informazioni di contesto necessarie per comprendere al meglio la storia dell'artista stesso, non solo, le aree relative alle "risorse collegate" consentono di ampliare il set di fonti utili alla ricerca.

Se su Google scriviamo le parole "Fortunato Depero Archivio" il motore di ricerca come primo risultato ci indicherà il sito ufficiale Depero.it[324], sito in cui però, contrariamente a quanto riportato nell'URL non conduce all'archivio, ma nella descrizione cita l'Archivio Depero custodito al MART. In questa pagina non si trovano link al MART e all'archivio dell'artista, questo probabilmente, a quanto si legge nel sito ufficiale, perché tra il MART e gli eredi dell'artista vi è in corso una trattativa sulla gestione delle opere.

Il primo risultato fornito da Google, nonostante ci dia delle indicazioni non è perfettamente funzionale alla nostra ricerca, o meglio, ci costringe a compierne una nuova.

Nel secondo risultato riportato da Google troviamo il sito del SIUSA, che come abbiamo visto è il Sistema Informativo Unificato per le Sovraintendenze Archivistiche, ecco quindi che Google ci suggerisce di accedere al portale di un sistema informativo. Aprendo il link suggerito da Google ci troveremo immediatamente nella pagina del SIUSA dedicata a Depero, qui avremo le prime informazioni sul fondo: estremi cronologici, datazione, consistenza, storia archivistica, descrizione, ordinamento, e ancor più interessante le indicazioni per strumenti di ricerca con link che rimandano sempre all'interno del SIUSA ma con le descrizioni dell'inventario Fortunato Depero e del CIM-Catalogo Integrato de Mart.

324 https://depero.it/archivio/ ultima consultazione 28/102021

La voce descrittiva che troviamo subito dopo “strumenti di ricerca”, è “siti web” dove troviamo i link per “Edizione nazionale epistolario Alcide De Gasperi” che se cliccato ci porta direttamente nella pagina in cui è possibile fruire il simulacro digitale di una lettera scritta da Depero a De Gasperi[325]. Sotto questo link ne troviamo un altro che ci dovrebbe indirizzare a “Mart. Archivio del 900”, ossia l'archivio che come abbiamo potuto vedere nell’esempio, conserva i documenti relativi all’artista, uso il condizionale perché in realtà questo link ci porta in un *404 Not Found*. Ovviamente noi sappiamo che la pagina Web dell’archivio del MART esiste, il problema probabilmente risiede in un possibile mancato aggiornamento del collegamento con un URL attivo. La cosa interessante è che se avessimo trovato un collegamento strutturato tramite RDF, dunque con triple che connettono oggetti informativi univoci composti da dati autoconsistenti, il collegamento tra le unità informative non avverrebbe tramite un link, non sarebbero le pagine web ad essere collegate tra di loro, ma direttamente gli oggetti digitali, che, se pur come atomi indipendenti portano al loro interno le informazioni relative anche al loro vincolo archivistico.

Affronteremo più avanti il tema del Web semantico, dei vocabolari comuni e della conseguente possibilità di costruire ontologie digitali, prima è necessario comprendere meglio cosa siano EAD e EAC-CPF - gli strumenti tecnici che ci consentono di interagire a livello digitale con gli standard di descrizione - e come funzionino gli standard di metadatazione.

Quello che per ora ci interessa è comprendere che il Web nelle sue articolazioni, se organizzato e ordinato, si mostra come uno strumento capace di rendere accessibili e fornire informazioni di tipo storico a tanti livelli differenti.

3.3 EAD

Per procedere nella nostra analisi è bene puntualizzare che i principi che regolano lo sviluppo di strumenti di ricerca, sono indipendenti dagli strumenti digitali oggi a disposizione o in fase di sviluppo, questo vuol dire che non è il mezzo tecnico che plasma il processo logico, ma al contrario è lo strumento che deve essere in grado di rispondere a specifici principi e funzionamenti.

I principi che regolano la preparazione di strumenti di ricerca come EAD e EAC sono:

a) consentire l’accesso al materiale archivistico fornendo agli utenti informazioni relative ad esso;

b) produrre strumenti di ricerca accurati, coerenti e auto-esplicativi;

c) presentare il contesto e il contenuto del materiale archivistico descritto applicando le regole della descrizione multilivellare.

325 https://epistolariodegasperi.it/#/archivio_digitale/cerca?search=true ultima consultazione 30/10/2021

Uno degli aspetti innovativi del processo di standardizzazione è l'affermazione del concetto di *finding aid system*, cioè della realizzazione di sistemi in cui ogni strumento, dalle tradizionali descrizioni di fondi alle liste d'autorità, gli indici, i tesauri, ecc., sia progettato in raccordo con gli altri così da completarsi, accrescersi e supportarsi reciprocamente, al fine di rendere i fondi conservati completamente accessibili in tutti i loro dettagli e tutti i loro aspetti. E' quasi lapalissiano sottolineare come e quanto il concetto di *finding aid system* sia fondamentale per iniziare a ragionare sulla costruzione di complessità informazionali collegate tra di loro allo scopo di produrre un insieme di informazioni di contesto di approfondimento delle fonti d'archivio, anche esse inserite all'interno di questa rete relazionale.

Come abbiamo già detto, gli standard nascono indipendentemente dalla tecnologia dell'informazione e come abbiamo potuto osservare direttamente forniscono modelli che sono in prima battuta concettuali. Nella trasposizione di questi modelli in un ambiente digitale, come ci ricorda Ilaria Barbanti, è necessario tenere conto che ciò avvenga mediante sistemi di codifica dei dati d'archivio che per essere affidabili devono rispettare specifici requisiti:

- l'esplicitazione della struttura logica che sottende il contenuto e la forma con cui esso viene rappresentato
- l'indipendenza dell'hardware, ovvero una particolare architettura elaborativa (processore), o da un particolare supporto digitale (disco magnetico, disco ottico, ecc.), e dai software, siano essi sistemi operativi o applicazioni deputate alla creazione, analisi, manipolazione visualizzazione di testi elettronici.[326]

Adottare uno standard dunque diventa un'occasione per facilitare lo scambio del contenuto informativo delle risorse in ambiente elettronico, seguendo una logica di accesso condiviso delle informazioni stesse.

L'avvento e la rapida diffusione di internet sta trasformando in tutto il mondo la comunicazione tra studiosi. Per cogliere appieno le opportunità che questo sviluppo offre, gli archivi, i musei e le biblioteche dovranno accettare e sviluppare standard che assicurino una comunicazione tanto utile quanto duratura. Uno standard in grado di codificare anche strumenti per la ricerca[327].

326 I. Barbanti, Il motore di un'evoluzione possibile. L'informatica spinge gli archivi, in AA.VV., A cura di, L. Pezzica, F. Valacchi, Dimensioni archivistiche. Una piramide rovesciata, Editrice Bibliografica, Milano 2021, p. 104

327 D. Pitti, *Una ricchezza in comune verso uno standard che codifichi strumenti per la ricerca*. Archivi & Computer 2:154-62, p. 154

Così scriveva nel 1994 Daniel Pitti, il padre del progetto che avrebbe portato all'elaborazione dell'Encoded Archival Description (EAD). A fronte della pluralità e della varietà di forme in cui si presentano inventari e guide, EAD ha cercato di definire una struttura comune di riferimento all'interno della quale registrare le informazioni di ciascuno strumento di ricerca.
Possiamo quindi considerare EAD un Data Structure Standard[328] pensato per favorire l'interscambio e l'accessibilità a lungo termine delle informazioni descrittive e per renderle disponibili ad utenti locali e remoti[329].
Nato per il recupero retrospettivo e per la pubblicazione in formato elettronico degli strumenti di ricerca cartacei, lo standard ha successivamente supportato l'elaborazione e lo scambio di descrizioni nativamente digitali, grazie all'impiego di tecnologie quali l'eXtensible Markup Language (XML), che consentono una *platform independence*[330] e garantiscono la persistenza della struttura e del contenuto delle descrizioni e la loro accessibilità e validità nel tempo. In tale prospettiva, EAD, la cui prima versione risale al 1998, sostituita poi da una seconda versione pubblicata nel 2002, ha assolto alla funzione di lingua franca della descrizione archivistica consentendo il dialogo a livello nazionale e internazionale e la realizzazione di molteplici progetti.
Come si legge sul sito dell'Istituto Centrale per gli Archivi, ICAR

> EAD si è proposto come uno strumento per la conversione e la pubblicazione in formato elettronico di strumenti di ricerca archivistici prodotti originariamente su supporto cartaceo e la loro pubblicazione in formato elettronico nonché per l'elaborazione e lo scambio di descrizioni archivistiche in formato nativamente digitale. L'adozione di standard e tecnologie come XML che consentono la conservazione e la comunicazione dei dati, indipendentemente da specifiche piattaforme hardware e software, mira a garantire la persistenza della struttura e del contenuto delle descrizioni, e la loro accessibilità e validità nel tempo[331]

Come abbiamo visto per ISAD e ICAAR, gli standard, nell'avanzare del tempo evolvono attraverso modifiche di varia natura, modifiche che spesso nascono dalla volontà di avviare un processo di semplificazione, questo anche perché è sempre più chiaro il legame biunivoco che esiste tra *front-end* e *back-end*, legame che porta sviluppatori e archivisti impegnati nell'elaborazione degli standard o dei codici, a dover o forse voler, guardare sempre con maggiore attenzione anche all'utente finale, dunque al fruitore che entrerà in relazione con il sito o il portale del sistema informavo. Non solo, l'evoluzione degli standard è legata a doppio filo anche con lo sviluppo di

328 K. Kiesling, *EAD as an Archival Descriptive Standard,* American Archivist 60/3: 344-354
329 D. Pitti, *Encoded Archival Description: The Development of an Encoding Standard for Archival Finding Aids,* The American Archivist 60, pp. 268-283, 1997
330 G. Michetti, *Il linguaggio per la descrizione archivistica: analisi di un* caso, Archivi & Computer 1, pp. 7-33, 2000
331 https://www.icar.beniculturali.it/standard/standard-internazionali/ead ultima consultazione 10/11/2021

nuove tecnologie come il Web Semantico e i Linked Data, sistemi informatici che portano gli standard di descrizione sempre di più verso nuove possibilità di interoperabilità tra le informazioni, per questo anche EAD ha subito nel tempo diverse modifiche.
Sul sito ufficiale di EAD è possibile vedere tutte le modifiche sviluppate sullo standard, dalla prima versione del 2002 alla corrente versione denominata EAD3[332]

Nei quasi quindici anni che ci separano dalla seconda versione le possibilità di offrire una descrizione archivistica più ricca e articolata sono notevolmente cresciute. Pensiamo al rilascio dell'Encoded Archival Context - Corporate bodies, Persons, and Families (EAC-CPF), alla comparsa e diffusione crescente di *collection management system*, ai vantaggi derivanti dal rapido evolvere del settore tecnologico e alle nuove opportunità e scenari cui sembrano aprire i Linked Open Data. Proprio in virtù di tali motivazioni, nel 2010 la Society of American Archivists (SAA) ha chiesto al Technical Subcommittee for Encoded Archival Description (TS-EAD) di procedere alla revisione dello standard con l'obiettivo di accrescerne chiarezza e coerenza e garantire una maggiore interoperabilità e un generale miglioramento delle funzionalità in ambienti internazionali e multilingue[333]. Si cercò allora la collaborazione dell'intera comunità archivistica invitata a contribuire al processo di revisione con osservazioni e suggerimenti maturati sulla base dell'esperienza di EAD 2002.
Tenendo conto dei 127 commenti pervenuti il TS-EAD ha avviato il lavoro di revisione e ha definito alcuni obiettivi cruciali:

- semplificare lo standard
- eliminare le ambiguità
- favorire l'interoperabilità
- collegare più facilmente EAD con altri standard, come EAC – CPF

Dopo cinque anni di lavoro, nell'estate 2015, la nuova versione EAD3 è stata adottata come standard dalla SAA[334].

Nelle pagine precedenti ho riportato integralmente la struttura di ISAD(G) e ISAAR(CPF) e come abbiamo potuto vedere il processo di strutturazione di questi standard è un processo di natura prettamente logica, la trasposizione di questi sistemi in XML, un metalinguaggio utilizzato per creare nuovi linguaggi atti a descrivere documenti strutturati, potrebbe essere più difficile da leggere e da comprendere.

332 https://www.loc.gov/ead/ ultima consutazine 10/11/2021
333 A. Bunnell, *Review of Encoded Archival Description Tag Library – Version EAD3,* «Journal of Western Archives» 7: 1-4 Jodi. 2016
334 F. Pala, *Lo standard EAD3 per la codifica dei dati archivistici: qualche novità e molte conferme*, JLIS.it 8, 3, DOI: 10.4403/jlis.it-12407, 2017 file:///D:/backup%2028_10_18/Documenti/unimore/dottorato%2017-18/tesi/fonti%20digitali/Pala_%20EAD_2017_12407-22919-3-PB.pdf ultima consultazione 15/11/2021

Gli elementi che compongono EAD sono circa 140, la maggior parte dei quali è opzionale. La ragione essenziale della presenza di un numero così elevato di elementi è da ricercare nello scopo che ha portato alla realizzazione dello standard, ossia avere un mezzo per poter codificare tutto quello che si può trovare all'interno di uno strumento di ricerca, che sino ad ora ha avuto sempre la veste grafica di un libro a stampa. In tal senso EAD consente di aggiungere, oltre agli elementi descrittivi conformi alle norme ISAD, anche quegli elementi che precedono la descrizione archivistica vera e propria, come la presentazione generale del lavoro svolto, le citazioni, le note a piè di pagina, e così via.
Di seguito, grazie al lavoro di sintesi svolto da Barbanti[335] riporto gli elementi di cui è composto lo standard e per ogni elemento, come già per gli standard precedentemente analizzati, vi è una brave spiegazione legata al funzionamento di ogni specifico elemento.

L'elemento contenitore più esterno, contiene i seguenti sottoelementi:

1. <eadheader> (EAD Header, intestazione dello documento - obbligatorio), per fornire le informazioni sullo strumento di ricerca (titolo, compilatore, data di compilazione, edizione, etc.);
2. <frontmatter> (Front Matter, frontespizio - opzionale): per strutturare gli elementi di un eventuale "frontespizio" personalizzato. Tale elemento è il più esposto a critiche, poiché non rende ragione di una partizione logica del documento, bensì formale;
3. <archdesc> (Archival Description, descrizione archivistica - obbligatorio) per le informazioni e le descrizioni relative al corpus (collezione, fondo, serie, etc.) di materiale archivistico vero e proprio. Ad un livello alto dunque è possibile codificare quelle informazioni che servono a descrivere la natura e la struttura generali dello strumento di ricerca, mentre le informazioni specifiche relative alla documentazione sono fornite scendendo al livello dell'elemento <archdesc>. Esso costituisce l'elemento contenitore dell'informazione riguardante il contenuto, il contesto e la consistenza della documentazione archivistica, come pure le informazioni supplementari che possono facilitarne l'accesso e l'utilizzo da parte degli utenti. In altre parole, tale sezione viene dedicata alla descrizione archivistica del complesso documentario per il quale è stato redatto lo strumento di ricerca.

335 I. Barbanti, *La codifica informatica dei dati di archivio: EAD ed EAC,* «Regesta» Scrinia, 2/3 2005 https://www.regesta.com/wp-content/uploads/documentazione/Barbanti-EADEAC2005.pdf ultima consultazione 22/11/2021

L’elemento <archdesc> presenta un attributo obbligatorio LEVEL che permette di specificare il livello dell’unità di descrizione codificata. Il valore dell’attributo LEVEL viene scelto all’interno di un set di valori predefiniti previsti dallo standard:

- collection
- fonds
- class
- recordgrp
- series
- subfonds
- subgrp
- subseries
- file
- item
- otherlevel

Nel caso in cui non sia possibile scegliere tra nessuno dei valori previsti, si riporta l’attributo OTHERLEVEL con, al suo interno, l’effettivo livello gerarchico dell’unità di descrizione.

Una volta terminata la descrizione generale dello strumento di ricerca sul complesso documentario, si giunge alla codifica dei raggruppamenti gerarchici e delle relative unità di descrizione del materiale attraverso l’elemento <dsc> (Description of Subordinate Components, descrizione delle componenti subordinate).

Per poter dichiarare la tipologia di descrizione effettuata, è previsto un attributo obbligatorio TYPE, il quale può assumere 4 valori:

1) analyticover (analytic overview, descrizione sommaria) segnala che la descrizione delle parti componenti sarà limitata alla codifica di un elenco di elementi presenti su un unico livello;

2) in-depth (in profondità) è riferito ad un tipo di descrizione che, arrivando fino al livello di item, segnala i numeri delle unità di condizionamento (container) per ciascuna delle unità di descrizione;

3) combined segnala che la descrizione di una serie (ovvero un raggruppamento logicoarchivistico “alto”) è immediatamente seguita dalla descrizione delle sue parti componenti;

4) othertype lascia infine la possibilità di indicare tipologie descrittive diverse da quelle previste. I livelli descrittivi di <dsc> si aprono al suo interno, nell’elemento <c> (Componente), o unità di descrizione.

<c> presenta la stessa potenzialità informativa disponibile in <archdesc>, ma in questo caso si accede alla vera e propria descrizione nel rispetto della gerarchia in cui si articola la

documentazione, dunque dalle serie si arriva all'unità archivistica di base (file), o all'unità documentaria (item).

All'interno di <c> la compilazione dell'attributo LEVEL, pur se raccomandata dalle Application Guidelines, non è obbligatoria.

Va notato come alla numerazione delle componenti subordinate non vada associata la designazione di un preciso livello logico-archivistico nell'ambito dello strumento di corredo (fondo, serie, sottoserie, fascicolo…): è plausibile infatti che nello stesso strumento di ricerca un elemento possa rappresentare, in un determinato punto della struttura gerarchica del materiale documentario in esso descritto, l'unità di descrizione "fascicolo", laddove in un punto diverso della gerarchia il medesimo livello di descrizione (cioè il fascicolo) possa trovarsi ad essere codificato come <c05>.

Sia <archdesc> (per la descrizione del complesso documentario nel suo insieme) che <c> (per la descrizione dettagliata della documentazione) hanno gli stessi sottoelementi.

Seguendo i principi della descrizione multilivello, i sottoelementi disponibili a livello di <c> si ripetono dunque inalterati ai livelli descrittivi inferiori, e l'informazione viene ad essere ereditata dal livello gerarchico più ampio a quello subordinato, conformemente anche a quanto indicato nelle ISAD(G).

Tra questi sottoelementi, <did> (Descriptive Identification, elementi di identificazione) contiene elementi informativi della descrizione archivistica del complesso documentario essenziali per garantire l'efficace descrizione di un'unità archivistica a qualsiasi livello. <did> può contenere:

- abstract: sommario del materiale;
- physdesc (Phisical Description, descrizione fisica): identifica la consistenza e tutte le informazioni sulle caratteristiche fisiche del materiale descritto. L'informazione può essere data come testo libero direttamente al suo interno, oppure suddivisa negli elementi <dimension> e <extent>, (= consistenza, quantità del materiale descritto), <genreform> (= genere e caratteristiche del formato fisico, stile e tecnica del contenuto, sia funzionale, sia intellettuale), <physfacet> (= informazione sull'aspetto esteriore/fisico del materiale es: colore, materia, sostanza, tecniche di creazione)
- note: campo note
- repository: istituzione o ente responsabile della conservazione e dell'accesso intellettuale al materiale descritto;
- origination (soggetto produttore): permette di indicare gli individui responsabili della creazione o dell'accumulo del materiale descritto. Nell'attributo LABEL è possibile indicare quale rapporto ha con le carte (ad esempio, soggetto produttore);

- unitdate (Date of the Unit, data dell'unità di descrizione): estremi cronologici del complesso. Nell'attributo NORMAL viene riportata la datazione nel formato normalizzato, nella dicitura "aaaa/mm/gg", mentre in TYPE è possibile specificare il tipo di data (inclusiva - "da …a …" - o datazione singola);
- unitid (ID of the Unit, codice identificativo);
- unititle (Title of The Unit, titolo o denominazione dell'unità di descrizione)
- container (contenitore) numero di unità di condizionamento (busta, scatola ecc.). L'attributo TYPE permette di specificare il tipo di unità (busta, faldone ecc.)
- physloc (Physical Location, localizzazione fisica) luogo della custodia/conservazione fisica del materiale descritto
- dao e daogrp (Digital Archival Object Groups – singolo oggetto e gruppo di oggetti digitali): servono per codificare uno o più collegamenti ipertestuali ad allegati digitali.

Una volta fornita attraverso <did> una descrizione basilare della documentazione, gli altri sottoelementi di <archdesc> e <c> hanno il compito di consentire la codifica, all'interno del livello di riferimento, di tutte quelle informazioni tipicamente presenti in un tradizionale strumento di ricerca archivistico, riguardanti il contenuto, il contesto, la consistenza, i criteri di ordinamento, e la gestione amministrativa in tutti i suoi aspetti, dall'acquisizione alle condizioni di accesso. Ecco di seguito quelli maggiormente significativi:

- accessrestrict (Restriction on Access, restrizioni di accesso) e userestrict (Restriction on Use, restrizioni di utilizzo del materiale)
- acqinfo (Acquisition Information, informazioni sull'acquisizione): informazioni su donazioni, trasferimenti, acquisto e deposito, in generale sulle fonti da cui sia pervenuto il materiale;
- altformavail (Alternative Form Available, disponibilità di formati alternativi): informazioni sulle copie dei documenti presenti su altri formati;
- appraisal (scarto): informazioni relative alla selezione del materiale e allo scarto, con descrizione di decisioni e metodologie adottate;
- accruals (versamenti): informazioni sui versamenti passati e sugli incrementi anche futuri del materiale;
- arrangement (sistemazione): descrizione della fascicolazione del materiale (in ordine cronologico o alfabetico);
- bibliography (bibliografia)
- bioghist (Biography and/or History, storia e/o biografia)

- controlaccess (Control Access Headings, intestazioni di accesso controllate) elemento in cui vengono raggruppati tutti i nomi, soggetti, argomenti che illustrano al meglio la natura, l'ambito e la rilevanza della documentazione descritta. Permette, in altre parole, di codificare le chiavi d'accesso alla documentazione
- custodhist (Custodial History, storia della custodia): informazioni sulla catena dei passaggi di proprietà e di custodia della documentazione, significativi in termini di provenienza, integrità e interpretazione delle fonti
- descgrp (Description Group, gruppo di elementi descrittivi): tale elemento raccoglie tutti gli elementi sopra indicati eccetto <dsc>, <dao> e <daogrp>;
- fileplan (File Plan): schemi di classificazione per il recupero del materiale (titolario);
- otherfindaid (Other finding aid, altro strumento di ricerca): riferimenti ad altri strumenti di ricerca (elenchi di versamento e di consistenza, inventari, ecc.);
- processinfo (Processing Information, informazioni sull'elaborazione): informazioni sull'accesso, l'ordinamento, la descrizione, la conservazione, la disposizione, o ogni altra operazione che rende disponibile il materiale all'utente;
- odd (Other Descriptive Data, altri dati descrittivi): utilizzato per codificare informazioni che non trovano posto negli altri elementi, o ancora permette di aggiungere informazione di varia natura, che risulta difficile scorporare nei vari elementi;
- organization: struttura della sedimentazione della carte o schema delle serie (livelli inferiori);
- scopecontent (Scope and Content, ambito e contenuto): corrisponde alla nota archivistica, e viene usato, ad esempio, quando la documentazione viene versata in momenti successivi. Ha come sottoelementi arrangement e organization;
- relatedmaterial (materiale correlato): materiali ritenuti utili per comprendere quello descritto, fisicamente e logicamente separati, non facenti parte del corpus documentario (altri enti produttori);
- separatedmaterial (materiale separato): documenti associati per provenienza dallo stesso ente produttore ma conservati altrove o separatamente

In seguito alle revisioni di EAD 2002, la comunità scientifica ha continuato a elaborare lo standard EAD3, per renderlo «lo standard degli standard»[336], dunque, oltre gli elementi segnalati da Barbanti, dobbiamo analizzare i cambiamenti nel corso dello sviluppo di EAD3.

336 F. Valacchi, *Pezzi di cose nel mondo. Il processo di integrazione delle descrizioni archivistiche nei sistemi interculturali*, «Jlis.it» 7,2 2016 Doi: http://dx.doi.org/10.4403/jlis.it-11529, https://www.jlis.it/article/view/11529/10911 ultima consultazione 22/11/2021

In tal senso deve essere ad esempio letta la rimozione di diversi elementi e attributi che facevano unicamente riferimento alla formattazione: si pensi a <frontmatter> o <runner>. La rimozione non è stata però completa: permangono infatti attributi ed elementi, fra i quali @align, @render, nonostante questo genere di istruzioni possa essere fornito ricorrendo ai fogli di stile. Le motivazioni, come è stato chiarito a chi scrive da Mike Rush nel corso di uno scambio informale di riflessioni, sono da ricercare nella volontà da parte del TS-EAD – a fronte della constatazione dell'uso frequente in alcuni contesti degli elementi di formattazione – di non operare un taglio radicale con la precedente versione. Sempre in relazione alla volontà di superare il "peccato originale", potrebbe essere letta anche l'introduzione dei nuovi e diversi elementi, fra i quali <maintenancehistory>, <maintenancestatus>, <maintenanceagency>, che consentono, o meglio impongono – visto che sono definiti elementi obbligatori – di tenere traccia di tutta una serie di metadati relativi all'istanza EAD […] Come auspicato da diversi esponenti della comunità archivistica internazionale che hanno fatto pervenire i propri suggerimenti al TS-EAD durante la fase di revisione di EAD 2002, il nuovo modello dati risulta semplificato specialmente nell'articolazione di alcuni elementi, e maggiormente allineato ad altri standard, quali EAC-CPF. Inoltre la nuova versione presenta una maggiore attenzione per la codifica di dati che siano machine-readable. Esemplificativa in tal senso è stata l'introduzione degli elementi fortemente "strutturati" <physdescstructured> e <unitdatestructured>. Infine, pur restando uno standard basato su XML, la nuova versione di EAD introduce nuovi elementi e attributi che supportano la creazione dei linked data. Oltre all'introduzione dell'elemento <relations>, che mette in luce come l'elaborazione di EAD3 non sia stata avulsa dai dibattiti internazionali sugli standard di descrizione, si segnala l'introduzione degli attributi @identifier e @relator nei sottoelementi di appartenenti al group element access. I due attributi lavorano in tandem per definire le relazioni necessarie a generare linked data, (Implementing EAD3: Search and Exploration)[337]

Di seguito una tabella che mette in evidenza le correlazioni tra gli elementi descrittivi per come gli abbiamo osservati in ISAD(G) sulla sinistra, e la loro trasposizione in EAD sulla destra, la tabella è fornita dal sito ufficiale di EAD[338]. Per non appesantire troppo la lettura, già di per sé complessa in questo passaggio così tecnico, riporto solo parte delta tabella, come abbiamo visto gli elementi che compongo gli standard son molto di più di quelli che osserviamo nell'immagine, ma già così credo si riesca ad avere un'idea di come funzioni il processo di "traduzione".

337 F. Pala, *Lo standard EAD3 per la codifica dei dati archivistici: qualche novità e molte conferme,* «Jlis» 2017, DOI: 10.4403/jlis.it-12407

338 https://www.loc.gov/ead/tglib/appendix_a.html ultima consultazione 22/11/2021

A.1. ISAD(G) to EAD

ISAD(G)	EAD
3.1.1 Reference code(s)	<eadid> with COUNTRYCODE and MAINAGENCYCODE attributes<unitid> with COUNTRYCODE and REPOSITORYCODE attributes
3.1.2 Title	<unittitle>
3.1.3 Dates	<unitdate>
3.1.4 Level of description	<archdesc> and <c> LEVEL attribute
3.1.5 Extent and medium of the unit	<physdesc> and subelements <extent>, <dimensions>, <genreform>, <physfacet>
3.2.1 Name of creator	<origination>
3.2.2 Administrative/Biographical history	<bioghist>
3.2.3 Archival history	<custodhist>
3.2.4 Immediate source of acquisition	<acqinfo>
3.3.1 Scope and content	<scopecontent>
3.3.2 Appraisal, destruction and scheduling	<appraisal>
3.3.3 Accruals	<accruals>
3.3.4 System of arrangement	<arrangement>
3.4.1 Conditions governing access	<accessrestrict>
3.4.2 Conditions governing reproduction	<userestrict>
3.4.3 Language/scripts of material	<langmaterial>
3.4.4 Physical characteristics and technical requirements	<phystech>
3.4.5 Finding aids	<otherfindaid>
3.5.1 Existence and location of originals	<originalsloc>
3.5.2 Existence and location of copies	<altformavail>
3.5.3 Related units of description	<relatedmaterial><separatedmaterial>
3.5.4 Publication note	<bibliography>
3.6.1 Note	<odd><note>
3.7.1 Archivist's note	<processinfo>
3.7.2 Rules or conventions	<descrules>
3.7.3 Date(s) of descriptions	<processinfo><p><date>

Quello che avviene con la costruzione di EAD è un processo di trascrizione in linguaggio informatico di un processo logico nato in seno alla discussione archivistica per realizzare degli standard internazionali. Il risultato di questo lavoro è che oggi, grazie a EAD ci è possibile usare ISAD come fosse uno strumento di ricerca aperto. Tutte le informazioni organizzate in modo così scrupoloso all'interno dello standard divengono accessibili tramite il Web. Oltre a ciò è importante ricordare il ruolo del finding aid sistem, che consente l'interazione tra tutti i sistemi descrittivi e tutti i record di descrizione. La conoscenza storica, intesa non come storiografia, ma come tracce che dal passato sono giunte a noi, viene dunque organizzata in modo tale divenire quasi autorappresentante a partire da semplicissime ricerche all'interno di sistemi digitali. Certo è che se per EAD è fondamentale un processo multlivellare gerarchico di ordinamento, come abbiamo potuto vedere osservando *LEVEL*, ciò che propone il Web semantico va ad intaccare in modo deciso questa base strutturale, dal momento che i principi del Web 3.0 portano a far interagire le informazioni tra di loro non attraverso sistemi gerarchici multilivellari, ma tra sistemi liberi e multidimensionali, dove le gerarchie vengo destrutturate. Di EAD si è però detto che vuole essere lo standard degli standard, o meglio la rappresentazione dello standard originario ISAD, ciò non viene realmante meno con l'avvento del Web semnatico, come vedremo infatti, il campo <archdesc> e il campo <c>, solo per citare quelli più importanti verranno incorporati all'interno dell'oggetto informativo.

3.4 EAC-CPF

L’elaborazione di EAC-CPF Encoded Archival Context for Corporate Bodies, Persons, and Families[339] un modello dati per la codifica elettronica in formato XML dei record di autorità archivistici risale al marzo del 2001, due mesi dopo l’avvio ufficiale delle attività di revisione della prima edizione delle ISAAR. Un gruppo di lavoro formato da specialisti del settore di ambito europeo e anglosassone, lavorò per creare un formato standard delle descrizioni sui soggetti produttori, in modo da renderle ricercabili e accessibili in ambiente elettronico, e da poterle collegare con altre entità soggetti e con la documentazione d’archivio da essi prodotta. Sebbene siano state pubblicate due versioni, alfa e beta, le attività di elaborazione del modello dati non hanno ancora portato all’emanazione di uno standard definitivo. Ciò nonostante, il modello dati EAC viene citato nella seconda versione dello standard ISAAR (CPF), nei confronti del quale risulta perfettamente compatibile, nel documento viene rilevata l’importanza che assume un modello informatico capace di avere una codifica elettronica delle informazioni d’autorità relative ad archivi per lo scambio e la condivisione delle stesse, grazie all’utilizzo di un adeguato formato di comunicazione quale è XML.

La codifica in formato XML permette di realizzare in un unico ambiente informatico il modello logico del collegamento tra le informazioni archivistiche o documentarie in genere e quelle delle voci d’autorità, così come prospettato dalle ISAAR.

EAC si pone come struttura che estende ed arricchisce il modello dati EAD, il quale, come abbiamo visto, presenta possibilità di codifica delle informazioni sui soggetti produttori limitate alla compilazione dell’unico elemento <bioghist >.

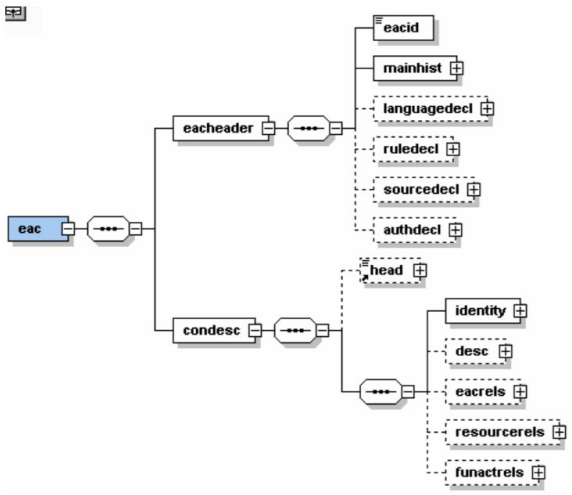
Una delle caratteristiche innovative di EAC è quella per cui viene dato rilievo non tanto all’intestazione quanto all’entità, come elemento univoco e identificativo del soggetto produttore. In più, le entità in quanto tali devono essere chiavi di accesso alla documentazione per individuare non soltanto il contesto in cui agirono i soggetti produttori, ma anche altri ruoli di responsabilità assunti, in conformità alle «linee guida per collegare record d’autorità archivistici alle descrizioni dei complessi archivistici prodotti dall’entità descritta e/o ad altre risorse informative delle quali è responsabile o che sono ad essa relative»[340].

Anche il modello dati EAC è composto da un elevato numero di elementi (80), molti dei quali sono aggiuntivi rispetto a quelli previsti dalle ISAAR, anche per accogliere la possibilità di codificare, con lo stesso modello dati, le voci d’autorità ed i relativi elementi descrittivi provenienti da altri ambiti disciplinari, non solo archivistici.

339 https://eac.staatsbibliothek-berlin.de/ pagina ufficiale progetto EAC, ultima consultazione 23/11/2021

340 H.L.P. Stibbe, K. Hawoth, *ISAAR (CPF) Standard internazionale per i record d'autorità archivistici di enti, persone, famiglie*, International Council on Archive, II edizione, paragrafo 6, 2004 https://www.ica.org/sites/default/files/CBPS_Guidelines_ISAAR_Second-edition_IT.pdf ultima consultazione 22/11/2021

Adesso non rimane che andare a vedere anche per EAC i diversi elementi di cui è composto, così da avere un quadro chiaro di come sia possibile anche tecnicamente, fruire informazioni archivistiche dettagliate rispetto ai soggetti produttori, utilizzando il Web come strumento di ricerca.
Per semplificare il nostro lavoro condivido lo schema base di funzionamento di EAC-CPF proposto da Barbanti[341]



L'elemento centrale che funge da cappello è l'elemento <eac> che si riferisce a Person per i nomi di persona, Corporatebody per gli enti e Family per le famiglie. A sua volta l'elemento <eac> è composto da due sezioni priciplai:

- <eacheader > che comprende gli elementi identificativi del record di autorità:
    1. <eacid> codice identificativo
    2. <mainhist> (Main History) elemento in cui sono riportate le informazioni di creazione e aggiornamento del record
    3. altri elementi opzionali, come la dichiarazione della lingua in cui sono riportate le informazioni, delle regole e delle fonti a cui si fa riferimento per la descrizione, e, infine, l'indicazione dell'authority a cui si vuole rimandare <languagedecl> <ruledecl> <sourcedecl> <authdecl>
- <condesc> in cui sono riportati gli elementi identificativi della voce <identity> e quelli descrittivi, che ne completano i dati sul contesto di provenienza <desc> e su eventuali relazioni <eacrels> <resourcerels > <funactrels>.

In particolare, all'interno di <identity> vi sono gli elementi in cui sarà possibile riportare la denominazione autorizzata e le varianti, a seconda della tipologia di entità:

- <pershead> per i soggetti produttori di tipo "persona";

341 I. Barbanti, *La codifica informatica dei dati d'archivio: EAD ed EAC1*, «Regesta.com. New Media and Historical heritage», 2-3 2005, p. 10 http://media.regesta.com/dm_0/regesta/regestacms/generic/documentazione/Barbanti-EADEAC2005.pdf ultima consultazione 22/11/2021

- <corphead> per i soggetti produttori di tipo “ente”;
- <famhead> per i soggetti produttori di tipo “famiglia”.

Il modello dati consente di codificare allo stesso livello tutte le denominazioni che nel tempo hanno permesso di identificare la stessa entità descritta nel record di autorità, potendo distinguere una denominazione con valore di intestazione principale da altre intestazioni parallele (se ad esempio vengono riportate denominazioni in diverse lingue), o anche da denominazioni secondarie. Per tali ragioni è anche possibile codificare e associare le date di esistenza a ciascuna denominazione con cui l’entità è stata identificata univocamente in un determinato periodo <usedate>.

Altri elementi descrittivi sono <place>, per l’indicazione del luogo o del territorio in cui ha operato ed è esistito l’oggetto descritto <nameadd>, per codificare denominazioni secondarie che ne completano l’identificazione, e <existdate>, per riportare gli estremi cronologici in cui è stato attivo.

All’interno dell’elemento <desc> vi sono gli elementi che permettono di codificare le informazioni storiche, di contesto e ogni altro dato che permetta di avere una scheda informativa completa, relativamente all’entità che si sta codificando. Tre sono gli elementi disponibili per raggruppare, al loro interno, la specifica descrizione di ogni tipologia di entità: <persdesc>, (per ogni nome di persona), <corphead> (per ogni nome di ente) e <famhead> (per ogni nome di famiglia).

Di particolare interesse, come abbiamo visto osservando la struttura di ISAAR, sono gli schemi che consentono l’interrelazione tra fonti pertinenti al soggetto produttore e altre entità in EAC-CPF queste vengono elaborate grazie agli elementi

- <eacrels> per la codifica delle relazioni tra più entità soggetto produttore
- <resourcerels> per la codifica delle relazioni tra entità soggetto produttore e le risorse documentarie, con ciò intendendo documentazione archivistica <archunit>, bibliografica <bibunit>, musueale <musunit> o di altro tipo, codificata elettronicamente
- <funactrels> per la codifica delle relazioni tra entità soggetto produttore e altre funzioni o attività

La tipologia di relazione viene espressa come valore dell’attributo RELTYPE, corredato da una lista di valori predefiniti: - per <eacreal>

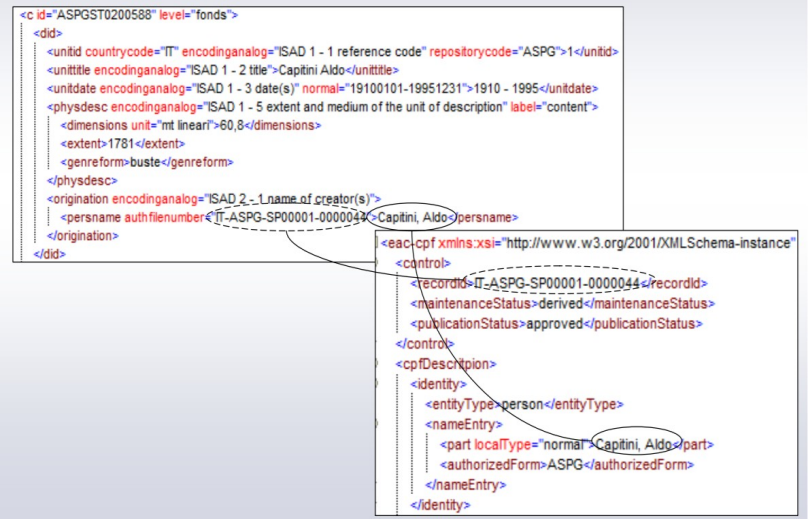
- superior
- subordinate
- earlier
- later
- associative
- parent

- child
- identity
- other

per <resourcerel>

- origination
- destruction
- control
- causa
- subject
- other

Nella figura seguente[342] possiamo osservare uno schema esemplificativo di come si costruisce un collegamento tra dati d'archivio e soggetto produttore, dunque come comunicano tra di loro gli standard EAD3 e EAC-CPF



L'interoperabilità tra sistemi archivistici e la possibilità di costruire relazioni strutturate tra informazioni fa sì che oggi, come abbiamo potuto vedere, la ricerca sul Web attraverso motori

342 Archivio di Stato di Perugia, *I modelli dati XML per la codifica elettronica dei dati d'archivio: EAD e EAC-CPF,* 2013, http://www.archiviodistatoperugia.it/sites/default/files/archivistica_informatica_2.pdf *ulti*ma consultazione *25/11/2021*

generalisti possa dare dei risultati capaci di indirizzare il fruitore su sistemi informativi, che grazie al granulare livello di descrizione consento di orientarsi tra diverse informazioni.

L'immagine condivisa sopra, se pur non ci aiuta ad approfondire ulteriormente i processi informatici che consentono il dialogo tra EAD e EAC, ci consente con grande semplicità di capire come questi sistemi possano comunicare al fine di garantire al fruitore un numero davvero notevole di informazioni in relazione alla risorsa digitale.

L'Istituto Centrale per gli Archivi ha predisposto un pacchetto integrato di formati di scambio in linguaggio XML allo scopo di permettere un elevato livello di interoperabilità fra i principali applicativi e sistemi di descrizione archivistica sviluppati nel nostro paese e consentire perciò l'esportazione e l'importazione dei dati da e verso di essi.

Tale pacchetto, oltre favorire una maggiore integrazione fra i dati residenti in quegli applicativi e i sistemi, si propone di incoraggiare la sempre maggiore diffusione di un approccio aperto al loro sviluppo, di ridurre i rischi di perdita di dati e di semplificare il loro recupero di fronte all'insorgere di problemi di obsolescenza dei software adottati o di loro abbandono da parte dei soggetti che ne curano lo sviluppo e la manutenzione, così come è capitato anche in tempi recenti, vedremo nel prossimo paragrafo il caso dell'archivio de «La Stampa».

Le entità prese in considerazione nei tracciati degli schemi, complessi archivistici, unità/sottounità archivistiche, unità documentarie, soggetti produttori, soggetti conservatori, strumenti di ricerca, entità/voci d'indice, sono quelle che tipicamente caratterizzano i sistemi descrittivi sviluppati nel nostro paese.

Prendiamo in esempio, come in altre circostanze, il nostro Sistema Archivistico Nazionale, SAN, per comprendere come sia costante il lavoro di analisi, modificazione e aggiornamento dei tracciati e dei sistemi di interazione tra di essi. Per l'interoperabilità tra EAD e EAC, SAN ha realizzato il tracciato CAT-SAN[343], che consente lo scambio di schemi XML e che costituisce ormai un punto di riferimento consolidato nell'assicurare l'interoperabilità fra i principali sistemi archivistici italiani e il Sistema Archivistico Nazionale.

Per favorire la più ampia compatibilità fra i vari sistemi e software, in molti casi sono stati predisposti sia elementi descrittivi altamente strutturati che elementi che permettano di fornire la medesima informazione con testi liberi. E' possibile utilizzare alternativamente l'una o l'altra tipologia di elementi descrittivi a seconda della maggiore o minore granularità dei dati del sistema di loro provenienza o destinazione.

343 Per approfondimenti rimando al sito dell'Istitituto Centrale per Gli Archivi https://www.icar.beniculturali.it/standard/standard-san/cat-san ultima consultazione 25/11/2021

Già da qualche tempo l'ICAR sta portando avanti alcune sperimentazioni di applicazione di tali tracciati alla esportazione e importazione di descrizioni archivistiche fra diversi applicativi. Essi sono stati ad esempio inclusi, con pieno successo, in Archimista[344], un pacchetto open source di interventi di manutenzione evolutiva del software di inventariazione.

I processi che ho descritto in queste pagine possono sembrare acquisiti per chi nel mondo dell'archivistica si sta confrontando ormai già da anni con gli sviluppi dell'informatica archivistica, ma risultano poco diffusi e discussi nel mondo degli storici e delle storiche, in pochi e poche infatti hanno mostrato interesse nel chiedersi come, all'interno di un sistema informazionale digitalizzato, le informazioni archivistiche vengono organizzate, come si manifestano a noi e in che modo sia possibile utilizzarle al meglio in una prospettiva di interoperabilità, così da poter immaginare nuovi processi di diffusione sociale delle fonti e dunque della storia.

Il concetto di rimediazione, da me ripreso da Grousin, è in sé un concetto molto potente che tende a descrivere il processo di modificazione di un contenuto che dal manifestarsi attraverso sistemi analogici arriva oggi a noi grazie a strumenti digitali, su queste poche parole è possibile comporre profonde analisi speculative che però non avrebbero nessun senso di esistere se non esistessero a monte procedure tecniche che consentono la trasformazione ontologica che tanto ci affascina.

Per dirlo con le parole di Popper

> Le argomentazioni metodologiche più feconde sono sempre ispirate da qualche problema pratico che sta di fronte allo studioso impegnato nella ricerca scientifica; e quasi tutte le argomentazioni metodologiche che non hanno una tale ispirazione sono caratterizzate da quella atmosfera di vana sofisticheria che ha tanto nuociuto alla reputazione della metodologia presso chi compie indagini pratiche[345].

A prescindere da ogni altra considerazione resta comunque il fatto che senza adeguati strumenti descrittivi gli archivi, o meglio le informazioni al loro interno, rimarrebbero conglomerati di informazioni inanimate e di difficile utilizzazione, qualunque sia l'uso che di queste informazioni si vorrà fare. Tutto ciò è ancora più vero se comprendiamo che questi processi sono finalizzati alla comunicazione e non solo alla catalogazione. Come ci ricorda sempre Valacchi, descrivere senza il fine di comunicare è un'operazione priva di senso. Ciò non di meno è necessario comprendere di cosa siano fatti gli archivi e quali siano i processi alchemici di trasformazione quando si impastano con i bit, sapendo che il futuro è già presente e che l'accesso ad inventari ontologici, per fare un

344 Per approfondimenti su Archimista si rimanda alla guida dell'ICAR elaborata da Costantino Landino nel 2018, *Archimista, nuove posisbilità per la descrizione e l'interoperabilità,* *https://www.icar.beniculturali.it/materiali/Presentazione_Archimista_20180919.pdf* ultima consultazione 25/11/2021

345 K.R. Popper, *Miseria dello storicismo*, Giangiacomo Feltrinelli Editore, Milano 2019, p. 70

esempio, sono non solo pensabili ma realizzabili grazie allo sviluppo dei Linked Opena Data e degli standard di metadatazione degli oggetti digitali.

Si tratta piuttosto di tarare sugli utenti quello che, con una certa compiaciuta retorica, si definisce il racconto degli archivi. In questo senso ciò che sembra necessario è una comunicazione archivistica tarata sulle tipologie di utenti reali e/o potenziali, capace di raggiungere bersagli precisi, ben identificati e che si ponga l'obiettivo di rompere il sostanziale accerchiamento che tanto penalizza gli archivi. Si potrebbe ancora riflettere sull'uso pubblico, attivo, degli archivi, nel tentativo di coglierne il ruolo di collante non solo culturale, ma di interi tessuti sociali. Gli archivi valutati quindi come patrimonio comune, funzionali ad esigenze civili, sociali e politiche, ancor prima che storiche e culturali hanno del resto un disperato bisogno di uscire dalle gabbie dei tecnicismi[346]

Queste parole offrono importanti spunti di riflessione per gli storici e le storiche, Valacchi infatti tocca due temi a mio avviso centrali rispetto a dei ragionamenti critici sulle pratiche di ricerca storica e di produzione storiografica. Ciò che Valacchi ci dice e che noi abbiamo potuto confermare nell'osservazione della composizione dei sistemi descrittivi di un archivio, è che le fonti, che sono appunto complessità narrative, non sono riposte in modo disordinato all'interno di un archivio, ma già nel loro ordinamento compongono un racconto. Fino ad oggi l'unico modo che si aveva per rendere pubblico questo racconto era smontandolo e ricomponendolo attraverso il lavoro degli storici e delle storiche, che interrogando e criticando le fonti e attraverso processi di narrazione e spiegazione, riescono ad elaborare contesti complessi, in cui i fatti, o meglio le informazioni relative ai fatti, vengono intrecciati in prima istanza attraverso un rapporto causale. Ma se l'archivio e i suoi enunciati divenissero di facile accesso per chiunque? E se i medium-documenti fossero collegati tra loro non solo tramite il metodo storico di ordinamento, ma anche tramite metadati che oltre a fornire informazioni descrittive della fonte, mettono gli oggetti digitali in relazione tra di loro anche tramite un rapporto tematico o meglio ancora concettuale? E se gli strumenti di comunicazione di un archivio fossero così sviluppati da rendere semplice l'interrogazione dell'archivio stesso anche a utenti non specializzati?
Per capire ciò di cui sto parlando è necessario meglio comprendere cosa siano i metadati e quali standard già oggi siano utilizzati per costruire delle articolazioni, che se coadiuvate dall'intervento di storici e storiche, potrebbero portare a mio avviso alla possibilità di costruire un nuovo tipo di storiografia, che non si sostituirebbe a quella già esistente, ma che consentirebbe di generare intorno

346 F. Valacchi, *L'applicazione degli standard agli strumenti archivistici nel web*, UniMc, 2020 https://docenti.unimc.it/federico.valacchi/teaching/2020/22027/files/standard-di-descrizione-archivistica, ultima consultazione 22/11/2021

alla fonte che si manifesta a noi come oggetto digitale, dei processi di spiegazione e narrazione basati sull'interoperabilità tra le fonti e l'utilizzo di metadati descrittivi.

Le mie domande sono frutto di intuizioni generate in prima istanza dal lavoro sul campo con fonti digitali (come abbiamo potuto vedere per esempio nel laboratorio svolto con l'Istituto Luce), ma spero che possano essere utili a sviluppare una riflessione che possa divenire collettiva, in relazione non solo alla possibilità di fruizione di fonti primarie nel Web, grazie a sistemi complessi di ordinamento e descrizione, ma anche in relazione al nuovo ruolo che storiche e storici potrebbero avere all'interno della rivoluzione digitale. Per meglio capire ciò di cui sto parando è però necessario andare ad osservare i metadati.

4 Metadati

Gli standard che abbiamo osservato, sono stati sviluppati al fine di costruire dei modelli condivisi per la strutturazione e per la comunicazione degli inventari d'archivio, dunque per organizzare e dare la possibilità di essere fruite tutte le informazioni disponili che si hanno in relazione alla fonte conservata. Grazie a questi standard e all'interoperabilità per i quali sono stati implementati è dunque possibile sapere quali fonti si trovino in quai archivi, quale sia la loro collocazione, le date di archiviazione e produzione, il formato, quali i vincoli archivistici, quale la storia archivistica della fonte e quali le relazioni con altre fonti. A corredo di queste e altre informazioni, vengo riportate informazioni di cotesto più generale e bibliografiche, come abbiamo potuto vedere per ISAAR/EAC, le informazioni riguardo il soggetto produttore arrivano a dettagliare anche le relazioni che l'ente o il soggetto ha avuto in vita, così da consentire di approfondire sempre di più la ricerca.

Sappiamo che attraverso una ricerca sul Web, anche utilizzando motori generalisti, è possibile trovare i link che ci rimandano a sistemi informativi, ad archivi digitali o a confederazione di archivi all'interno dei quali, prima ancora di osservare la fonte primaria, possiamo ottenere informazioni valide e utili per la ricerca o per la costruzione di un contesto di approfondimento. Standard, archivi digitali e Sistemi Informativi, sono quindi strumenti di organizzazione per la conservazione, ma anche di comunicazione delle informazioni, strumenti che nel caos della rete si mostrano come potenziali e potenti argini alla misinormation. Certo, perché i Sistemi Informativi possano davvero assumere un valore sociale diffuso e generalizzato, dobbiamo in prima istanza educare noi stessi all'utilizzo, così che gli algoritmi di Google ce li mostrino come risultati possibili già nella prima pagina di ricerca, non solo, se a scuola, ma anche all'università, si condividesse la conoscenza di questi strumenti e si desse indicazioni utili e semplici per usarli, questi sarebbero davvero strumenti formidabili e dal grande impatto sociale e culturale. E' impossibile pensare di poter controllare e disinnescare ogni informazione falsa o sbagliata che gira in internet, riguardante la storia, è forse più facile insegnare dove e come è possibile trovare informazioni affidabili.

Nelle pagine precedenti ci siamo concentrati sull'analisi di sistemi quali EAD e EAC, come rappresentazione in XML di ISAD e ISAAR, ma oltre a questi vi sono altri standard nati in seno alle diverse organizzazioni archivistiche nazionali[347] e non solo.

347 Abbiamo già avuto modo di osservare se pur brevemente CAT-SAN come modello di interscambio con EAD e progettato direttamente per il Sistema Archivistico Nazionale

Tra questi non posso non citare NIERA. Le norme italiane per l'elaborazione dei record di autorità archivistici di enti, persone, famiglie NIERA (EPF) sono l'esito di un lungo percorso di elaborazione, avviato nel quadro dello sviluppo del Sistema Archivistico Nazionale con l'istituzione con decreto 29 gennaio 2009, della Direzione generale per gli archivi, di una sottocommissione tecnica per l'elaborazione delle liste di autorità dei soggetti produttori e dei soggetti conservatori degli archivi presenti all'interno del Sistema Archivistico Nazionale. La sottocommissione ultimò i propri lavori nell'aprile del 2009 con la redazione del documento *Tracciato descrittivo dei record di autorità dei soggetti produttori dei complessi archivistici*. Con decreto del Direttore generale per gli archivi del 7 aprile 2010, fu successivamente istituita la Commissione nazionale per l'elaborazione del codice normativo per i soggetti produttori d'archivio, composta da archivisti di Stato, rappresentati delle Regioni, dell'Università, del mondo bibliotecario e delle istituzioni culturali.
Una prima versione delle NIERA (EPF) fu presentata nella conferenza nazionale *I poli archivistici e le reti informative*, tenutasi a Pescara dal 15 al 17 dicembre 2011. Nel 2012-2013 sono state illustrate e discusse in numerose sedi e sono state aggiornate sulla base delle osservazioni e dei contributi ricevuti e della sperimentazione nel quadro dei sistemi archivistici nazionali e locali. Nel luglio 2014 ne è stata rilasciata la seconda versione, pubblicata in «Rassegna degli Archivi di Stato»[348]

Le NIERA, basate sul modello rappresentato dall'International Archival Autority Records (Corporate bodies, Persons and Families), propongono alla comunità archivistica italiana uno standard per:

- la descrizione delle entità nel record di autorità archivistico e in qualsiasi altro ambito (sia cartaceo che informatico) e indipendentemente dalla natura della relazione che intercorra tra le entità e la documentazione (produzione, conservazione, ecc.);
- la composizione delle intestazioni di autorità normalizzate delle entità e la visualizzazione delle liste di autorità nei sistemi informativi archivistici
- la descrizione di collegamenti e relazioni tra entità e documentazione, entità e altre entità - entità e risorse di altro tipo sempre nel contesto di sistemi informativi archivistici.

Le NIERA forniscono regole e vocabolari controllati per la standardizzazione degli elementi strutturati del record di autorità, in modo che essi costituiscano chiavi di accesso normalizzate, e norme redazionali per la compilazione del record di autorità.

Come scrive Vitali

348 Rassegna degli archivi di stato, Nuova serie, VIII-IX, Roma 2013
https://www.icar.beniculturali.it/fileadmin/risorse/Niera_Rassegna_AS.pdf ultima consultazione 28/11/2022

Dalle Niera emerge chiaramente che una maggiore omogeneità delle descrizioni dei soggetti produttori, la loro comparabilità e l'interoperabilità semantica fra i vari sistemi descrittivi può essere conseguita non tanto grazie all'identica formulazione dell'intestazione d'autorità, quanto alla presenza nei sistemi dei medesimi elementi in grado di identificare univocamente le entità descritte (denominazioni, date, luoghi, professioni, contesto politico statuale ecc.) rappresentati secondo un'identica scala di dettaglio, o come si suol dire, con la medesima granularità.[349]

Le parole di Vitali ci aiutano ad avvicinarci ad un tema centrale per il nostro ragionamento.
La definizione più semplice di metadati è di «dati relativi ad altri dati»[350], se andiamo a prendere la definizione che da Dublin Core nel suo glossario del 2005 leggiamo:
«In general, "data about data;" functionally, "structured data about data." Metadata includes data associated with either an information system or an information object for purposes of description, administration, legal requirements, technical functionality, use and usage, and preservation.»[351]
Questa definizione è molto importante poiché non descrive i metadati solo come dei dati su altri dati, ma ci da indicazione di come questi abbiano il "potere" di essere classificati in base alla funzione che hanno, dandoci così un prezioso indizio: i metadati sono uno strumento necessario all'organizzazione e alla specificazione delle informazioni direttamente collegate alla fonte primaria o meglio all'oggetto digitale.
Utilizzare le parole oggetto digitale in riferimento alle fonti primarie di per sé ci aiuta ad ampliare il modo che fino ad oggi abbiamo sempre avuto di immaginare e comprendere una fonte storica.
Negli ultimi anni, grazie anche al lavoro svolto dalle associazioni nazionali e internazionali di Public History e ai diversi prodotti realizzati dai public historian, il concetto di fonte si è sempre più aperto e non afferisce più in modo univoco all'idea di documento di archivio come frequentemente siamo stati abituati in ambito storico. Ogni elemento afferente al passato può e viene osservato come una fonte da criticare e comparare ad altre, per riuscire ad elaborare un'analisi e una narrazione storiografica. Fotografie, filmati, tracce sonore, opere d'arte, testimonianze orali, prodotti di grafica, film, diari e autobiografie, unite a documenti amministrativi, comandi militari, leggi, minute delle sedute parlamentari ecc., fanno parte del complesso mondo informativo con cui storiche e storici si devono confrontare.
Un oggetto digitale, è ognuna delle fonti che abbiamo elencato, ma esiste solo se collegato a dei metadati.

349 S. Vitali, *La descrizione degli archivi nell'epoca degli standard e dei sistemi,* in, a cura di, L. Giuva, M. Gercio, *Archivistica. Teoria, normativa e prassi*, San Miniato, Archilab, 2002, p. 207
350 S. Vitali, *Pasato digitale. Le fonti dello storico nell'era del computer*, Bruno Mondadori, Milano 2004, p. 186
351 https://dublincore.org/specifications/dublin-core/usageguide/glossary/ ultima consultazione 3/12/2021

Come si legge nel Glossary of Digital Library Terms, della Californai Digital Library, la definizione di oggetto digitale è «An entity in which one or more content files and their corresponding metadata are united, physically and/or logically, through the use of a digital wrapper»[352]

Oltre a ciò dobbiamo tenere in considerazione che il concetto di digital object viene declinato in due modi: simple digitl object e complex digital object.

Simple digital object: «Comprised of a single content file (and its format variants or derivatives) and the metadata for that file. For example, a TIFF of the Mona Lisa, a user JPEG, a reference GIF, and the appropriate metadata would comprise a simple digital object.»[353]

Complex digital object: «Includes two or more content files (and their format variants or derivatives) and corresponding metadata. The content files are related as parts of a whole and are sequenced logically, such as pages. For example, a complex digital object could consist of a multi-page diary scanned as TIFF images, from which are generated display images (JPEGs and GIFs), plus a transcription of the diary and the metadata for each file.»[354]

La prima cosa che notiamo da queste definizioni è che non esiste oggetto digitale senza metadati, da qui comprendiamo in modo chiara l'importanza dello strumento metadato e la conseguente importanza per storici e storiche di comprendere cosa siano.

La definizione di complex digitl object porta alla nostra attenzione la possibilità, sempre grazie ai metadati, di generare un oggetto digitale che è rappresentazione di sé, ma non solo. Dunque un oggetto digitale complesso non si limita ad essere la rappresentazione simulacrale della fonte materica con in aggiunta uno schema di metadati, ma è qualcosa in più. Pensiamo al progetto sul Codice Pelavicino o al progetto riguardante i Rotoli del Mar Morto di cui abbiamo parlato nei capitoli precedenti, l'oggetto digitale non si limita ad essere la riproduzione del documento, ma anche una sua edizione critica. Senza metadati tutto ciò non potrebbe essere, o meglio non potrebbe essere fruito e ordinato.

La necessaria presenza di metadati nella produzione di oggetti digitali, da storici e storiche non può che farci porre alcune riflessioni sul fatto che una fonte da materica a digitalizzata, oltre a subire un processo di ri-mediazione - che di per sé comporta l'utilizzo di un nuovo linguaggio per la rappresentazione della fonte stessa – esiste, in quanto oggetto digitale, solo grazie ad un processo di sovrascrittura; sono infatti i metadati ad assicurarne una validità come oggetto digitale e a renderne possibile la fruizione.

La fonte materica, dal che viene digitalizzata al che si rende disponibile al fruitore in un ambiente digitale, vive un processo prima di riscrittura e poi di sovrascrittura.

352 https://cdlib.org/resources/technologists/glossary-of-digital-library-terms/#D ultima consultazione 10/12/2021
353 ibidem
354 ibidem

E’ dunque sbagliato pensare che le fonti storiche vivendo questo processo si modifichino non solo nella loro natura ontologica, ma anche in quella epistemologica?

Più avanti proveremo ad affrontare questo discorso, prima è necessario confrontarci ancora con l'oggetto metadato e con gli standard di metadatazione.

Di per sé il termine non indica nulla di specificamente nuovo, poiché metadati possono essere tranquillamente definite le schede di un catalogo, qualunque sia la normativa seguita o la tecnologia applicata. Nuovo è semmai il fatto che essi entrino a far parte della struttura stessa del documento e siano finalizzati a rendere possibile, idealmente, la catalogazione “automatica” del documento al momento della sua indicizzazione da parte di strumenti di ricerca appositamente predisposti per interpretarli.

La storia del termine metadato è di per sé molto particolare, fu infatti Jack E. Myeres che nel 1973 utilizzò il termine metadati con intenti dichiaratamente pubblicitari. Nel 1986 Myeres registrò il termine metadati come marchio di fabbrica della società di cui era presidente la Metadata Copany, e quando da li a poco la Federal Geographic Data Committee incominciò ad usare il termine nella sua accezione attuale, Myere aprì un contenzioso sostenendo che la Metadata Company avesse diritto all’esclusiva, dal momento che la società si era costituita per sviluppare prodotti basati sui concetti di metamodello di descrizione dell’insieme di dati[355].

La richiesta di esclusiva di Myere non fù accolta e oggi è proprio sui metadati che si sta sviluppando l’ultima visione del Web, quello che potremmo chiamare Web 3.0.

DESIRE, Development of a European Service for Information on Research and Education, un progetto di sviluppo cooperativo, volto a promuovere l'uso del World Wide Web nella comunità di ricerca europea, che fa parte del settore "Telematica per la ricerca" del programma specifico di RST della Comunità nel campo delle applicazioni telematiche e coordinato da SURFnet bv, - fornitore di reti accademiche nei Paesi Bassi, che conta 23 partner in otto paesi[356] - da come definizione di metadati «dati relativi ad oggetti che consentono agli utenti di conoscere l’esistenza di una fonte informativa e le sue caratteristiche». Questa definizione sposta la nostra attenzione sul valore dei metadati non solo per la descrizione atta ad assicurare la corretta esistenza del documento digitale, ma su lato della conoscenza. Uno strumento sviluppato pensando ad un possibile pubblico che vuole fruire informazioni tramite il Web, specificando anche che i metadati «sono dati associati ad

355 K. Hodgson, *Metadata foundations, potential and applicationes,* Edmonton: School of Library and Infomation Studies, University of Alberta, 1998

356 https://cordis.europa.eu/article/id/7486-desire-development-of-a-european-service-for-information-on-research-and-education ultima consultazione 10/12/2021

oggetti che esentano il potenziate fruitore dal dover disporre preventivamente di una conoscenza completa dell'esistenza o delle caratteristiche degli oggetti»[357].

Secondo DESIRE i metadati fanno sì che il possibile fruitore sia esentato dall'avere conoscenza preventiva delle caratteristiche degli oggetti digitali. Questo implica che i metadati all'interno di un sistema come il Web, se ben organizzati si possono mostrare come degli strumenti atti al facile reperimento delle informazioni anche per chi non ha conoscenze relative alle fonti che sta cercando, ciò implica che nel processo di metadatazione si debba tener conto della possibilità di risemantizzare la fonte al fine di renderla semanticamente accessibile anche da parte di chi non ha conoscenze specifiche, come vedremo questo processo è possibile implementarlo se la definizione dei metadati avviene in relazione ad una comunità di riferimento.

Tim Berners-Lee definisce i metadati come informazione relativa a risorse Web comprensibile dalla macchina[358], sottolineando la novità dei metadati rispetto alle precedenti strutture referenziali. I metadati vengono inseriti nell'intestazione dell'oggetto digitale, tramite diversi insiemi di etichette e di campi di codifica per rendere possibile il trattamento automatico da parte di sistemi di elaborazione e di tutti i soggetti che, a diverso titolo e con finalità differenti, utilizzeranno quella risorsa elettronica. Non appena l'oggetto digitale nasce, dunque, è già automaticamente catalogato, vi è quindi un processo di automatizzazione delle procedure di indicizzazione.

E' indubbio che i metadati si differenzino e non poco dalle consuete pratiche di indicizzazione e catalogazione.

I metadati, se da un lato sono finalizzati a rendere automatizzata la pratica di catalogazione della risorsa al momento della sua indicizzazione, da parte di strumenti di ricerca appositamente predisposti; dall'altro lato forniscono notizie riguardanti la struttura, l'autenticità, la disponibilità, le modalità di accesso e di riuso e la storia della risorsa stessa.

Oltre a ciò, Paul Weston ci da indicazione chiara del fatto che i metadati siano anche molto di più.

Affinchè i metadati possano assolvere alle diverse funzioni, è indispensabile che siano correttamente applicati dal punto di vista della completezza e della correttezza sintattica e semantica. Solo a tale condizione essi consentono l'identificazione univoca, l'individuazione e la localizzazione di una risorsa, l'aggregazione e l'organizzazione di risorse con caratteristiche comuni e la gestione dei processi di archiviazione e conservazione che supportano e descrivono.[359]

357 M. Day, *Metadata. Mapping between metadata formats*. Uklon: The UK Office For Library and Information Networking, *http://www.ukoln.ac.uk/metadata/*, ultima consultazione 10/12/2021

358 T. Berners-Lee, *Metadata Architecture: documents, metadata and links, axiomof Web architecture 3*. Cambridge, Mass.:Mit, 1997, https://www.w3.org/DesignIssues/old/Metadata.html ultima consultazione 10/12/2021

359 P.G. Weston, L. Sardo, *Metadati*, Associazione Italiana Biblioteche, Milano 2017

Pur rimanendo legati ad un'analisi di tipo tecnico vorrei tornare alla domanda che ci siamo posti poco fa, ossia se sia corretto pensare che i metadati implichino la necessità di una nuova critica epistemologica al concetto di fonte storica. Rileggiamo alcune parole di Weston a mio avviso particolarmente significative: i metadati richiedono una «correttezza sintattica e semantica», per assicurare un'identificazione univoca della fonte, una sua individuazione e la sua aggregazione con altre fonti. Ciò che avviene dunque è che la fonte oltre alle informazioni che già porta con se, per esistere come oggetto digitale ha bisogno di altre informazioni che noi le attribuiamo, e queste devono essere corrette non solo a livello sintattico, ma anche semantico. Vi è dunque la possibilità che i dati collegati all'oggetto digitale assommino dei significati alle informazioni presenti nella fonti e che questi significati, ma in senso metasintattico, possano essere letti e utilizzati dalle macchine al fine di mettere quella fonte in correlazione con altre fonti alle quali è stato attribuito anche lo stesso significato. E' in questo passaggio, che nelle prossime pagine approfondiremo anche da un punto di vista tecnico, che si intravede la possibilità di generare all'interno di sistemi di archiviazione digitale quel processo che richiamerebbe alla spiegazione, che prima, nell'analisi degli standard come EAD e EAC pareva mancasse.

Se alle fonti si attribuiscono significati capaci di creare connessioni con altre fonti il discorso storiografico arriva a comporsi di uno degli elementi che maggiormente lo contraddistingue.

Seguendo questo ragionamento è impossibile non portare le nostre riflessioni a confrontarsi con le idee del Web Semantico – che grazie all'utilizzo di triple è realmente in grado di dare competenza semantica alle macchine -, ma per ora è meglio rimanere ancorati all'analisi dei metadati e delle procedure necessarie al loro funzionamento.

Il rapporto tra metadati e fonte digitale o digitalizzata non si sviluppa sempre nello stesso modo, con questo intendo che l'applicazione dei metadati alla fonte, al fine di poterla far diventare un oggetto digitale, può avvenire nel momento della realizzazione della fonte stessa o successivamente. A seconda della procedura in base alla quale vengono creati i metadati, possono avere differenti collocazioni ed essere collegati alla risorsa elettronica che rappresentano secondo più modalità. Possono essere incapsulati nella risorsa, collocati all'interno del documento elettronico nella sua parte iniziale, che non viene visualizzata dal lettore, detta intestazione head o Headher (ripesiamo ad uno dei campi che abbiamo appena analizzato in EAC-CPF). Esempi di questa tipologia di collocazione sono le intestazioni di Dublin Core in un file HTML o XML.

Questa è la procedura più ricorrente ogni qualvolta la creazione dei metadati sia contestuale alla produzione del documento, cioè sia operata dal soggetto produttore stesso. Oppure possono essere collocati esternamente alla risorsa, ossia far parte di un archivio elettronico distinto da quello che

ospita il documento che i metadati descrivono, con la predisposizione di un opportuno sistema di puntamento fra i metadati e le rispettive risorse.
Un esempio di questa tipologia di collocazione è il formato MARC, con il quale vengono prodotte notizie catalografiche conformi a standard di descrizione e mediante l'applicazione di norme che ne definiscono gli elementi e ne prescrivono le fonti. La locazione della risorsa, le sue caratteristiche tecniche e la modalità di accesso vengono indicate nel campo 856, che funge da dispositivo per il collegamento diretto alla risorsa[360].
La compilazione dei metadati in questo caso è conseguente all'intervento di un catalogatore umano o artificiale, che descrive la risorsa elettronica in un momento successivo alla sua creazione.
I metadati possono essere collocati esternamente alla risorsa, ma essere forniti in risposta a una specifica richiesta di informazione. In questo caso i metadati vengono prodotti e gestiti all'interno di un archivio elettronico separato dalla risorsa. Questa possibilità riporta dei vantaggi per l'aggiornamento. Ripensiamo ai funzionamenti di Google e ai processi di estrapolazione di parole dai contenuti selezionati dagli spider. In quel caso le parole selezionate per indicizzare la risorsa e per rendere possibile il ritrovamento al fruitore, vengono collezionate in file esterni, che attraverso un sistema di puntamento sono collegati con la risorsa che si vuole indicizzare.
Infine i metadati possono costituire la struttura all'interno della quale i dati vengono inseriti, questo avviene come abbiamo visto in EAD che non è uno standard per il contenuto dei dati, né fornisce indicazione su come esprimerli, ma è uno schema sviluppato per codificare e strutturare gli strumenti di ricerca di raccolte archivistiche in livelli gerarchici, all'interno dei quali i metadati sostituiscono la risorsa. Ovviamente in questa modalità la creazione dei metadati coincide con la procedura di creazione della risorsa.
Le informazioni che costituiscono i metadati sono scandite in campi predefiniti, o elementi, che rappresentano gli attributi, ciascuno dei quali può assumere, di volta in volta, uno o più valori.
Uno schema di metadati presenta le seguenti caratteristiche:

- un numero limitato di elementi
- il nome di ciascun elemento
- il significato di ciascun elemento
- indicazioni circa l'obbligatorietà o meno di ciascun elemento
- istruzioni e o esempi d'uso

La gestione dei metadati richiede un'attenzione e una costanza maggiore rispetto a come eravamo abituati a gestire i documenti e le informazioni in un ambiente analogico.

360 https://www.loc.gov/marc/ ultima consultazione 15/12/2021

Il documento digitale viene infatti, sottoposto a continui interventi di ampliamento, modifica, rielaborazione e aggiornamento dei dati da parte di soggetti differenti, nei vari momenti della propria esistenza e in ragione delle diverse funzioni svolte. Le informazioni che lo descrivono devono essere ospitate in una struttura flessibile e modulare, in cui vengano registrati, nella successione delle procedure, elementi ricavati di preferenza direttamente dal documento stesso.[361]

Generalmente gli elementi della struttura dei metadati fanno riferimento alle caratteristiche fisiche, alla tipologia e al formato della risorsa descritta, a cui vanno aggiunti i punti di accesso ai documenti, oltre al titolo l'autore o il creatore e le coordinate spazio temporali.
Nel corso degli anni numerose comunità di utilizzatori hanno messo appunto schemi di metadati atti a rappresentare nel modo più efficace le risorse di loro interesse. Dobbiamo considerare che accanto a questi metadati, che possiamo definire descrittivi, ce ne sono altri che rispondono all'esigenza di assicurare la preservazione dei documenti elettronici e altri ancora che rendono possibile l'esecuzione di procedure di tipo amministrativo gestionale.
Ciò che maggiormente interessa alla nostra ricerca è comprendere quali reali funzioni abbiano o possano avere schemi di metadati descrittivi soprattutto in relazione allo spunto che ci ha offerto Waston nella citazione precedente, in cui viene detto che nelle strutture è possibile individuare elementi di riferimento presi direttamente dalla fonte.
L'integrazione della descrizione con la riproduzione digitale del documento rappresenta un cambiamento di non poco conto nel modo di intendere la funzione svolta dallo strumento di ricerca. Se tradizionalmente per fare ricerca si utilizzava strumenti di tipo referenziale, aventi come primo obiettivo quello di consentire la localizzazione dei singoli oggetti o delle diverse collezioni, nel momento in cui si rende possibile l'accesso diretto al documento, o al suo simulacro digitale, lo strumento di ricerca diventa fonte di informazione anche primaria e, di conseguenza, uno strumento di tipo ibrido, il cui obiettivo va oltre la semplice localizzazione dei documenti.
Nel 2003 all'interno del report sulla conservazione digitale del seminario *Metadata and digital preservation* i metadati vengono così definiti:
«Metadata are for documents or other information resources as water is to humane beings. Lack of water will lead to dehidration and malfunctioning»[362]
L'esigenza di allestire una struttura - innanzi tutto logica, ma finalizzata al trattamento elettronico dei dati - all'interno della quale raccogliere, descrivere, conservare e rendere disponibile il patrimonio culturale sia nella forma originaria, sia in quella digitale, integrando le due versioni

361 P.G. Weston, L. Sardo, *Metadati*, Associazione Italiana Biblioteche, Milano 2017, p. 10
362 Erpanet Traing Seminar, *Metadata in digital preservation – final report*, Marburg3-5- settembre, 2003, p. 7 https://www.yumpu.com/en/document/read/30354260/metadata-in-digital-preservation-erpanet ultima consultazione, 15/12/2021

all’interno dei nuovi meccanismi indicizzatori, ha condotto gli specialisti di diversi ambiti ad occuparsi dei metadati.
Grazie ai linguaggi che consentono la realizzazione di strutture come EAD, con la conseguente integrazione di metadati descrittivi direttamente collegati all'oggetto digitale, macchine e esseri umani si trovano ad interagire in una costante interpretazione biunivoca.
Secondo Gorman «la funzione dei metadati non è tanto quella di descrivere la risorsa, quanto quella di garantire una qualche forma di continuità nel passaggio della risorsa da una procedura o da un ambiente all'altro»[363]
Possiamo dire che i metadati abbiano quindi un valore di collegamento tra il processo mentale di chi svolge la ricerca e le risorse necessarie alla ricerca. Questo collegamento avviene unicamente attraverso il linguaggio, o meglio i linguaggi di cui è composta questa complessa rete informazionale. Non è solo l’utente che deve modificare i suoi meccanismi mentali per doversi confrontare con una ricerca attraverso il Web, ma sono i metadati stessi che devono riuscire ad accogliere e “comprendere” i meccanismi mentali di chi fa la ricerca.

L’utente che avvia la ricerca in un sistema informativo, la formula in base a uno schema mentale determinato da fattori quali le sue conoscenze linguistiche, la familiarità con la disciplina, l’esperienza accumulata in precedenti ricerche, la conoscenza diretta di uno o più documenti desiderati, nonché dalle finalità della ricerca stessa. Le corrispondenze fra lo schema mentale dell'utente e quello applicato dalla risorsa e le loro interrelazioni sono spesso difficili da identificare e da rendere esplicite.
Un primo compito dei metadati è, quindi, quello di consentire l’individuazione dei documenti mediante la loro rappresentazione.[364]

Torniamo per un attimo ad immaginarci all’interno di quell'archivio senza pareti di cui nei capitoli precedenti abbiamo provato a definire i Livelli di Astrazione, dove i processi di conoscenza telepistemica possono avvenire attraversato una ricerca per parole chiave così da riuscire ad individuare gli oggetti digitali e a portarli nel nostro spazio di lavoro.
Le parole chiave, come ci fa capire Waston, non sono altro che la rappresentazione sintattica e semantica dei processi mentali che sottostanno all’azione di ricerca e che devono trovare una corrispondenza nei metadati descrittivi legati all’oggetto digitale. Non solo, in un’ottica di interoperabilità, gli standard di metadatazione consentono all’utente che compie la ricerca, di utilizzare specifiche parole per individuare non una sola fonte, ma un set di fonti che rispondono

363 M. Gorman, *Metadata or Cataloging? A false Choice*, «Journalof Internet Catloging». 2/1, pp. 5-22
364 P.G. Weston, L. Sardo, *Metadati*, Associazione Italiana Biblioteche, Milano 2017, p. 11

alla chiave di ricerca; una descrizione formalizzata consente all'utente di stabilire se accedere a tutte o ad un parte delle fonti e delle risorse individuate.

«I metadati consentono di connotare i diversi elementi della descrizione e di circoscrivere quindi l'ambito della ricerca a quelle risorse che presentino una caratteristica comune, come la ricorrenza di un determinato termine nel titolo, il nome del creatore della risorsa o l'argomento.»[365]

Le relazioni tra risorse informative e metadati, dunque, appaiono anche solo ad uno sguardo veloce tanto cruciali quanto particolarmente articolate e dinamiche. Come abbiamo visto i metadati possono essere statici o dinamici, possono essere *embedded* cioè inclusi come parte integrante della risorsa, oppure ad essa collegati sulla base di meccanismi identificativi stabili e sicuri nel tempo; diversi schemi di metadati con diverse funzioni possono essere associati allo stesso oggetto o più oggetti possono essere associati tra loro tramite metadati, e infine sappiamo che durante il ciclo di vita di una risorsa è necessario che il corpus di metadati sia aggiornato di continuo, per documentare tutti gli eventi ad essa occorsi.

Non si pensi dunque che la digitalizzazione conduca ad una semplificazione nella gestione delle informazioni, soprattutto se si parla di fonti che necessitano un ordinamento all'interno di un archivio, sia questo corrente, di deposito o storico.

4.1 Standard

Per riuscire a costruire "infrastrutture" capaci di gestire questo rapporto costante tra dati e metadati negli anni sono nati diversi progetti anche molto simili tra loro e con finalità praticamente identiche, alcuni di questi modelli e schemi con il tempo sono diventati degli standard per l'utilizzo e l'applicazione dei metadati. Se avere standard condivisi di metadatazione è fondamentale per riuscire a costruire un processo di omologazione all'interno di quello che è un grande sistema complesso, dall'altra parte il proliferare degli standard ha chiaramente avuto un risultato opposto.

Questo ci deve far chiedere cosa si intenda propriamente per standard quando ci si riferisce al variegato panorama di modelli, profili applicativi, formati, linguaggi, protocolli, e in genere degli schemi di riferimento più o meno applicativi delle scienze dell'informazione. Da che momento in poi, e tramite quali meccanismi un modello tecnico opportunamente documentato diviene uno standard? Pur esistendo degli organismi di certificazione degli standard riconosciuti

365 Ivi, p 12

internazionalmente e a diversi livelli, come W3C, ISO, NISO, UNI (che peraltro non adottano tutti la stessa politica per la diffusione pubblica della documentazione), questi non sono certo gli unici a rilasciarne. Si pensi, ad esempio, agli standard prodotti dalla Library of Congress statunitense, un'istituzione talmente meritoria e autorevole a livello globale da costituire un punto di riferimento indiscutibile nel panorama delle scienze dell'informazione.
Vi sono poi i formati rilasciati da associazioni tecniche e imprenditoriali, come ad esempio il Moving Pictures Expert Group, meglio conosciuto come MPEG[366], che si occupa della definizione dei requisiti tecnici fondamentali degli oggetti digitali (digital item, nell'ottica di questo standard) e della possibilità di interazione da parte dell'utente con i medesimi oggetti.

Possiamo dunque sostenere che perché si possa parlare di uno standard, per i metadati, è necessario che una comunità sufficientemente ampia e autorevole accolga un modello documentato e lo sperimenti e applichi, evidenziandone eccellenze e criticità, così da correggerlo e mantenerlo aggiornato rispetto a eventuali nuove esigenze, questo processo, soprattutto in una fase di costante sperimentazione tecnologica, rende il concetto di standard qualcosa di un po' più liquido di quello che ci potremmo normalmente immaginare.[367]

Sapendo, come ci ricorda Guercio,[368] della complessità che sottende alla descrizione dei diversi livelli di applicazione dei metadati, può essere utile elencare sinteticamente alcuni tra i molti punti critici. Prima di tutto, dato che la conservazione a lungo termine degli archivi deve garantire non solo l'accesso, ma anche l'identità e l'integrità dei documenti e delle loro relazioni, è necessario adottare, accanto a modelli e schemi gestionali di livello generale, schemi dettagliati e differenziati per le singole tipologie documentarie e per i contesti applicativi, con la massima attenzione a documentare e conservare, oltre che le risorse documentarie, tutto ciò che documenti le scelte tecnico-organizzative e gli schemi adottati.
Prima di procedere nell'osservazione dei diversi schemi di metadati che ad oggi sono disponili come standard per l'identificazione degli oggetti digitali, dobbiamo far emergere alcune problematiche che benché non inficino il valore intrinseco di organizzazione della conservazione e di ricerca offerta dai metadati, ci sono utili per comprendere come l'utilizzo di standard di metadati sia ben lungi da essere un processo chiuso, ma anzi, come abbiamo già accennato è necessario osservarlo come processo in divenire, dal momento che è legato alle modificazioni logiche e

366 https://mpeg.chiariglione.org/ ultima consultazione 18/10/2021
367 P. Feliciati, *Gestione e conservazione di dati e metadati per gli archivi: quali standard?*, Unimc Pubblicazioni Aperte Digitali, 2010 http://eprints.rclis.org/13657/1/atti-testoFeliciati09.pdf ultima consultazione 18/10/2021
368 M. Guercio, G. Marinelli, *La conservazione delle memorie digitali: un quadro di riferimento per un progetto applicativo*, in «Archivi & Computer», a. XVII, fasc. 1/07, 2007 pp. 30-58

tecniche che con il passare del tempo sempre coinvolgeranno i sistemi informativi e l'organizzazione in generale del Web.

Sull'identificazione sicura e stabile degli oggetti digitali in ambiente di rete, per esempio, dobbiamo in prima istanza sottolineare che per via delle caratteristiche dell'attuale Internet governance sono oggi in vigore contemporaneamente troppi standard "globali", con l'effetto di vanificare l'efficacia universale delle funzioni identificative, indispensabili perché il Web, nella sua versione più avanzata di sistema dei sistemi informativi, sia davvero ambiente di interoperabilità negli ambiti scientifici, culturali e anche dell'amministrazione[369], ricordiamoci il caso del link presente su SIUSA che non era più in grado di portarci sul sito del MART.

Altra questione spinosa è comprendere la tipologia funzionale dei metadati, sempre meno facile da distinguere.

Tra le tassonomie, quella adottata più di frequente e che qui si proverà a seguire è quella di Wendler[370], che distingue i metadati in tre categorie funzionali:

- descrittivi: per l'identificazione ed il recupero degli oggetti digitali; sono costituiti da descrizioni normalizzate dei documenti fonte (o dei documenti digitali nativi), risiedono generalmente nelle basi dati dei sistemi di Information Retrieval all'esterno degli archivi degli oggetti digitali, e sono collegati a questi ultimi tramite appositi link;
- amministrativi e gestionali: per le svariate operazioni di gestione degli oggetti digitali all'interno dell'archivio, comprendono di norma anche i metadati tecnici, relativi agli oggetti digitali nella loro natura più strettamente tecnica come formati, mezzi e procedure di creazione, elaborazione, post elaborazioni;
- strutturali: per descrivere la struttura interna dei documenti come introduzione, capitoli, indice di un libro, e gestire le relazioni fra le varie parti componenti degli oggetti digitali.

Di questi, i primi, sono i più simili agli schemi su cui si basano le classiche attività di descrizione, inventariazione o catalogazione riguardanti oggetti analogici, i secondi somigliano agli elementi gestionali che pure sono presenti in alcune tipologie di descrizioni dell'analogico come i riferimenti topografici, i numeri di catalogo, i riferimenti alla responsabilità della descrizione, etc., infine i terzi sono separati dai primi solo per le specifiche proprietà dell'ambiente digitale, mentre nel mondo del tangibile la struttura o è parte integrante della descrizione, ad esempio per gli archivi o per gli spogli delle riviste, o non è affatto, ad esempio nell'articolazione in capitoli o pagine di un oggetto bibliografico, che non è oggetto di catalogazione.

369 M. Sebastiani, *Identificatori persistenti per gli oggetti digitali*, «DigItalia», 2005, pp. 62-84
370 https://www.iccu.sbn.it/export/sites/iccu/documenti/manuale.html ultima consultazione 15/12/2021

La logica dei metadati, tiene conto dunque, del fatto che le procedure necessarie per creare e poi aggiornare i documenti digitali implicano una continuità maggiore rispetto alle procedure messe in atto per documenti che applicano tecnologie analogiche. Il documento digitale è costituito da un nucleo iniziale di dati che viene ampliato, modificato, rielaborato, aggiornato da entità differenti, nei vari momenti della sua esistenza ed in ragione delle diverse funzioni svolte. Le informazioni che lo descrivono devono, quindi, rientrare in una struttura flessibile e modulare, in cui vengano organizzati, nella successione delle procedure, elementi ricavati di preferenza direttamente dal documento.

Più tipologie di informazioni concorreranno a descrivere un documento: quelle di tipo semantico o formale ne renderanno possibile la ricerca e l'individuazione; quelle di natura tecnologica ne consentiranno il recupero, il trasferimento e la rielaborazione; quelle amministrative ne faciliteranno la gestione. Le fonti con i metadati si stratificano e divengo sempre più dinamiche.

In relazione alla tassonomia espressa da Wendler andiamo adesso ad osservare gli standard di metadatazione più utilizzati con particolare interesse verso gli standard di metadati descrittivi.

Come riporta Pigliapoco in *Progetto archivio Digitale* e con lui concordano Feliciati, e Waston, nonostante la necessaria tara che come abbiamo visto è necessario fare sul concetto di standard applicato ai metadati, questi son gli schemi più utilizzati:

- Dublin Core Metadati Initiative (DCMI)
- Metadata Encoding and Transimission Standard (METS)
- Metadati Amministrativi e Gestionali (MAG)
- Preservation Metadata: Implementation Strategies (PREMIS)

Per quanto MAG sia uno standard sviluppato in Italia e METS sia lo standard sul quale ha elaborato un proprio linguaggio il SAN con il suo standard METS-SAN, DCM è lo standard che trovo maggiormente interessante, fosse anche solo a livello speculativo. Lo ritengo lo standard che apre maggiori riflessioni in ambito storiografico e a mio avviso indica la strada per soluzioni davvero innovative. Invero questa non è solo la mia opinione, DCMI è infatti anche un vocabolario utilizzabile all'interno del Web semantico e come vedremo la sua flessibilità ne consente un utilizzo davvero molto diffuso. Per questo lascerò l'analisi di Dublin Core alla fine, avendo così maggior libertà argomentativa e in prima battuta descriverò gli altri standard, a partire da METS.

4.1.1 METS

Metadata encoding and transmission standard, nasce nel 2001 su iniziativa della Digital Library Federation[371], ed è supportato dalla Library of Congress. Si propone come un formato di documenti XML - formato che come abbiamo già potuto vedere nell'osservazione di EAD e EAC viene utilizzato da praticamente tutti gli standard - utilizzabile per codificare i metadati descrittivi, amministrativi e gestionali degli oggetti digitali conservati in un archivio digitale e per lo scambio di contenuti tra diversi depositi e utenti. Anche METS come altri standard, consente di realizzare dei pacchetti informativi poi applicabili al modello OAIS, un modello concettuale per la realizzazione e gestione dei sistemi informativi[372], modello di cui approfondiremo in seguito il funzionamento.

METS è costituito da un unico schema in grado di ospitare elementi provenienti da altri schemi integrandoli attraverso due meccanismi di inclusione: l'incapsulamento degli elementi di metadati <mdWrap> nel file METS o il riferimento esterno ai file che li contengono <mdRef>.

Un documento METS è costituito da sette sezioni principali

1. Intestazione METS <metsHdr>

   Permettere di memorizzare i metadati descrittivi minimi su l'oggetto all'interno di un documento METS. Fra questi la data di creazione del documento METS, la data della sua ultima modifica e lo status del documento METS. In questa sezione si possono memorizzare le entità che hanno avuto un ruolo rispetto al documento e identificatori alternativi del documento
2. Metadati descrittivi <dmdSec>

   contiene i metadati descrittivi relativi all'oggetto METS nel suo insieme o in una delle parti che lo costituiscono, possono essere inseriti anche in documento METS esterni e collegati con un link
3. Metadati amministrativi <amdSec>

   contengono i metadati amministrativi relativi all'oggetto digitale, alle sue componenti e al materiali originale dal quale l'oggetto digitale deriva. La sezione è divisa in quattro sottosezioni: i metadati tecnici riguardanti la creazione, il formato e le caratteristiche di utilizzo, i metadati sulla proprietà intellettuale come copyright e licenze d'uso, metadati sull'eventuale provenienza analogica e i metadati sulla provenienza digitale
4. Sezione file <fileSec>

371 https://www.diglib.org/ ultima consultazione 15/12/2021
372 S. Pigliapoco, *Progetto archivio digitale. Metodologia, sistemi professionalità*, Civita Editoriale, Lucca, 2018 p. 71

contiene gli oggetti digitali o i link ad essi collegati, contiene uno o più elementi <fileGrp> che consente di raggruppare i file correlati.

5. Mappa strutturale <structMap>
   Collega gli elementi di questa struttura ai contenuti della sezione file è l'unica sezione obbligatoria, definisce la struttura gerarchica degli oggetti digitali da presentare all'utente, in modo da permettere la consultazione
6. Link strutturali <structLink>
   contiene un singolo elemento <smLink> che consente di memorizzare la presenza di hyperlink tra i diversi componenti delineati nella mappa strutturale
7. Sezione comportamento <behaviorSec>
   utilizzata per associar uno o più codici eseguibili al documento METS

L'importanza dei metadati tecnici e amministrativi è esplicitata perfettamente nell'introduzione all'utilizzo di METS, più ci si inoltra nell'osservare le strutture degli schemi e più ci è chiaro come e quanto i metadati rappresentino un cambiamento davvero profondo nelle possibilità di archiviazione e ricerca, gli oggetti digitali senza metadati rimarrebbero un insieme di bit in un universo in espansione, i metadati rappresentano la fonte, la rendono possibile da trovare e ne aumentano il portato informativo. Quale dunque il reale valore epistemologico di questi dati sui dati, quanta la nostra dipendenza da questa sovrascrittura, possiamo noi in quanto storici non sviluppare un'analisi e una critica verso gli strumenti che oggi, ma soprattutto domani consentiranno lo svolgere del nostro lavoro?

Curare una biblioteca di oggetti digitali significa gestirne, necessariamente, i relativi metadati. I metadati necessari, ad una gestione e ad un uso efficace degli oggetti digitali, sono diversi ed anche più numerosi, rispetto a quelli usati per manutenere collezioni di lavori a stampa o di altri oggetti fisici. Se una biblioteca può conservare i metadati descrittivi di un libro, il libro non si dissolverà in una serie di pagine sconnesse, se la biblioteca dovesse incorrere in qualche errore nel memorizzare i metadati strutturali su come è organizzato il libro, né gli studiosi saranno impossibilitati nel valutare il libro se la biblioteca dovesse sbagliare nell'annotare che il libro e' stato prodotto usando una stampa offset Ryobi. Lo stesso non può dirsi per la versione digitale dello stesso libro. Senza i metadati strutturali, le immagini delle pagine o i file di testo di cui e' costituito un lavoro digitale sono di scarsa utilità, e senza i metadati tecnici riguardanti il processo di digitalizzazione, gli studiosi potrebbero essere insicuri su quanta rispondenza ci sia nella versione digitale rispetto all'originale. Inoltre, per necessita' di gestione interna, una biblioteca deve avere accesso a metadati

tecnici appropriati per rinnovare e migrare periodicamente i dati, assicurando che la risorsa abbia un valore durevole[373].

Questo passaggio è davvero centrale nel dare valore alla nostra ricerca e al nostro ragionamento. Senza metadati il lavoro di storiche e storici non potrebbe avvenire in un ambiente digitale e se questo per alcuni oggi può sembrare un problema relativo, sapendo che molta della documentazione del passato è in formato materico, cosa dire però, per chi farà il nostro lavoro tra cinquanta anni?
METS come anticipato è lo standard utilizzato dal Sistema Archivistico Nazionale, METS-SAN, per creare dei file capaci di codificare tutte le informazioni presenti all'interno dell'archivio e consentendo una interoperabilità non solo tra gli oggetti digitali ma tra gli standard stessi.

Ogni record METS in SAN avrà obbligatoriamente una sezione descrittiva in Dublin Core, finalizzata a garantire la massima compatibilità con altri sistemi, sia in importazione che in esportazione. Per la descrizione degli oggetti archivistici digitalizzati si prevede inoltre una descrizione specifica, che utilizza elementi di EAD. Ove presente, alcuni elementi di tale descrizione EAD-like, saranno ripetuti anche nella sezione Dublin Core[374].

4.1.2 MAG

Rimanendo sull'analisi di un progetto italiano, MAG, standard di Metadati Amministrativi e Gestionali, nasce per la gestione di fonti originariamente materiche, anche se può essere utilizzato per la gestione di fonti native digitali. MAG è un application profile[375] che ha l'obiettivo di fornire le specifiche formali per la fase di raccolta, di trasferimento e disseminazione dei metadati e dei dati digitali nei rispettivi archivi. Lo Schema MAG è realizzato e mantenuto dal Comitato Mag.
L'istituto Centrale per il Catalogo Unico, ICCU, tra le numerose attività relative all'utilizzo dei metadati, ha costituito nel 2000, un Gruppo di studio sugli standard e le applicazioni di metadati nei beni culturali a cui hanno partecipato rappresentanti di biblioteche, musei e archivi. Il Gruppo è nato con l'obiettivo di coordinare a livello nazionale le implementazioni di metadati nei progetti di digitalizzazione nei diversi settori dei beni culturali e di raccordare le iniziative italiane con quelle europee e internazionali.
Il Gruppo di studio, che ha operato nel periodo 2000 - 2002, si è suddiviso al suo interno in due sottogruppi, uno relativo ai metadati descrittivi, l'altro ai metadati amministrativi e gestionali.

373 https://www.loc.gov/standards/mets/METSita.html ultima consultazione 18/12/2021
374 http://ims.dei.unipd.it/data/san/met/documents/2010-05-20_oggetti-digitali-san20100425-2.pdf ultima consultazione 15/12/2021
375 R. Heery, M. Patel., *Application profiles: mixing and matching metadata schemas*, http://www.ariadne.ac.uk/issue/25/app-profiles/ ultima consultazione 15/12/2021

Da luglio 2003 si è formalmente costituito un gruppo di lavoro permanente, il Comitato Mag, che prosegue le attività del Gruppo di studio sugli standard e le applicazioni di metadati nei beni culturali con particolare riferimento alle attività connesse alla diffusione ed evoluzione del set di Metadati Amministrativi Gestionali (MAG). L'ultima versione dello standard risale al 2006.

MAG integra diversi elementi provenienti da vari schemi di metadatazine internazionali come Dublin Core o METS, e ha come finalità quella di rendere «più affidabile e sicura l'informazione digitale documentando le modalità di generazione, immissione, archiviazione e manutenzione degli oggetti digitali nel sistema di gestione dell'archivio digitale»[376]

Lo schema MAG fornisce specifiche formali per le fasi di raccolta e archiviazione dei dati e fornisce elementi per

- identificare in maniera univoca gli oggetti digitali
- certificare l'autenticità e l'integrità dei contenuti
- documentare la catena di custodia degli oggetti digitali
- documentare i processi tecnici eseguiti per la conservazione permanente degli oggetti digitali
- informare sulle condizioni e i diritti di accesso agli oggetti digitali da parte degli utenti finali

Lo schema risulta composto di diverse sezioni, utilizzabili a seconda del contenuto digitale e dell'impiego dello stesso:

- GEN: informazioni generali sul progetto e sul tipo di digitalizzazione
- BIB: metadati descrittivi dell'oggetto analogico
- STRU: metadati strutturali dell'oggetto digitale
- IMG: metadati specifici per le immagini fisse
- OCR: metadati specifici relativi al riconoscimento ottico del testo
- DOC: metadati specifici per oggetti digitali in formato testo che possono essere derivati o born digital
- AUDIO: metadati specifici per file audio
- VIDEO: metadati specifici per file video
- DIS: metadati specifici per la distribuzione di oggetti digitali.

In questo schema, per quanto riportato in modo sintetico, è possibile rintracciare i processi tecnici da cui dipende la realizzazione di alcuni esempi che abbiamo portato nei capitoli precedenti. I

376 MAG, *manuale utente*, https://www.iccu.sbn.it/export/sites/iccu/documenti/manuale.html#-b3

campi OCR, DOC, AUDIO e VIDEO, ci mostrano in modo inequivocabile l'adattamento dei metadati alle diverse tipologie di fonti e alle diverse modalità che si hanno per fruirle.
Quando si è parlato di oggetti digitali complessi si è fatto riferimento alla possibilità, grazie alla tecnologie OCR, di fare una ricerca ottica all'interno di un documento digitalizzato che si manifesta a noi come un PDF. Queste tecnologie ci consentono non solo di osservare il documento nella sua rappresentazione simulacrale, ma di navigarne il contenuto attraverso una ricerca testuale, andando così a creare un unione tra le due modalità di digitalizzazione di cui abbiamo parlato in precedenza, una relativa alla trascrizione con metalinguaggi di marcatura del testo e una relativa alla scansione digitale del documento che trasforma questo in un medium-documento. Nel primo caso, come abbiamo visto, è possibile sviluppare percorsi di ricerca anche di natura linguistico computazionale, ma apre la strada a possibili errori di trascrizione e ad una più semplice modificazione della fonte digitale, mentre la scansione del documento e la sua trasformazione in un medium-documento, ci assicurano una perfetta riproduzione della fonte ma ne limitano, in teoria, le possibilità di analisi testuale. Con i sistemi di Optical Character Recognition, è possibile ottenere l'affidabilità della riproduzione per scansione e l'approfondimento attraverso l'analisi testuale.

La sezione OCR raccoglie i metadati amministrativi e gestionali relativi a file di testo ottenuti mediante riconoscimento ottico automatico del contenuto. I file classificabili all'interno di tale sezione si differenziano da quelli descritti dalla sezione DOC perché non presuppongono alcun intervento correttorio manuale e quindi qualsiasi intervento editoriale è assente.
La sezione OCR utilizza il namespace niso: che fa riferimento allo schema che traduce le linee guida del Data Dictionary NISO. Anche se lo standard NISO è stato concepito per la descrizione delle caratteristiche tecniche delle immagini digitali, lo schema niso-mag.xsd include anche specifiche tecniche relative a documenti text-oriented (riconoscibili grazie al prefisso doc)[377]

La stratificazione di dati consente di realizzare progetti incredibilmente innovativi, ma ancor più interessante è che nonostante la complessità di questa stratificazione da un punto di vita tecnico e logico, per il fruitore tutto avviene in modo così lineare da non accorgersi minimamente della complessità che rende tutto ciò possibile.
Per questo è necessario andare a scoprire meglio anche gli ultimi standard di metadati; prima di arrivare a Dublin Core, che come abbiamo potuto vedere è lo standard con il quale sia MAG che METS dialogano, andiamo ad approfondire PREMIS.

377 Comitato MAG, *Manuale utente*, 2006, https://www.iccu.sbn.it/export/sites/iccu/documenti/manuale.html#2 ultima consultazione 18/12/2021

4.1.3 PREMIS

Preservation Metadata Implementation Strategies è uno standard di metadati di conservazione ed è stato elaborato dal gruppo di lavoro internazionale Preservation Metadata Imlementation Strategies Working Group costituito nel 2003 dal research library group dell'OCLC.

PREMIS si mostra come uno strumento differente da quelli che abbiamo osservato fino ad ora, questo infatti non definisce tutti i metadati presi in considerazione, ma solo quelli di maggior uso che un deposito digitale deve conoscere; pertanto può essere visto come un insieme di elementi che racchiude al proprio interno porzioni di metadati descrittivi, informazioni sulle regole di amministrazione del deposito digitale, dati tecnici, indicazioni su enti e diritti. La maggior parte dei dati PREMIS può essere valorizzata automaticamente dal sistema di conservazione, ma alcuni di essi devono essere specificati manualmente. Tra questi figurano le informazioni relative all'accessibilità e all'uso di contenuto digitale, alla provenienza digitale intesa come conoscenza dell'ambiente di produzione e delle attività di gestione che hanno interessato l'oggetto prima del trasferimento nel deposito, e alle caratteristiche significative dell'oggetto digitale che devono essere mantenute nel tempo[378].

PREMIS si mostra come uno standard per l'integrazione degli altri standard all'interno di un deposito digitale.

Il dizionario di dati PREMIS definisce unità semantiche ciascuna delle quali «is mapped to an entity that is organized within a simple data model. A semantic unit therefore can be understood as a property than an entity»[379]

Le funzioni di conservazione possono essere diverse a seconda del tipo di sistema di deposito, ma generalmente includono azioni per assicurare che gli oggetti digitali rimangano accessibili per i media di supporto e possano essere visualizzati, eseguiti o altrimenti interpretati dalle applicazioni software, come anche assicurare che gli oggetti digitali nel deposito non vengano inavvertitamente alterati, e che i cambiamenti legittimi agli oggetti vengano opportunamente documentati.

Il Dizionario dei Dati non è stato concepito per definire tutti gli elementi di metadati di conservazione possibili, ma solo per definire quelli che, il più delle volte, la maggior parte dei depositi avrà bisogno di sapere.

Diverse categorie di metadati sono escluse in quanto ritenute fuori obiettivo:

- Metadati per formati specifici, ad esempio i metadati che riguardano solo un formato di file o di formati classi come audio, video o di grafica vettoriale.

378 S. Pigliapoco, *Progetto archivio digitale. Metodologia, sistemi professionalità*, Civita Editoriale, Lucca, 2018 p. 73
379 https://www.loc.gov/standards/premis/ ultima consultazione 23/12/2021

- Metadati per implementazioni e regole business specifiche, come per esempio i metadati che descrivono le politiche o le pratiche di un deposito su come esso fornisce accesso ai materiali.
- Metadati descrittivi. Sebbene la descrizione delle risorse è ovviamente di pertinenza della conservazione, molti standard indipendenti possono essere usati a tale scopo, come il MARC, il MODS e il Dublin Core.
- Informazioni dettagliate su media o su hardware. Di nuovo, sebbene chiaramente di pertinenza della conservazione, la definizione di tali metadati vengono lasciati alla competenza di altre comunità.
- Informazioni sugli agenti (persone, organizzazioni o software), tranne che per il minimo necessario all'identificazione.
- Informazioni aggiuntive sui diritti e sui permessi; il focus è incentrato su quelli che hanno un impatto diretto sulle funzioni di conservazione.

Se si pensa che ogni metadato sia utile all'organizzazione che gestisce un deposito digitale, PREMIS può essere visto come definizione di un sottoinsieme che sta nel centro. Non è coinvolto nella ricerca e nell'accesso, e non definisce metadati dettagliati per formati specifici. Esso definisce solo i metadati comunemente necessari ad eseguire funzioni di conservazione su tutti i materiali.
PREMIS dunque a differenza degli standard finora descritti non è stato sviluppato per la costruzione di uno schema di metadati, ma per il mantenimento di questi. Da qui emerge chiaramente l'importanza di questo standard senza il quale si rischia di muoversi verso un futuro digitale in cui la conservazione a medio e lungo termine non ha il supporto adatto ad esistere e di conseguenza il lavoro di digitalizzazione e metadatazione rischia di perdersi come pulviscolo nella rete se non di scomparire proprio a causa dell'obsolescenza.
PREMIS consente ai sistemi di archiviazione digitale di immaginarsi una vita a lungo periodo.
In seguito a ciò trovo davvero molto interessante che nel manuale del 2021 si tenda a sottolineare come PREMIS non lavori su metadati, ma su unità semantiche.

Il Dizionario dei Dati PREMIS definisce unità semantiche e non elementi di metadati. La distinzione è sottile ma importante. Un'unità semantica è un pezzo di informazione o di conoscenza. Un elemento di metadati è un modo definito di rappresentare quelle informazioni in un record di metadati, in uno schema o in un database. PREMIS non specifica come i metadati dovrebbero essere rappresentati in qualsivoglia sistema, definisce soltanto cosa è necessario sapere per poterli esportare su altri sistemi.[380]

380 Priscilla Caplan, *Capire PREMIS*, 2021, p. 5
https://www.loc.gov/standards/premis/understandingPREMIS_italian_2021.pdf ultima consultazione 23/12/2021

L’attenzione di PREMIS si focalizza, pertanto, sul sistema informativo, sul deposito e sulla gestione, e non sugli autori dei contenuti digitali, sulle attività di digitalizzazione o sul personale che si occupa delle licenze, nella selezione degli elementi di metadati non vengono inclusi tutti gli elementi esistenti, ma soltanto quelli funzionali ala conservazione[381].

Uno dei principi più importanti sottesi all’iniziativa PREMIS è l’esigenza di avere un’evidenza immediata di cosa viene descritto. PREMIS definisce cinque tipi di cose, chiamate Entità, che possono essere descritte: Entità Intellettuali, Oggetti, Agenti, Eventi e Diritti.

Gli Oggetti sono ciò che realmente viene archiviato e gestito nei depositi di conservazione e buona parte di PREMIS è dedicata a descrivere oggetti digitali.

L’entità Evento, aggrega informazioni sulle azioni che riguardano gli Oggetti del deposito. Un record di eventi accurato ed affidabile è necessario per mantenere la provenienza digitale di un oggetto, che è importante per dimostrare l’autenticità dell’oggetto stesso. Interessante sottolineare come in questo caso con il termine evento non si definisca l’evento dal quale prende forma la fonte, ma gli eventi a cui è sottoposta la fonte stessa una volta all’interno del processo di digitalizzazione.

Gli Agenti sono attori che hanno dei ruoli negli Eventi e nelle dichiarazioni sui Diritti, e negli Oggetti Ambiente. Gli Agenti possono essere persone, organizzazioni o applicazioni software. PREMIS definisce solo un numero minimo di unità semantiche necessarie a identificare gli Agenti, dal momento che ci sono diversi standard esterni che possono essere usati per memorizzare informazioni più dettagliate. Un deposito potrebbe scegliere di usare uno standard a parte per memorizzare informazioni addizionali sugli Agenti, o potrebbe usare l’identificatore dell’Agente per puntare alle informazioni memorizzate esternamente.

Anche sul tema dell'Agente è interessante proporre un breve ragionamento, in questo caso non è l’agente come lo intende Dray, ossia l’individuo che da solo o in relazione con altri individui pensa e agisce in modo da generare l’evento, ma è l’Agente in quanto archivista 2.0 (mi si consenta la libertà espressiva). Ciò ci deve rimandare al ragionamento svolto nel secondo capitolo, dove si è posto l’attenzione sul ruolo dell’archivista e del suo apporto nella costruzione del contenuto dell’archivio, dal momento che la selezione delle fonti da un archivio di deposito ad uno storico è in carico ad agenti/archivisti che con le loro azioni contribuiscono a costruire quell’ordito che abbiamo detto essere la prima rappresentazione di una narrazione storica basta sulle fonti.

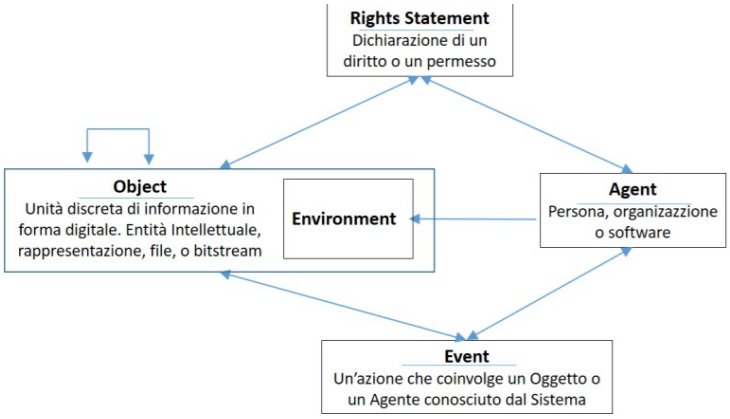
La maggior parte delle strategie di conservazione comporta l’esecuzione di copie identiche e di versioni di derivazione degli oggetti digitali, azioni che per la legge sul copyright sono permesse solo ai proprietari dei diritti. L’Entità Diritti aggrega informazioni sui diritti e permessi che sono direttamente pertinenti agli oggetti nel deposito di conservazione, in modo tale che il deposito possa

381 P.G. Weston, L. Sardo, *Metadati*, Associazione Italiana Biblioteche, Milano 2017, p. 97

fare ciò che necessita, per conservarli. Ogni dichiarazione sui Diritti asserisce due cose: le azioni che il deposito ha diritto ad intraprendere e in base a quale diritto può farle.

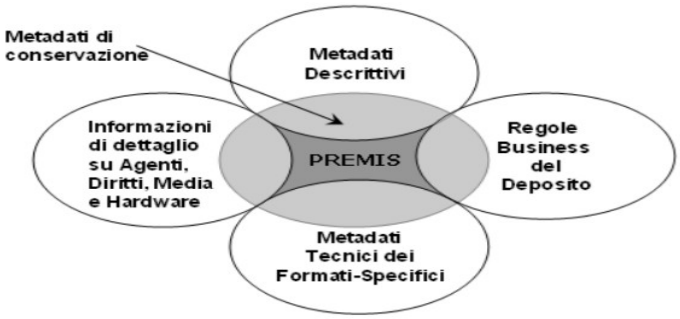
L'elemento diritti, è di chiaro interesse in una prospettiva di riuso delle fonti d'archivio, sempre più spesso, soprattutto in un'ottica di Public Digital History, vi è l'intenzione di utilizzare e riusare materiali d'archivio nella realizzazione di progetti e prodotti. Ma non sono solo i public historian a porsi questa domanda, non di meno specifici archivi stanno provando a condividere i loro contenuti con un pubblico sempre più ampio, cercando con l'esterno un'interazione attiva. Un esempio lampante è il progetto PH-Remix, realizzato dal Laboratorio di Cultura Digitale dell'Università di Pisa, la Fondazione Sistema Toscana (FST) e il Festival dei Popoli, un progetto che si è posto l'obiettivo di sviluppare soluzioni innovative per la fruizione e il riutilizzo da parte del pubblico di archivi audiovisivi tramite l'estrazione semiautomatica di contenuti, una catalogazione che tenga in considerazione i contenuti estratti e l'accesso facilitato in una piattaforma che consenta all'utenza di "rimescolarli" per creare nuovi materiali. Tutto ciò però può avvenire solo nel rispetto dei diritti vigenti[382].

PREMIS è uno standard fondamentale nel processo di digitalizzazione che sta investendo il mondo degli archivi in quanto standard per la conservazione a lungo termine delle fonti digitali, ci fa riflettere sui processi che garantiscano la corretta procedura del funzionamento di un deposito. Le due immagini che possiamo vedere di seguito ci fanno riflettere su due diversi livelli di PREMIS, la prima è una immagine schematizzata della relazione in cui stanno le diverse Entità descritte, è dunque uno schema che ci consente di comprendere il funzionamento interno di PREMIS. Interessante notare che l'Entità Oggetto è contraddistinta da una freccia che punta verso se stessa, poiché è previsto che un Oggetto possa essere in relazione con altri oggetti, mentre nessuna delle altre Entità può essere in relazione con un'altra Entità del medesimo tipo.



382 http://www.labcd.unipi.it/ph-remix/ ultima consultazione 23/12/2021

La seconda immagine invece ci mostra l'importanza di PREMIS come standard di gestione di metadati di differente natura, come possiamo vedere il cerchio che rappresenta PREMIS include alcuni metadati descrittivi, alcune regole business, alcuni metadati tecnici di dettaglio e alcune informazioni dettagliate su agenti, diritti, media e hardware, PREMIS non è coinvolto nella ricerca e nell'accesso, e dall'altra non definisce metadati dettagliati per formati specifici. Esso definisce solo i metadati comunemente necessari ad eseguire funzioni di conservazione su tutti i materiali.



4.1.4 Dublin Core Metadata Initiative

L'ultimo standard di cui voglio parlare, è quello che ritengo maggiormente interessante e sul quale torneremo anche in relazione all'utilizzo del sistema di gestione OAIS e alla possibilità di costruire ontologie digitali grazie a vocabolari comuni, sto parlando di Dublin Core Metadata Initiative (DCMI)

Il Progetto DCMI è stato sviluppato nell'ambito dell'Online Computer Library Center, OCLC, con l'obiettivo di creare un insieme di strumenti condivisi per l'accesso alle risorse digitali. La prima versione, rilasciata nel 1996, era composta da 15 elementi di base, ciascuno dei quali descritto con attributi in conformità con la norma ISO[383].

Il progetto nasce in seguito alla necessità di escogitare uno schema descrittivo molto agile, nel quale i dati richiesti per la compilazione della descrizione possano variare a seconda delle necessità e che possano essere utilizzati direttamente da coloro che creano le risorse digitali, pur senza essere in possesso di specifiche competenze catalografiche. È con questi presupposti che nel 1995 si tenne il primo workshop, ospitato presso il quartier generale dell'OCLC, a Dublin, Ohio, città dalla quale lo schema Dublin Core deriva il proprio nome. Dal momento che lo sviluppo di uno schema risponde alle esigenze funzionali della comunità di utilizzatori che lo elabora, gli schemi già esistenti hanno

383 S. Pigliapoco, *Progetto archivio digitale. Metodologia, sistemi professionalità*, Civita Editoriale, Lucca, 2018 p. 70

in genere poco in comune, al di là del fatto di essere stati ideati per il trattamento e la descrizione di risorse informative elettroniche. Ciò che contraddistingue lo schema conosciuto come Dublin Core Metadata Element Set è l'assenza intenzionale di una specifica destinazione d'uso. Mentre, infatti, gli altri metadati sono stati pensati in funzione di utilizzazioni ben definite e per tipologie di materiali chiaramente individuate, DCMI può essere utilizzato per descrivere risorse di qualunque natura.

I principi da cui nasce DCMI non possono che affascinarci, specificatamente se public historian. La prima cosa sulla quale dobbiamo porre la nostra attenzione è il fatto che lo standard nasca dal basso in seguito ad un workshop che ha coinvolto una collettività che non rispondeva ad un unico interesse ma cercava al contrario di farne emergere di diversi. Se noi consideriamo i metadati come elementi di sovrascrittura delle fonti, capaci di attribuire a queste dei significati estrapolati con attenzione dal loro contenuto o dal loro contesto, immaginare di avere a che fare con uno standard così agile e semplice da usare, e così diffuso da renderlo punto di riferimento per altri standard, ci può far pensare che davvero la digitalizzazione delle fonti e l'ordinamento di archivi digitali si possano mostrare come un processo capace di proporre innovative modalità di intendere la storia e la storiografia e chiaramente anche il modo che si ha di intendere il lavoro della storica.

Per favorire la diffusione di Dublin Core e renderne possibile l'adozione come standard *de facto*, il gruppo promotore ha stabilito di definire un numero minimo di elementi dal significato immediatamente comprensibile e di dar loro una struttura talmente flessibile da consentire la descrizione di risorse di qualunque natura e relative a un qualsiasi ambito disciplinare. Gli elementi di cui è composto lo standard sono stati definiti tenendo presenti essenzialmente le esigenze della ricerca e delle individuazione delle risorse. Di conseguenza l'insieme degli elementi che costruiscono lo schema Dublin Core è il risultato di un consenso ampio e interdisciplinare sui dati che sono necessari per l'effettuazione della ricerca di una risorsa. La domanda che continuo a pormi è quale potrebbe essere l'apporto degli storici nell'implementazione di questo standard?

Nell'elaborare lo schema i promotori si sono attenuti ai seguenti criteri:

- L'intrinsecalità, cioè il fatto di descrivere le proprietà intrinseche della risorsa, come il contenuto intellettuale o la forma fisica, piuttosto che quelle estrinseche, che descrivono il contesto nel quale la risorsa viene utilizzata
- L'estensibilità, cioè la possibilità di introdurre un limitato numero di personalizzazioni per adeguare la descrizione a esigenze specifiche
- L'indipendenza dalla sintassi, per non vincolare gli utilizzatori al rispetto di definizioni formali

- Le opzionalità degli elementi, per consentire agli utilizzatori di redigere descrizioni di qualunque estensione, anche minima
- La ripetibilità degli elementi, per non imporre limiti al numero di occorrenze del medesimo elemento: autori, date, soggetti e così via
- La modificabilità degli elementi, cioè la possibilità di introdurre dei qualificatori per rendere più specifico il campo di applicazione di un determinato elemento

I principi ai quali si ispira la formulazione del nuovo standard sono rimasti nel tempo gli stessi, anche se il ruolo che esso ha assunto nei confronti delle risorse digitali e degli strumenti del web è significativamente mutato rispetto agli obiettivi individuati nei primi tempi

- Diffusione attraverso il consenso: DCMI ha inteso profilarsi come uno standard *de facto*, raccogliendo le sollecitazioni provenienti da ambienti disciplinari, linguistici e geografici eterogenei all'interno di un dibattito che prevede consultazioni pubbliche e costanti procedure revisionali e offrendosi come uno strumento gratuito
- Partecipazione internazionale: benché le prime riflessioni si siano svolte sotto gli auspici di OCLC, l'obiettivo è stato quello di presentare da subito lo strumento come un'iniziativa internazionale. Paesi sempre diversi hanno ospitato convegni e seminari, che si sono tenuti periodicamente, e anche negli organismi che coordinano le attività istituzionali si è fatto in modo che le diverse aree regionali fossero adeguatamente rappresentate. Ampio spazio ha ricevuto la tempestiva traduzione dei documenti più significativi .
- Neutralità degli scopi: nell'evoluzione dello standard si è adottato un approccio agnostico rispetto alle finalità d'uso dei metadati e alla natura pubblica o privata dei soggetti che li utilizzano, incoraggiandone l'adozione anche nell'ambito di standard *de iure* come è avvenuto nel caso della biblioteca digitale italiana, perché non in conflitto con la politica dell'accesso aperto .
- Neutralità nei confronti della tecnologia: la principale finalità di Dublin Core è fornire uno strumento a carattere semantico, in grado di chiarire la natura e la tipologia delle informazioni riguardanti la risorsa digitale. Poiché l'infrastruttura tecnologica sulla quale si appoggiano è destinata a mutare nel corso del tempo, vi è l'impegno a consentire che la codifica dello standard evolva di conseguenza per iniziativa delle istituzioni e dei settori che ne avvertono l'opportunità.
- Approccio transdisciplinare: fin dalle prime discussioni negli anni ‘90 l'interesse sulla realizzazione di uno strumento che consentisse di mettere in relazione tra loro risorse create in ambiti disciplinari diversi contribuendo alla loro interoperabilità nel web.

Leggendo i principi generativi di DCMI emergono diversi spunti e questioni molto utili alla nostra riflessione, ma per completezza di informazione e per poter avere un quadro più chiaro, è utile andare ad osservare anche le caratteristiche che hanno contribuito alla sua popolarità:

- La semplicità: lo schema è utilizzabile da chiunque senza necessità di una formazione specifica
- La comprensibilità: i descrittori utilizzati sono sostanzialmente tratti dal linguaggio naturale, a differenza di quelli utilizzati nell'ambito di schemi definiti da comunità specialistiche di utenti, e nessuna lingua è predefinita per la loro presentazione, per cui essi sono stati ampliamene tradotti
- L'interoperabilità semantica: lo schema può essere utilizzato per descrivere risorse di qualunque natura appartenenti a qualsiasi disciplina
- La flessibilità: non sono stabiliti a priori criteri riguardanti l'obbligatorietà e la ripetibilità di ciascun elemento e non è necessario ricorrere a vocabolari controllati predeterminati[384]

Gli elementi concettuali e i propositi di progettazione di DCMI, per quanto sicuramente datati visto che l'elaborazione dello schema risale al 2000, credo però ci debbano far porre alcune domande su quale possa essere o sarebbe potuto essere, il ruolo di storici e storiche nell'elaborazione di questo schema. Essendo uno standard di metadati descrittivi, più di altri standard DCMI ha e può avere un ruolo determinate nei processi di sovrascrittura delle fonti. DCMI è stato pensato per agire direttamente sull'oggetto digitale che ha una sua identificazione univoca all'interno del "caos" del Web[385], la flessibilità di DCMI consente di utilizzare questo schema per qualsiasi risorsa digitale così da consentire un'interoperabilità semantica che potenzialmente rende possibile un'aggregazione di diverse fonti attraverso l'utilizzo di semplici parole chiave. Non solo, essendo più semplice da gestire rispetto ad altri standard richiede un livello di specializzazione inferiore nella codifica dei metadati e nell'ampliamento dei set descrittivi, ricordiamo infatti che DCMI consente di utilizzare anche solo una parte dei sui quindici livelli descrittivi e ancora più interessante consente di attribuirne di nuovi.

È in relazione a questa apertura e modificabilità che credo ci si possa chiedere se i principi di DCMI non ci debbano far riflettere sul lavoro dello storico e delle storiche e su quale possa essere il loro ruolo oggi all'interno di una società in cui ogni tipo di informazione viene cercata e trovata prevalentemente nel Web grazie a diversi motori di ricerca.

384 P.G. Weston, L. Sardo, *Metadati*, Associazione Italiana Biblioteche, Milano 2017, p. 26
385 T. Berners-Lee, *Network Working Group*, 1998, https://www.ietf.org/rfc/rfc2396.txt ultima consultazione 2/1/2022

Se condividiamo l'idea che il lavoro dello storico si possa non limitare alla produzione di storiografia, ma che nella sua accezione più sociale possa essere veicolo per diffondere il metodo di ricerca e analisi delle fonti, credo che sia lecito chiedersi cosa possa avvenire se lo storico, o la storica, intervenissero già nelle fasi di descrizione della fonte.

La nuova accessibilità alle fonti primarie che si ha grazie al Web, a mio avviso può modificare in modo sensibile il rapporto tra storia e storiografia, queste due accezioni trovano infatti un nuovo punto di congiunzione. La storia, di cui l'unica reale rappresentazione sono le fonti primarie, non emerge più solo grazie al lavoro delle storiche, ma anche grazie allo sviluppo di articolati processi di ordinamento, sovrascrittura e comunicazione delle fonti digitali.

Un lavoro storiografico per prendere forma ha bisogno oltre che dell'analisi critica e comparativa delle fonti, anche della spiegazione e della narrazione che come ci spiega Topolski è inevitabilmente guidata anche dal *savoir* dello storico o della storica[386]. Ma se noi immaginiamo che storici e storiche in seguito all'analisi delle fonti ne estrapolassero dei concetti, da trasformare in metadati, questi se pur in forma limitata potrebbero avere il ruolo di spiegazione, mentre l'interoperabilità delle fonti assumerebbe il ruolo di narrazione.

La narrazione storiografica però, non avverrebbe attraverso un intreccio per come lo intende White[387], ma attraverso la spiegazione che è avvenuta grazie ad una concettualizzazione. Il fruitore non sarà guidato da un registro narrativo che lo storico ha ritenuto adeguato per la narrazione di specifici eventi, ma dalla concettualizzazione di questi. Sempre facendo riferimento a White, per provare a spiegare meglio ciò che sto provando a dire, se si deve affrontare un argomento come un genocidio, lo storico o la storica saranno portati ad utilizzare un registro narrativo nobile, come l'epica o la tragedia[388], in questo caso non ci sarà questa necessità perché saranno le fonti stesse, come complessità narrative, a mostrarsi al fruitore, che a sua volta potrà generare un proprio intreccio narrativo in base alla reazione cognitiva e emotiva che avrà trovandosi difronte alla fonte primaria.

La conoscenza storica si manifesterebbe al fruitore attraverso un nuova tipologia di intermediazione, che tiene al centro della narrazione la fonte stessa, che a sua volta avrà un ruolo differente; non più solo solido appiglio scientifico per la costruzione storiografica da parte di un autore, ma vero e proprio elemento narrativo che si manifesta attraverso dei collegamenti, caricati di significato da chi in una fase antecedente a quella della fruizione ha composto la metadatazione.

Ciò che sto immaginando presuppone che sarà l'utente stesso a decidere l'ordine e la sequenzialità con cui navigare le fonti presenti nell'archivio o nella confederazione di archivi. Questo ci riporta ai

386 Jerzy Topolski, *Narrare la storia. Nuovi principi di metodologia storica*, Bruno Mondadori, Milano, 1997 p. 112
387 H. White, *Forme di storia. Dalla realtà ala narrazione*, Carocci Editore, Roma 2006, p. 88
388 Ivi, p. 91

funzionamenti della realtà immersiva, in cui i contenuti vengono mostrati in modo particellare e non ordinati in un processo sequenziale. Un esempio concreto di questi processi di fruizione degli oggetti digitali è dato da Thinglink[389], un' app che consente di realizzare ambienti di apprendimento in realtà immersiva dove è il fruitore, inserito sensorialmente nell'ambiente attraverso un Oculos o un meno sofisticato visore VR, che decide come muoversi all'interno dello spazio virtuale, quali informazioni fruire e in quale ordine.

Questa ipotesi apre a due problematiche non indifferenti, la prima, di ordine più strettamente metodologico, riguarderebbe un modificazione al paradigma che da sempre si ha nei processi di analisi delle fonti e di costruzione del processo storiografico. Muovendosi tra le fonti attraverso un legame semantico/concettuale che si sviluppa tra queste, si andrebbe ad osservare un rapporto non più solo di ordine causale, ma anche tematico. Non più quindi solo causa e effetto ma anche relazioni tematiche come filo conduttore della spiegazione.

La seconda questione riguarda un tema non solo metodologico, ma anche etico sociale. Come estrapolare i concetti? quali parole chiave usare? Come sapere che le parole scelte hanno effettiva rilevanza per i fruitori che si avventurano in un archivio digitale? Questa domanda credo possa trovare una risposta, o quanto meno una proposta, nelle pratiche della Public History, se si osserva questo metodo non solo per la diffusione del contenuto ma anche - o soprattutto mi verrebbe da dire - per la costruzione dal basso di analisi e contenuto storiografico. Questa riflessione trova una ricaduta nell'ambito pratico dal momento in cui, come vedremo, DCMI è uno standard adatto a strutturare un oggetto informativo con l'attribuzione di metadati in relazione ad una comunità di riferimento.

Immaginiamo dunque di essere delle storiche o degli storici presenti al primo working group che ha definito la struttura dello schema di DCMI, quali proposte avremmo potuto portare? Personalmente se fossi stato lì avrei proposto di immaginare la creazione di un campo relativo a dei metadati che definirei concettuali. Per formulare la proposta avrei argomentato che dal momento in cui i metadati descrittivi possono riportare argomenti presenti nelle fonti primarie, in seguito ad un attenta analisi delle fonti da parte delle storiche e degli storici, si potrebbe immaginare che il dato da riportare possa non essere solo di contenuto, ma di concetto.

Rispetto all'idea di usare i concetti per spiegare la storia, credo che ci possa essere utile questo passaggio di Koselleck

Che un'elucidazione storica dei concetti via via non debba risalire solo alla storia della lingua, ma anche a dati della storia e della società, è ovvio, poiché ogni semantica, in quanto tale, ha a che fare con contenuti

389 https://www.thinglink.com/ ultima consultazione 5/1/2022

extra linguistici. Questa è la ragione della posizione precaria e marginale che la semantica occupa tra le scienze linguistiche, ma spiega anche perché essa aiuti egregiamente la storia. Attraversando il setaccio della delucidazione concettuale, gli enunciati del passato vengono precisati, le situazioni e i rapporti intesi nel passato diventano più chiari ai nostri occhi nella loro versione linguistica[390].

Oltre a ciò i concetti consentono, o meglio semplificano, un'operazione che già Carr definiva fondamentale nella comprensione della storia, ossia di risemantizzare nell'oggi ciò che appartiene al passato e viceversa.

Per meglio comprendere cosa intendo mi rifarò nuovamente all'esperienza didattica sviluppata in collaborazione con l'Istituto Luce Cinecittà in un liceo classico di Roma[391].

Con gli studenti dovevamo realizzare una mostra multimediale che indagasse il fascismo, per farlo ho creduto fosse metodologicamente più coretto non procedere ripetendo gli stilemi di una didattica top down, dunque indicando io i temi da affrontare, ma ho cercato di aprire in classe una discussione che portasse studenti e studentesse a ragionare sul fascismo in modo autonomo, così poi da poter recuperare le loro analisi e tramite quelle dare avvio al lavoro.

La discussione si è mossa fin da subito sul filo dei concetti, quello che maggiormente ha coinvolto la classe è stato il concetto di appartenenza.

Questo concetto ha un significato nel nostro tempo: si appartiene ad un gruppo politico, ad una famiglia, ad un gruppo di tifo organizzato, ad un collettivo, ad una comunità, ad un quartiere, ad un gruppo di amici e ciò modifica il modo che si ha di vedere noi stessi all'interno della società. Lavorare su questo concetto dunque vuol dire lavorare anche sul nostro presente, vuol dire osservare la storia non come qualcosa che non riguarda il nostro quotidiano, ma al contrario ci aiuta a comprendere come la storia sia in divenire e non sia divenuto. Durante il fascismo il concetto di appartenenza assunse un significato molto più radicale, fino ad arrivare alle estremizzazioni razziali e nazionalistiche.

Ciò che mi colpì fu che in seguito all'analisi del concetto di appartenenza studenti e studentesse, che dovevano svolgere un ricerca all'interno dell'Archivio Storico dell'Istituto Luce, decisero di sviluppare il lavoro andando ad analizzare dei temi che con il concetto di appartenenza avevano una stretta connessione. La classe cercò le fonti audiovisive che potessero spiegare il processo di fascistizzazione dell'Italia indagando il tema della scuola, della comunicazione, dello sport, dell'urbanistica e dell'opposizione. Tranne l'ultimo tema, che chiaramente si pone in antitesi al concetto dell'appartenenza, ma ciò non di meno lo carica di significato, gli altri temi sono tra quelli

390 R. Koselleck, *Futuro passato. Per una semantica dei tempi storici*, Clueb, Bologna 2007, p. 97

391 https://www.camera.it/_dati/leg18/lavori/documentiparlamentari/indiceetesti/015/134/00000005.pdf ultima consultazione 03/08/2021

attraverso i quali è più semplice comprendere come si fosse sviluppato il concetto di appartenenza all'interno dell'Italia fascista.

Poter pensare di utilizzare il concetto di appartenenza come metadato, definiamolo appunto metadato concettuale, da attribuire unitamente ad altri metadati descrittivi ad alcuni documenti relativi al fascismo, consentirebbe di importare all'interno della fonte un altro livello informazionale oltre a quelli già esistenti, così da consentire una comprensione delle fonti e del presente attraverso l'utilizzo di un concetto che assumerebbe la funzione di spiegazione e consentirebbe di comprendere come il concetto di appartenenza, se anche può apparire innocuo, possa portare a conseguenze drammatiche, mettendoci dunque in una posizione critica rispetto all'utilizzo che ne facciamo nel nostro presente.

Guerra, liberazione, culto della personalità e vita quotidiana, sono i concetti sui quali in collaborazione con altre storiche[392] abbiamo realizzato un'installazione multimediale sulla storia del fascismo: dalla marcia su Roma alla liberazione. Ogni tema era veicolato attraverso uno schermo che riproduceva in loop una serie di fonti che durante la ricerca avevamo ritenuto specificatamente attinenti ai concetti scelti. L'installazione è stata realizzata in una piazza a Roma e il risultato è stato che cittadini e cittadine si potevano confrontare direttamente con le fonti attraverso un percorso strutturato unicamente in senso concettuale. Non vi erano processi cronostorici, non vi erano voci fuori campo, se non quelle montate in un contenuto audio realizzato unicamente con fonti primarie che si aggiungeva alla narrazione. Quello che si è provato a fare è stato di far esplodere in una piazza il contenuto di diversi archivi dando così la possibilità a chiunque passasse di confrontarsi con le fonti primarie, essendo guidati però solo dai concetti che apparivano sospesi sui relativi schermi, e da un "libretto delle fonti" dove per ogni schermo erano riportate tutte le fonti usate e la loro collocazione all'interno dei diversi archivi.

In questo caso i metadati concettuali capaci di clusterizzare i contenuti di diverse fonti al fine di sviluppare una narrazione, erano analogici. Si presentavano come elementi sovrainformativi che davano ai fruitori e alle fruitrici dell'installazione la possibilità di comprendere e contestualizzare ciò che stavano osservando.

Questo esempio ci aiuta a comprendere come si potrebbero usare dei metadati in un ambiente digitale, ma non mostra le reali possibilità di processi di metadatazione sviluppati in seguito all'analisi critica delle fonti all'interno di un archivio digitale.

L'installazione è stata realizzata con un processo preliminare di ricerca e selezione delle fonti da parte delle ricercatrici, a questo processo se ne è aggiunto uno di curatela, che ha portato

392 A. Foschi, M. Parma, O. Orsini, M. Lepre, M. Scanagatta, https://luceperladidattica.com/2017/04/26/prove-di-public-history-in-piazza/ ultima consultazione 12/07/2022

all’individuazione delle modalità di messa in scena delle fonti stesse. I fruitori dell’opera dunque, se pur attraverso un'intermediazione atipica, erano guidati all’osservazione del tema principale, la storia dell’Italia fascista, da una serie di interventi preliminari svolti da ricercatori e ricercatrici.
La metadatazione delle fonti all'interno di un archivio digitale, oltre a coinvolgere una mole documentale enormemente maggiore rispetto a quella proposta per l’installazione, consentirebbe al fruitore dell’archivio di interagire con le fonti e i contenuti in esse riportati in modo ancora più libero. I metadati concettuali potrebbero assumere il ruolo di spiegazione ed essere gli elementi di clusterizzazione/interoperabilità dei contenuti, ma all’interno di questi cluster consentirebbero ai ricercatori di muoversi liberamente tra le fonti.

4.2 Omeka, una sperimentazione

Durante la Summer School internazionale in Digital Humanities organizzata dall’Università di Modena e Reggio Emilia, con Matteo De Cristofaro, linguista che sta studiando le potenzialità dei metadati nell’analisi di corpora linguistici, abbiamo avuto l’opportunità di tenere un workshop dal titolo *metadata.*[393]

With many misconceptions surrounding it, metadata has come to play a central role in the interaction between humans and computers, making exchanges between human knowledge and computer information possible. From search engines to semantic web, metadata allows the retrieval of information from digital data, as well as allowing machines to represent “traces” of meaning across pieces of information stored as strings of 0s and 1s, de-facto translating human knowledge into digital objects that can be interpreted by computers. Consequently, how to correctly translate and represent knowledge into digital format becomes central to the possibility of accessing and querying digital data, and to ensure that this possibility is not hindered in the future due to technological or conceptual obsolescence. Similarly metadata should provide ways to interact with the data that are not limited to one single perspective or purpose, but rather that allow for multidisciplinary (re-)use of its contents. The workshop presents an experimental approach to the creation of metadata, combining theories and practices from historiography and linguistics in an effort to enrich digital objects with metadata derived from the analysis of both the context and the contents of digital sources.

Durante il workshop si è deciso di sperimentare la realizzazione/prototipazione di un “archivio corpus partecipato” andando a lavorare sulla metadatazione di documenti storici nati materici e

393 https://www.summerschooldigitalhumanities.unimore.it/2022-edition/scientific-programme/

digitalizzati relativi al conflitto Jugoslavo del 1991, con un focus specifico sulla guerra dei dieci giorni scoppiata in seguito alla richiesta di indipendenza da parte della Slovenia, e documenti nati digitali, relativi al conflitto ora in atto in Ucraina. Per questo secondo set di fonti abbiamo utilizzato i documenti prodotti dal Parlamento Europeo[394].

Si è deciso di lavorare su questi due eventi in quanto la prima e l'ultima guerra in Europa dopo la fine della seconda guerra mondiale.

L'analisi sui documenti si è sviluppata seguendo sia un approccio storiografico di analisi e critica delle fonti, con il quale abbiamo estrapolato dalle fonti metadati di interesse storico, sia con un approccio linguistico andando a lavorare attraverso l'annotazione/tagging sui testi.

I contenuti delle fonti, dunque, sono stati individuati attraverso due processi molto differenti, si è lavorato sia attraverso l'analisi semantica, umana, delle fonti, cercando nei contenuti parole chiave e temi che i gruppi di lavoro ritenessero interessanti per la descrizione delle fonti, sia sintattica, utilizzando LancsBox: Lancaster University corpus toolbox[395] un software di analisi computazionale, per individuare all'interno del corpora le ricorrenze e le relazioni sintattiche delle parole. Agendo su questo doppio livello abbiamo potuto ottenere risultati sia di ordine qualitativo, sia di ordine quantitativo.

Questa fase è stata seguita da un lavoro di comparazione e sintesi grazie al quale sono state estrapolate le keywords da utilizzare per la creazione di metadati a partire dai contenuti degli oggetti digitali.

Visto il tempo a disposizione, non era possibile far svolgere la ricerca dei documenti ai partecipanti e alle partecipanti del workshop, per questo con il dott. De Cristofaro abbiamo fatto una selezione preliminare dei documenti, che poi abbiamo condiviso ai partecipanti e alle partecipanti all'interno di una cartella Drive.

Ritengo necessario specificare questo passaggio per un duplice motivo. E' importante comprendere che l'intento del workshop non fosse quello di implementare le competenze di ricerca e le conoscenze storiografiche dei partecipanti e delle partecipanti, ma quello di farli e farle ragionare, attraverso la pratica, sull'importanza dei metadati nell'interazione tra uomo e macchina e sulle implicazioni possibili per la costruzione e l'organizzazione dell'informazione, in questo specifico caso di natura storica. L'altro elemento per cui ho ritenuto importante specificare che in una prima fase i documenti selezionati fossero all'interno di una cartella Drive, è che fin tanto che si trovavano in quel "luogo" digitale non li si poteva definire oggetti digitali, in quanto privi di una

394 Il portale del Parlamento Europeo si mostra come uno strumento di incredibile interesse, oltre ad avere documenti tradotti in tutte le lingue, ciò a cui si ha accesso è niente di diverso da un archivio istituzionale corrente o in certi casi di deposito. Questo ci dovrebbe far riflettere ulteriormente sui principi dell'accessibilità in epoca digitale e sulla necessità di definire sistemi di ordinamento e metadatazione già nelle fasi correnti e di deposito di un archivio.

395 http://corpora.lancs.ac.uk/lancsbox/ ultima consultazione 18/07/2022

metadatazione. È stato solo importando i documenti all'interno di Omeka, un programma open source di web-publishing, per la creazione e la fruizione di archivi digitali[396], che i file dopo essere stati metadatati sono diventati oggetti digitali.

Il workshop si è svolto dividendo i partecipanti in gruppi di lavoro, abbiamo dato loro le credenziali per scaricare Lanxbox, in modo da far processare al programma le diverse fonti digitali e digitalizzate, mentre per l'utilizzo di Omeka abbiamo creato un'istanza apposita per la Summer School, sulla quale i partecipanti di ogni gruppo di lavoro hanno potuto importare i documenti che decidevano di metadatare.

La scelta di utilizzare Omeka è stata presa perché oltre la semplicità di utilizzo, questo strumento consente di interagire con le fonti utilizzando diversi standard di metadatazione e vocabolari per costruire ontologie.

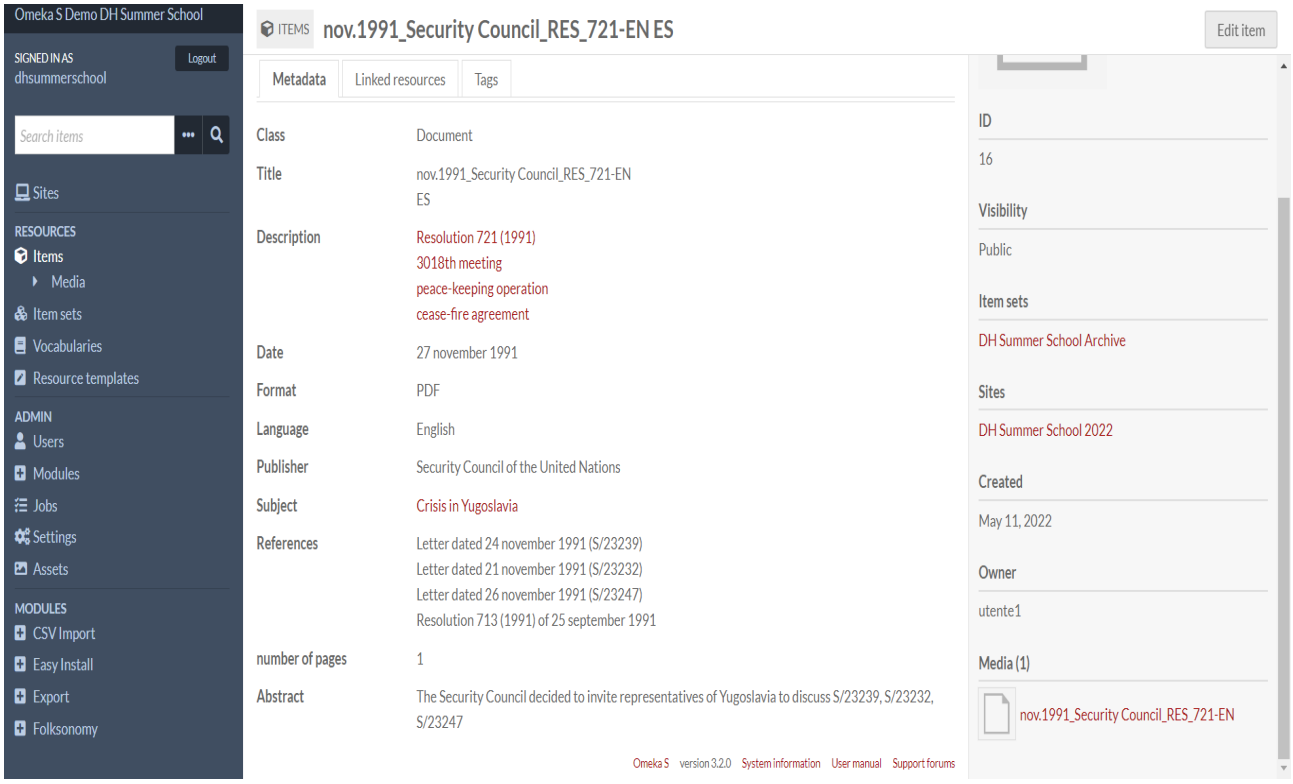
Nella voce "Vocabularies" il sistema consente di scegliere tra Bibliographic Ontology, Dublin Core, Dublin Core Type e Frinds of a Friends. La nostra scelta è ricaduta sull'utilizzo di Dublin Core, che su Omeka si presenta con uno schema composto da 55 elementi. Come abbiamo detto nelle pagine precedenti uno dei grandi vantaggi di DCMI è la sua flessibilità, dunque la possibilità di scegliere quali elementi utilizzare per la metadatazione delle fonti. Durante il workshop si è deciso di lasciare libera scelta ai frequentanti sia sui documenti da analizzare sia sui campi di metadati da importare, sottolineando però la necessità di importare metadati descrittivi relativi al titolo del documento, l'anno di produzione, il soggetto produttore, l'oggetto e la tipologia di documento. Indicazioni archivistiche necessarie per validare l'esistenza e la tracciabilità del documento stesso.

La parte più interessante è emersa nel momento in cui i partecipanti hanno iniziato a chiedersi come poter utilizzare il campo "Description" nello schema di Dublin Core. Questo elemento infatti consente di indicare descrizioni archivistiche per la gestione del documento, così come consente di provare a sviluppare degli elementi descrittivi estrapolati dal contenuto del documento stesso.

Come possiamo vedere nell'immagine sottostante, uno dei gruppi di lavoro ha deciso di metadatare la Risoluzione 721 del Security Council delle Nazioni Unite rilasciata nel novembre del 1991, utilizzando come descrizione anche i concetti di "peace keaping operation" e "cease fire agreament". Questi termini non sono esplicitamente espressi nel titolo del documento, quindi potrebbero non essere formalmente necessari all'indicizzazione del documento stesso, ma in seguito all'analisi computazionale e semantica del testo, per i partecipanti rappresentavano elementi significativi per la comprensione e per la ricerca del documento all'interno di un sistema di archiviazione digitale, che come abbiamo detto è anche strumento di comunicazione. Omeka in

396 https://omeka.org/ ultima consultazione 18/07/2022

particolare consente di utilizzare gli oggetti digitali e i relativi metadati per creare un sito Web con dei template che mette a a disposizione.



Ai fini del workshop non era necessario realizzare un sito Web, ma anche nell'interfaccia di back office, in basso a destra nel campo definito "Media" è possibile cliccare sul link in rosso per trovarsi davanti alla fonte primaria, che con l'azione di metadatazione è stata trasformata in un oggetto digitale e in questo specifico caso, visto la ricorsività dei descrittori, in un oggetto informativo.

Poter utilizzare anche le parole "cease fire" per fare una ricerca all'interno di un archivio o del Web generalista, può aumentare le possibilità di individuare documenti che raccontino come e se il Consiglio di Sicurezza delle Nazioni Uniti si fosse espresso in merito al verificarsi del conflitto jugoslavo. L'utilizzo di questo tipo di metadatazione non è utile solo a dare chiavi di ricerca funzionali per fruitori che non necessariamente conoscono l'oggetto o il titolo dei documenti di cui hanno bisogno, è uno strumento incredibilmente utile anche per sviluppare un'intreroperabilità tra documenti attraverso chiavi di lettura tematiche e concettuali.

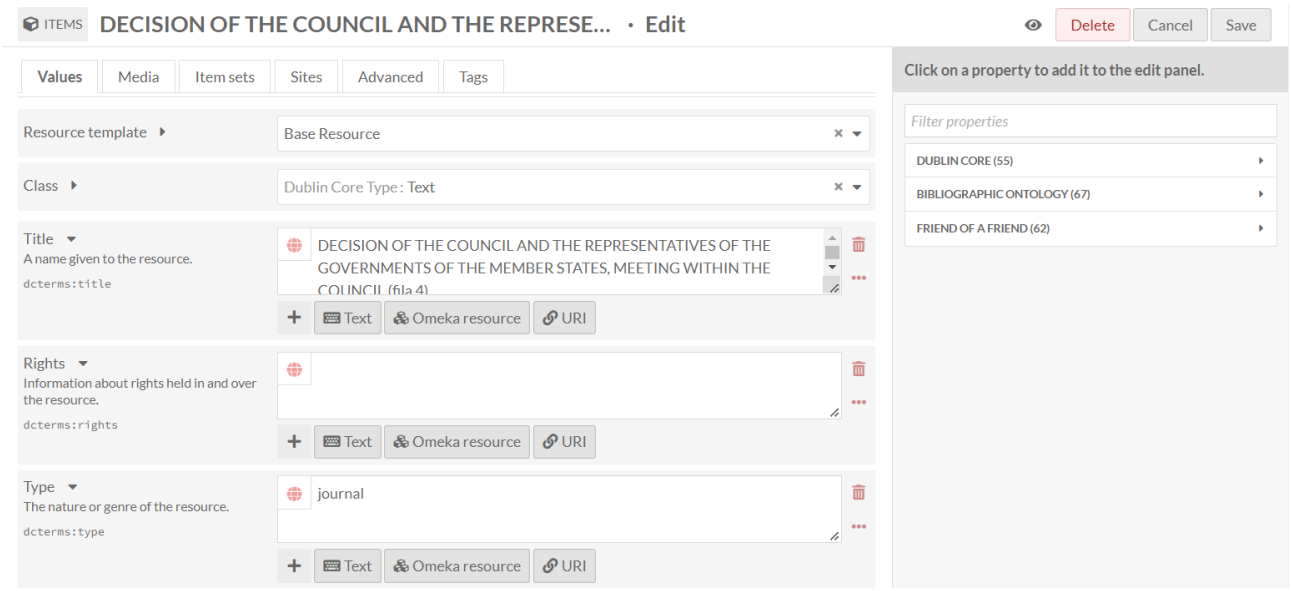
Se nella barra di ricerca del sistema di Omeka si scrivono le parole "peace" o "cease fire", sempre per rimanere su questo esempio, il sistema restituirà come risultato tutti i documenti che sono stati metadatati con questi termini, consentendo al fruitore di trovarsi davanti ad un'eterogeneità di elementi informativi accomunati da termini o concetti specifici.

In questo caso possiamo vedere che utilizzando il concetto “cease fire” il sistema risponde fornendoci l’accesso a tre documenti differenti, caricati e metadatati da due diversi gruppi di lavoro: un verbale previsionale, una risoluzione del Consiglio di Sicurezza e un articolo di giornale.

L’utilizzo del concetto “cease fire” in relazione a tre documenti differenti tra di loro può aiutare il fruitore non solo a caricare di significato il documento che è stato metadatato, ma consente anche di poter meglio comprendere il concetto di “cease fire” mettendolo in relazione con i contenuti informativi dei diversi documenti.
Il campo “Media” consente di inserire all’interno di Omeka documenti di ogni tipo, come abbiamo più volte ripetuto video, testi, immagini o audio, una volta subito il processo di rimediazione o se nati digitali, sono tutti composti di bit, per questo per il sistema digitale sarà possibile accogliere al proprio interno e rendere interoperabili tipologie di documenti molto differenti tra di loro e una volta metadatati e dunque dopo aver subito un processo di sovrascrittura, ogni documento risponderà alle chiavi di ricerca espresse dal ricercatore o dalla ricercatrice.
Quando si accede nella pagina di edit dello schema di metadati in relazione al documento da descrivere, vediamo che in ogni campo di metadatazione che si decide di utilizzare, è possibile indicare un URI.

Gli URI costituiscono la tecnologia di base ideale con la quale costruire un Web globale. Possiamo definire un URI per un qualsiasi oggetto, e qualsiasi cosa che ha un URI può essere considerato sul Web. Gli URI sono il fondamento del Web: mentre ogni parte del Web stesso può essere rimpiazzata, gli URI no. Poiché il Web è troppo esteso per essere controllato da una qualsiasi organizzazione, gli URI in massima parte sono decentralizzati. Nessuna persona o organizzazione controlla chi li produce o cosa ne fa. Un URI non è un insieme di direttive che indicano al computer dove trovare un file specifico nel Web (sebbene lo faccia anche), ma è un nome per una risorsa (una cosa), accessibile o meno attraverso Internet. L'URI può o no fornire un modo per ottenere più informazioni su una risorsa.[397]

397 A. Dorati, S. Costantini, *Approcci al Web semantico*, in, https://www.websemantico.org/articoli/approcciwebsemantico.php



Andando ad assegnare degli URI agli oggetti informativi, stiamo ponendo le basi per creare l'autoconsistenza del dato, che consente di andare a realizzare le triple del Web semantico. All'interno delle triple, soggetto e oggetto sono necessariamente degli URI, così come lo sono i concetti rappresentati dai predicati.

Il workshop è stato un primo esperimento molto utile per comprendere meglio il funzionamento dei metadati e i possibili processi di metadatazione in relazione ad una comunità di riferimento. Senza nessuna pretesa di essere una sperimentazione sufficientemente valida, grazie a questo workshop si è comunque potuto testare un processo d'analisi dei documenti sia in senso qualitativo che quantitativo, il coinvolgimento di una comunità di riferimento molto eterogenea ci ha anche mostrato le possibilità e le complessità di un metadatazione partecipata, durante la quale siamo riusciti ad estrapolare diversi significati dalle fonti, facendone dunque emergere la complessità informativa. L'interoperabilità delle fonti ci ha mostrato come sia possibile costruire delle narrazioni all'interno di un archivio digitale utilizzando i metadati come elemento portante della spiegazione. Non solo, in una visione di Public History, un processo come quello descritto ha consentito ad un pubblico non specialistico di confrontarsi con le fonti primarie e interrogarsi sui processi di analisi critica tipici del lavoro di storiche e storici, una vota conclusa la fase di metadatazione i frequentanti hanno potuto osservare e sperimentare un nuovo processo di narrazione della storia.

Ciò che è anche emerso dal workshop è la difficoltà di elaborare molti e diversi contenuti, il tempo è un elemento fondamentale al fine di raggiungere un buon obiettivo, così come lo è una conoscenza degli argomenti a cui fanno riferimento le fonti. Queste difficoltà, ci fanno comprendere come siano complessi i processi di costruzione di un archivio digitale tramite metadatazione. Per

metadatare anche con contenuti e concetti un intero corpus archivistico è necessario che le diverse competenze coinvolte nella metadatazione dei documenti portino avanti un costante lavoro di confronto e condivisione delle proprie analisi. Nel caso specifico del workshop il tempo a disposizione non ci ha permesso di sviluppare in modo adeguatamente approfondito un lavoro di equipe[398] che consentisse, prima di arrivare al processo di metadatazione, di aprire una discussione collettiva su quanto era emerso in seguito all'analisi dei documenti. La mancanza di questo passaggio ha portato, per esempio, a non metadatare con i concetti di “cease fire” o “peace keaping” anche il documento relativo alla “Brioni Declaration”, per quanto questi concetti siano alla base del documento stesso.

È importante sottolineare però, che nonostante l’assenza di una metadatazione concettuale la Brioni Declaration esiste come oggetto digitale all’interno dell’archivio ed è possibile rintracciarlo utilizzando le parole che ne compongono l’oggetto o la data del documento. Basta dunque scrivere “1991” all’interno della barra di ricerca del sistema per ottenere questo risultato:

Title	Class
nov.1991_Security council_RES_721-EN (1)	Decision
nov.1991_Security Council_RES_721-EN ES	Document
DECISION OF THE COUNCIL AND THE REPRESENTATIVES OF THE GOVERNMENTS OF THE MEMBER STATES, MEETING WITHIN THE COUNCIL of 11 November 1991 (FP)	Legal Document
Stern EC warning over Brioni agreement	Text
PROVISIONAL VERBATIM RECORD OF THE THREE THOUSAND AND NINE MEETING (1)	Text
Stern EC warning over Brioni agreement (fila 4)	Text
Stern EC warning over Brioni agreement - IS	Article
Brioni declaration (Fila 4)	Legal Document
DECISION OF THE COUNCIL AND THE REPRESENTATIVES OF THE GOVERNMENTS OF THE MEMBER STATES, MEETING WITHIN THE COUNCIL (fila 4)	Text
Stern EC warning over Brioni agreement (fila 4)	Newspaper
Resolution 721 (Fila 4)	Text

398 Quando parlo di lavoro di equipe in ambito di ricerca storica e costruzione storiografica mi rifaccio alle teorizzazioni espresse dall’Equipe Sperimentale di Storia nel saggio *Fuggite sciocchi. Riflessione su ciò che non esiste*. https://www.academia.edu/17729970/Fuggite_Sciocchi_riflessione_su_ci%C3%B2_che_non_esiste

Da questo screenshot possiamo notare due cose: la prima è una ridondanza di alcuni documenti, questo è dovuto al fatto che i gruppi di lavoro hanno lavorato senza dividersi i documenti da analizzare, e ciò ha portato inevitabilmente ad avere lo stesso documento metadatato da diversi gruppi.

Al fine di questa sperimentazione questo però non è un problema, anzi ci consente di osservare come ogni gruppo abbia deciso di andare a lavorare sui documenti facendo emergere le diverse scelte prese.

La seconda cosa che possiamo notare è che è bastato inserire una data per avere accesso a tutti i documenti importati in Omeka e prodotti nel 1991. Come indicazione iniziale avevamo detto ai e alle partecipanti del workshop che fossero tenute e tenuti ad utilizzare il campo “Date”, in quanto elemento necessario alla descrizione archivistica di ogni documento. Soffermarsi su questo passaggio ci aiuta a comprendere che il possibile utilizzo di metadati concettuali in nessun modo va a ledere o entra in conflitto con le prassi di indicizzazione archivistica, ma consente, potenzialmente, di aumentare il portato informativo. Dublic Core ammette la ripetibilità dei propri elementi all’interno dello schema di metadati e ciò proprio per consentire una maggiore libertà descrittiva.

In una prospettiva laboratoriale in un’ottica di diffusone sociale della storia, penso possa essere formativo far lavorare le classi con uno strumento come Omeka, che nella sua versione base è gratuito. A partire dallo studio della letteratura storiografica si potrebbe andare a lavorare con alcune fonti relative all’evento o al periodo storico preso in esame. Se Hilda Girardet con il suo testo sull’insegnamento della storia con le fonti fa emergere le positive implicazioni di una didattica sviluppata non solo con i manuali, ma anche attraverso le risorse dello storico, la possibilità di utilizzare Omeka per costruire dei piccoli archivi inventati con l’ausilio di metadati concettuali, richiederebbe a studenti e studentesse di doversi confrontare con i contenuti delle fonti per estrapolare da esse dei significati da attribuire all’oggetto digitale. Ciò renderebbe il processo di apprendimento inevitabilmente collegato alla comprensione dei testi e delle fonti e potrebbe aiutare a rendere la storia come qualcosa da capire e non solo da imparare. Sarebbe interessante osservare la scelta dei concetti e dei contenuti selezionati per la metadatazione anche in una prospettiva narrativa. Una volta consapevoli delle possibilità date dall’interoperabilità delle fonti tramite temi e concetti, studenti e studentesse potrebbero sperimentare la costruzione di una innovativa forma di storiografia dove appunto la narrazione avviene tramite l’interoperabilità che a sua volta prende forma attraverso l’utilizzo dei concetti come spiegazione e contestualizzazione.

Prima di procedere nell’analisi di OAIS, un sistema di gestione per gli archivi digitali interamente strutturato sullo scambio informazionale tra diversi livelli che si snodano dalla produzione della

fonte fino alla comunità di riferimento, vorrei provare a condividere un ragionamento che credo ora abbia delle basi tecniche abbastanza solide per essere affrontato.

In base a ciò che abbiamo scritto fino ad ora, credo che la rivoluzione digitale per come ha investito le fonti e gli archivi ci offra la possibilità di riattualizzare nell'oggi il concetto di *Geschichte* declinato come l'unione tra ciò che è accaduto (*Res Gestae)*, la narrazione della storia o dell'avvenimento accaduto (*Historia Rerum Gestarum o Historie*) e la conoscenza degli avvenimenti accaduti, la conoscenza della storia (*Geschichtkunde)*[399], con la consapevolezza che mai noi potremmo avere un accesso "puro" all'evento del passato.

Da sempre abituati a confrontarci con un processo storiografico che come emerge dal termine stesso vede nella scrittura l'elemento centrale e maggiormente funzionale per la comunicazione della storia, siamo abituati a confrontarci con pagine di libri dove storici e storiche in seguito alle loro ricerche elaborano una struttura narrativa in cui, attraverso l'utilizzo di fabula e intreccio, riportano analisi e argomenti derivanti dall'osservazione e lo studio delle fonti primarie e di altra storiografia. In questo processo il fruitore poteva avere come unico accesso alla fonte primaria, delle note con la collocazione dei documenti all'interno dell'archivio. Negli ultimi anni diversi progetti di Public History e di Digital Public History hanno prodotto dei contenuti storiografici in cui la fonte viene messa in scena, dando dunque per la prima volta la possibilità al fruitore di confrontarsi non solo con la narrazione dello storico o della storica, ma con la fonte primaria. Ciò che però abbiamo descritto nelle pagine precedenti, ci consente, forse, di immaginare per il futuro una relazione maggiormente strutturale tra il lavoro della storica e l'ambiente digitale. Ciò che è possibile immaginare, grazie ai metadati, è di poter aiutare o se vogliamo guidare il fruitore all'interno dell'archivio stesso, è l'archivio che si trasforma nel libro che viene aperto, i metadati nelle parole e le fonti nelle frasi.

Oggi gran parte dell'informazione la fruiamo attraverso device che teniamo costantemente in tasca e capaci di riprodurre ogni forma di oggetto digitale. Avere la certezza che esistono sistemi di ordinamento e metadatazione delle fonti che consentono una fruibilità sicura e affidabile mi convince che oltre allo smartphone non sia necessario avere in tasca anche public historian sempre disponibili a guidarti nella ricerca, ma se il pensiero e le capacità dei public historian venissero integrate all'interno dell'oggetto digitale grazie ai metadati si potrebbe davvero pensare di essere entrati in una nuova era della diffusione del contenuto storico. Non dimentichiamoci che siamo all'interno di una rivoluzione.

399 R. Koselleck, *Storia, la formazione del concetto moderno*, Clueb, Bologna 2009 p. 40

5 Da OAIS al Web semantico

Open Archival Information System, OAIS, che come scrive Silvio Salza è oggi unanimemente considerato il riferimento per i sistemi di conservazione digitale, è stato introdotto come raccomandazione nel 2002 nell'ambito delle agenzie spaziali, interessate alla conservazione dell'enorme mole di dati proveniente dai satelliti, è successivamente diventato uno standard ISO nel 2003, ed ha subito un aggiornamento nel 2012.
Questo modello mette in relazione tutti i passaggi che determinano la vita di un archivio: produzione, conservazione/gestione, fruizione.
OAIS è un sistema di gestione e archiviazione digitale, in quanto tale deve essere in grado di contenere e elaborare differenti dati che conducano alla definizione di quelle che possiamo sostenere essere le quattro componenti maggiormente significative nei processi di digitalizzazione e gestione delle fonti: creazione del Digital Object; Rappresentation of Information; individuazione della Designated Community; comprensione e condivisione della Knoweldge Base. Ognuna di queste componenti richiede di essere analizzata dal momento che coinvolgono le fasi significative del lavoro di archiviazione, ricerca, analisi e diffusione delle fonti.
Per ognuna di queste fasi si aggiungono e si pensa come organizzare delle informazioni che entrano in relazione diretta con le informazioni che determinano l'esistenza e il valore stesso del documento storico. Ancora una volta vediamo come i livelli informativi riferiti al documento si stratificano e richiedono una differente e nuova lettura e comprensione rispetto alla consueta archiviazione, ricerca e analisi di fonti materiche.

Il modello di riferimento OAIS (Open Archival Information System) si propone come quadro concettuale unitario per descrivere oggetti, processi, strategie e tecniche finalizzati alla conservazione digitale a lungo termine, nonché per comprendere le loro reciproche relazioni e per analizzare e confrontare soluzioni conservative diverse. Il modello si basa su un'idea dinamica della conservazione, intesa come processo permanente e mai concluso di monitoraggio del contesto in cui sono immersi gli oggetti; e disegna una complessa architettura di funzioni e oggetti informativi fondata sull'individuazione delle risorse necessarie per ricostruire il significato degli oggetti, assumendo il bit come unità minima del sistema concettuale[400].

Gli elementi che maggiormente caratterizzano OAIS possono essere rintracciati nel fatto che il sistema si mostra come aperto, ossia disponibile per una comunità internazionale e ottimizzabile con il contributo di tutti, non solo, in questo sistema è centrale il concetto di conservazione a lungo

400 G. Michetti, *Il modello OAIS*, in «Digitalia», Anno III, n° 1, 2008

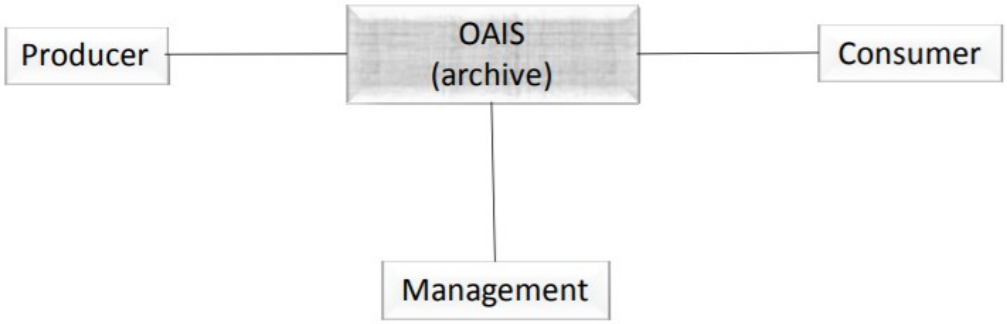
termine, dunque prevede che vi possano essere diversi cambiamenti tecnologici sia riguardo ai formati elettronici, sia riguardo i supporti di memorizzazione. Questo passaggio è fondamentale per prevenire l'obsolescenza tecnologica. Spesso infatti si è convinti che la conservazione digitale sia più sicura rispetto a quella materica, in realtà la fonte dematerializzata è molto più "fragile" in quanto rimane intellegibile per i diversi fruitori solo se sempre aggiornata alle tecnologie in uso in quel momento. Per fare un semplice esempio, se oggi un ricercatore dovesse cercare delle fonti digitali registrate in un floppy disk avrebbe grande difficoltà a reperire un hardware e un software capace di leggere le informazioni all'interno di quel supporto. Pensiamo al caso, precedentemente accennato, dell'archivio digitale della testata giornalista «La Stampa» che nel 2020 ha rischiato completamente di sparire perché sviluppato con una tecnologia Java ormai obsoleta.

Tra poche settimane, 138 anni di storia italiana scompariranno dal web, a meno di qualche miracolo finanziario e amministrativo, e non si sa con precisione quando o come potranno ricomparire. L'archivio storico della Stampa rischia di rimanere vittima della rapida obsolescenza degli strumenti digitali e di una serie di complicate circostanze proprietarie e tecnologiche. Stiamo parlando di oltre 12 milioni di articoli e di 1.761.000 pagine pubblicati a partire dal 9 febbraio 1867, quando uscì a Torino il primo numero della Gazzetta Piemontese (diventata La Stampa nel 1894), fino al 2005[401]

Fortunatamente, grazie all'intervento della Regione Piemonte[402] l'archivio è stato salvato, ma questo episodio ci deve far riflettere sull'importanza di costruire sistemi di archiviazione a lungo termine, altrimenti, contrariamente a quanto sostenuto da Ferraris, l'oblio potrebbe essere maggiore che in un passato/presente in cui le fonti materiche una volta collocate all'interno di un archivio storico, se conservate in modo adeguato, dovevano temere quasi solo alluvioni o incendi.
L'importanza di OAIS però non risiede solo nella sua flessibilità per la conservazione a lungo termine, questo sistema di gestione dei dati ha infatti un ruolo determinante anche nella possibilità di fruire i dati conservati. OAIS infatti è in sé anche una ontologia, in quanto offre un modello concettuale e funzionale che fornisce un vocabolario condiviso, ma senza la necessità di una specifica piattaforma tecnologica. Ciò rende possibile e semplice la realizzazione di sistemi di conservazione distribuiti e interoperabili.

401 M. Tedeschini-Lalli, *Tecnologia digitale obsoleta. Un secolo e mezzo di Storia a rischio*, 27 nov. 2020, https://tedeschini.medium.com/tecnologia-digitale-obsoleta-un-secolo-e-mezzo-di-storia-a-rischio-1bb75bf68c2f ultima consultazione 13/1/2022
402 M.Tedeschini Lalli, *Un secolo e mezzo di Storio di nuovo online*, 23 feb. 2021, https://tedeschini.medium.com/un-secolo-e-mezzo-di-storia-di-nuovo-online-594a9c1e6358, ultima consultazione 13/1/2022

Tutto ciò assume ancora più significato se pensiamo che questo sistema «è stato immaginato per essere personalizzato per una comunità designata che presuppone lo studio preliminare delle esigenze degli utenti del sistema di conservazione e utilizzo di una terminologia appropriata»[403]
Prima di approfondire il concetto di comunità designata che può essere analizzato in modo davvero complesso e interessante - sopratutto in una visione di diffusione sociale del contenuto storico - credo che possa essere utile andare a vedere, grazie a schemi esemplificativi, quale sia il funzionamento di OAIS.
Dal momento che non siamo archivisti informatici mi limiterò a condividere gli schemi maggiormente comprensibili e a mio avviso esplicativi di quello che è un sistema dalle grandi potenzialità e da un funzionamento logico tutto sommato abbastanza semplice.



Questa immagine, tratta dal documento *Reference model for an open archival infromation sistem*[404], mostra come OAIS si incarni nell'archivio e si ponga come strumento di mediazione e gestione informazionale rispetto a tre enti differenti ma tutti egualmente fondamentali: il producer, ossia il soggetto produttore; il management che si pone sotto all'archivio perché è l'ente tecnico amministrativo che deve prendere in carico la gestione di dati e metadati in accordo con il produttore, e in fine sulla destra quindi alla fine del percorso, il consumer, ossia, chiunque potrà fruire dei documentari, dunque banalmente, anche noi storiche e storici, ma in un'otica più public è possibile immaginare come consumer, utenti non specializzati e in particolar modo studentesse e studenti.
Questa prospettiva, trova però alcune difficoltà nell'idea di comunità di riferimento, dal momento che teoricamente la comunità di riferimento dovrebbe essere rappresentata da coloro che hanno le

403 S. Pigliapoco, *Progetto archivio digitale. Metodologia, sistemi professionalità*, Civita Editoriale, Lucca, 2018 p. 51
404 The consultative committee for Space Data System Practices, *Reference model for an open archival infromation sistem*, Recommendend Practice, 2012, https://public.ccsds.org/pubs/650x0m2.pdf

competenze per comprendere ciò che l'archivio contiene. Questo problema può trovare soluzione ancora una volta solo nel rinnovato ruolo sociale di storiche e storici che dovranno essere in grado di formare studenti e studentesse alla lettura e fruizione di un archivio, a maggior ragione se come vedremo la comunità di riferimento ha in OAIS un ruolo attivo e non solo passivo. È infatti attraverso un'interazione con il sistema che si definisco i pacchetti informativi relativi alla parte di disseminazione, questo implica che studenti e studentesse, potrebbero avere un ruolo fondamentale nella costruzione di un complesso percorso informazionale, ripensiamo ai risultati ottenuti con il workshop sui metadati.

In prima istanza andiamo ad osservare in che modo OAIS genera ciò che nel suo schema viene definito Information Object.

La definizione che anche Feliciati da di Information Object non differisce molto da quella che abbiano già dato di Digital Object, con la differenza che in questo caso, la natura dei metadati che vengono attribuiti alla fonte digitale ha un chiaro ruolo di rappresentazione e dunque di informazione.

Per Pigliapoco «le informazioni di rappresentazione contengono indicazioni sia di natura sintattica che semantica e svolgono un ruolo analogo a quello che il dizionario della lingua italiana e le regole grammaticali assumono per la comprensione di un testo scritto in italiano.»[405]

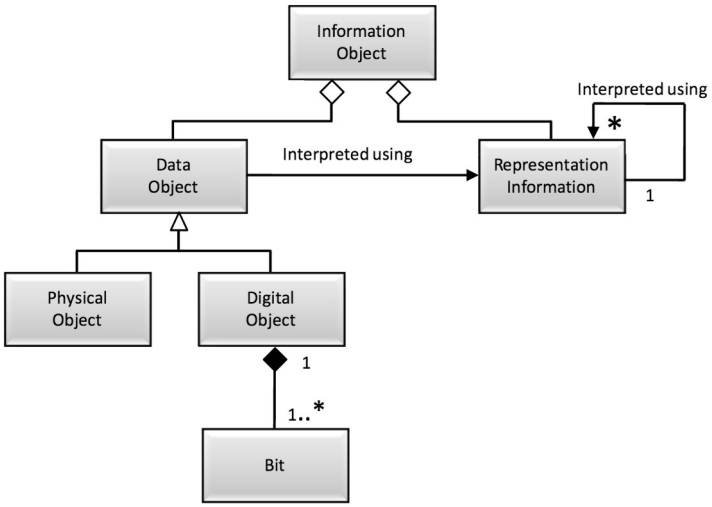


Figure 4-10: Information Object

405 S. Pigliapoco, *Progetto archivio digitale. Metodologia, sistemi professionalità*, Civita Editoriale, Lucca, 2018 p. 152

Come vediamo dallo schema, preso dalla National Digital Stewardship Residency[406], la parte di rappresentazione dell'informazione consiste in una fase in cui il dato sull'oggetto digitale viene interpretato dalla macchina al fine di poter essere leggibile da un essere umano, consentendo così la rappresentazione e la comprensione al livello dell'uomo.

In sostanza, la codifica dell'informazione digitale, a differenza di altre, non è mai né auto-sufficiente né auto-esplicativa, ma deve sempre e necessariamente documentare sé stessa al livello minimo del singolo atomo di informazione, aggiungendo al dato contenuto vero e proprio molte informazioni necessarie per la decodifica, l'identificazione, il recupero, l'accesso e uso, etc. Un vero e proprio doppio delle risorse insomma[407].

Il tema della rappresentazione dell'informazione per la creazione di un information object è assolutamente centrale in un sistema come OAIS che ha come scopo ultimo quello di consentire ad una comunità di riferimento la fruizione dei documenti, per questo si considera indispensabile associare agli oggetti digitali delle informazioni di natura anche semantica comprensibili e sviluppati in sinergia con la suddetta comunità.
Prima di andare ad osservare il funzionamento e l'interazione tra i tre Information Packages vorrei soffermarmi ancora sul concetto e l'importanza della comunità d riferimento, questo perché appunto ritengo che sia il nodo centrale per immaginare un processo di diffusione delle fonti.
Leggendo Pigliapoco si può avere l'impressione che il sistema OAIS sia di particolare interesse per l'archiviazione di fonti con una specifica valenza tecnico amministrativa, indirizzando così l'idea della comunità di riferimento a tecnici che hanno già competenze specifiche rispetto alla tipologia dei documenti conservati. In prima battuta si ha la stessa impressione anche leggendo Lavoi, che in modo molto interessante sottolinea come «il modello di riferimento precisa una classe speciale di utenti indicata come la "comunità di riferimento": il sottoinsieme degli utenti in grado di comprendere autonomamente l'informazione archiviata nella forma in cui è conservata e resa disponibile dall'OAIS.»[408]

406 National Digital Stewardship Residency, https://ndsr.nycdigital.org/tag/oais-reference-model/
407 P. Feliciati, *I metadati nel ciclo di vita dell'archivio digitale e l'adozione del modello PREMIS nel contesto applicativo nazionale*, UniMc, Pubblicazioni aperte digitali, 2015
408 F.B. Lavoi, *Il modello di riferimento per un Sistema informativo aperto per l'archiviazione,* Associazione Italiana Biblioteche, 2017, p. 9 http://www.soprintendenzaarchivisticatoscana.beniculturali.it/fileadmin/risorse/materiali/Lavoie_-_OAIS.pdf

Andando avanti nella lettura del testo però si comprende come il concetto di comunità di riferimento possa essere estendibile senza confini predeterminati, anzi, lo stesso Lavoi arriva a parlare di «tutta la popolazione». Questo ampliamento del concetto di comunità di riferimento non ha una ricaduta solo in un ambito teorico, più si allarga la comunità di riferimento e maggiori dovranno essere i metadati da utilizzare, la rappresentazione dell'informazione dovrà divenire comprensibile dunque ad un numero di persone molto ampio e diversificato. Questo comporta un maggior lavoro nell'attribuzione di metadati ma anche una contestualizzazione talmente ampia da rendere un archivio gestito da un modello come OAIS un luogo sempre più semplice da fruire.

E' stato prima ricordato che i confini della "comunità di riferimento" hanno un impatto sulla quantità dei metadati richiesti per sostenere il processo di conservazione. Si tratta appunto di metadati che costituiscono la *Representation Information*. In generale, più ampio è il confine della "comunità di riferimento", meno specializzata è la base di conoscenze della "comunità di riferimento". In questo caso occorrerà includere un maggior numero di metadati nella *Representation Information* per assicurare che l'informazione conservata resti fruibile e comprensibile alla "comunità di riferimento" nel lungo periodo[409].

Michetti riesce in modo esemplare a spiegare come funzioni la stratificazione di dati al fine di generare un information object con una rappresentation of information altamente sviluppata, quale sia il reale ruolo di una comunità di riferimento e quali l'importanza di riuscire a far muovere in modo sincronico un flusso informazionale necessario per garantire una conservazione e un fruizione delle fonti sul lungo periodo.

Il modello muove quindi dai dati per costruire su di essi una struttura concentrica di metadati tramite il meccanismo ricorsivo: i dati richiedono altri dati per poter essere correttamente interpretati; questi a loro volta possono richiedere altri dati interpretativi; e così via, finché non si raggiunge un livello autosufficiente, determinato dalla base di conoscenza della comunità di riferimento. Qui si arresta un'iterazione che altrimenti condurrebbe paradossalmente alla conoscenza del mondo: ad esempio, nel caso di un testo italiano in formato digitale, trascurando per brevità tutti gli altri aspetti, non sembra necessario annoverare un vocabolario tra le informazioni di rappresentazione; ma se ci si pone nell'ottica della conservazione a lungo termine un orizzonte temporale così ampio può giustificare la conservazione delle informazioni relative a semantica e sintassi del linguaggio naturale, data la possibilità di evoluzioni significative nell'uso. Il problema è determinare il momento in cui la cultura dell'utenza di riferimento non è più adeguata a interpretare un oggetto con le consuete informazioni di rappresentazione: quando è necessario introdurre

409 F.B. Lavoi, *Il modello di riferimento per un Sistema informativoaperto per l'archiviazione,* Associazione Italiana Biblioteche, 2017, p. 18 http://www.soprintendenzaarchivisticatoscana.beniculturali.it/fileadmin/risorse/materiali/Lavoie_-_OAIS.pdf

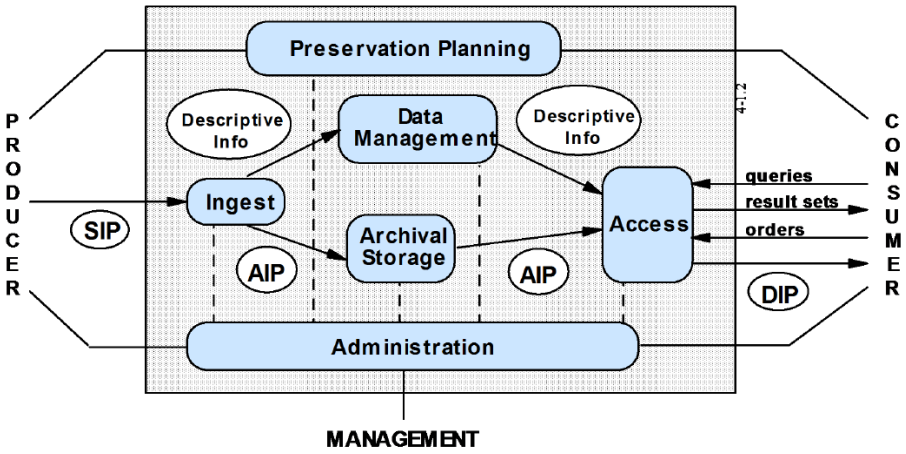
questo ipotetico vocabolario? fra cinquanta anni? o cento? La risposta è ovviamente vincolata alle caratteristiche della comunità di riferimento, ma l'aspetto che qui si vuole sottolineare è la dinamicità della funzione conservativa, che richiede una sensibilità costante nei confronti dell'utenza al fine di aggiornare la risorsa archiviata.[410]

I livelli di Infomation Package che determinano il funzionamento di OAIS sono tre e devono essere gestiti in collaborazione dall'ente che produce i documenti e da quello che ne garantisce la conservazione e la fruizione. Questo ci mostra un inedito collegamento tra il soggetto produttore e il possibile fruitore

In modo sintetico il modello funziona nell'interazione tra:

- Submission Information Package (SIP), che viene trasmesso nella fase di versamento dal produttore al deposito;
- Archival Information Package (AIP), che viene generato a partire dal SIP in fase di accettazione (Ingestion) e poi diventa oggetto diretto della conservazione deposito;
- Dissemination Information Package (DIP), che viene generato a partire dall'AIP per essere distribuito alla Comunità Designata per la fruizione

I tre livelli di Information Package sono i livelli che maggiormente ci aiutano a riflettere su quale sia e possa essere in ambito digitale la relazione tra fonte-archivio-fruitore che se riportati in un'ottica di Public History o di Digital Public History richiamano al rapporto tra storia-storico-pubblico



410 G. Michetti, *Il modello OAIS*, in «Digitalia», Anno III, n° 1, 2008 p. 42

Questo schema ci aiuta a comprendere ciò che succede alle fonti quando entrano all'interno di un processo di archiviazione elettronico con l'ausilio di metadati.
In una struttura così organizzata il fruitore assume un ruolo centrale, l'accessibilità alle fonti non è data solo da un accesso facilitato, ma dalla ricerca e l'analisi di una comunità di riferimento con la quale costruire una conoscenza di base che permetta la fruizione migliore del fondo archivistico. Osservando lo schema vediamo che la relazione tra il "consumer" e il "management" è una relazione dialogica, il DIP si genera in seguito a una interlocuzione tra i possibili fruitori dell'archivio e l'archivio stesso.
Se è vero che anche all'interno degli archivi materici, indici, protocolli e cataloghi sono immaginati anche per una fruizione esterna è altresì vero che uno schema come quello rappresentato apre a nuove e affascinanti pratiche di ricerca. Quello che ci appare subito chiaro è che l'utilizzo dei pacchetti informativi determina una stratificazione di dati sulla fonte, sia nel processo di submission sia in quello di dissemination e ciò comporta inevitabilmente che aumenti il livello di informazione. Se per esempio, come per il workshop, ipotizziamo di utilizzare lo standard Dublin Core, che è compatibile con il modello OAIS, si può immaginare di inserire nei pacchetti informativi sia nella fase di gestione, quindi in AIP, sia nella fase di disseminazione quindi in DIP, dei metadati descrittivi che per il fruitore possono avere un valore semantico utile a meglio trovare e comprendere le fonti che sta cercando.

5.1 Interoperabilità

L'interoperabilità tra archivi digitali e oggetti digitali come abbiamo più volte detto, è uno degli elementi maggiormente innovatiti che la digitalizzazione dell'informazione, inserita in un contesto ordinato di archiviazione ci può offrire.
OAIS come modello di gestione a lungo termine dell'informazione digitale consente di costruire processi di interoperabilità tra diversi archivi.
Il concetto di interoperabilità assume un valore ancora più forte nell'ambito del Web semantico e della costruzione di ontologie guidate da vocabolari comuni, in questo contesto i linked data hanno un ruolo determinante a tal punto da far scrivere a Guerrini e Possemato che

I linguaggi per la trasmissione della conoscenza registrata si evolvono in continuazione sulla base della tradizione culturale e linguistica acquisita, delle caratteristiche sempre più variegate dell'universo

bibliografico, della disponibilità di tecnologie più efficaci e delle esigenze più puntuali delle persone. I linked data rappresentano l'ultima tappa di tale percorso[411].

«Il concetto di interoperabilità può essere spiegato semplicemente con l'intento di costruire una infrastruttura di comunicazione tra i sistemi di conservazione di due o più enti»[412]

E' chiaro che questa definizione, se pur corretta, non fa emergere una serie di problemi non indifferenti nella costruzione dell'infrastruttura di comunicazione. È necessario infatti che gli archivi per essere in grado di comunicare tra di loro adottino lo stesso standard o decidano di usarne un terzo come lingua franca capace di costruire il rapporto comunicativo.

Come ha giustamente sottolineato Feliciati, ad oggi il concetto stesso di standard è un concetto liquido, nel senso che l'esistenza di diversi standard rende molto complesso riuscire ad ottenere una omogeneità capace di generare infrastrutture estese, ciò non di meno l'utilizzo sempre più diffuso di OAIS consente di facilitare questo processo, come ha detto Lavoi, autore della guida all'utilizzo di OAIS

«Probabilmente il principale risultato conseguito dal modello OAIS è di essere divenuto, quasi ovunque, la lingua franca della conservazione digitale»[413]

All'interno di OAIS possiamo individuare quattro tipologie differenti di associazione tra archivi:

- Indipendent: motivati da problematiche locali
- Cooperating: archivi che hanno producer comuni e standard comuni di submission e dissemination
- Federated: insieme di archivi con una extended designated community e con strumenti di ricerca comuni
- Shared Resources: archivi che hanno sottoscritto convenzioni per condividere risorse e costi di gestione, questo richiede l'adozione di standard comuni[414]

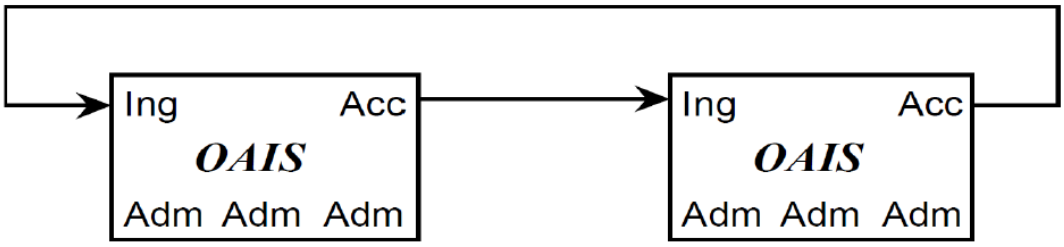
La prima tipologia identifica un archivio che utilizza il modello OAIS per servire una comunità di riferimento senza ricercare forme di collaborazione con altre strutture di conservazione, è chiaro che questo processo non coglie a pieno il nostro interesse in un ottica in cui sia possibile immaginare la strutturazione di un'infrastruttura di comunicazione capace di mettere in relazione, per esempio, tutti gli archivi regionali e l'Archivio Centrale dello Stato.

411 M. Guerrini, T. Possemato, *Lnked data per biblioteche archivi e musei*, Editrice Bibliografica, Milano 2015 p. 13

412 AGID Agenzia per l'Italia Digitale, Prgetto poli di comservazione. Definizione di un modello di riferimento per i Poli di Conservazione e della relativa rete nazionale, 2021, https://www.agid.gov.it/sites/default/files/repository_files/definizione_di_un_modello_di_riferimento_per_i_poli_di_conservazione_e_della_relativa_rete_nazionale_0.pdf ultima consultazione 23/1/2022

413 OAIS Introductory Guide (2nd Edition), https://poloarchivistico.regione.emilia-romagna.it/notizie/oais-introductory-guide-2nd-edition

414 S. Salza, *Il modello OAIS*, Centro di Eccellenza Italiano sulla Conservazione Digitale



L'architettura dei Cooperating Archives, da questo punto di vista, si mostra come più interessante, in quanto consente di rendere operabili due o più archivi OAIS attraverso la condivisione dei modelli dei pacchetti informativi di versamento SIP e di disseminazione DIP.

Con questo tipo di architettura informatica un archivio può inviare ad un altro sistema di conservazione un DIP, con la richiesta di materiale informativo e ottenere un altro DIP contenente il pacchetto cercato[415].

E' come se due archivi differenti, con interfacce utente proprie entrassero in comunicazione condividendo i pacchetti informativi relativi alla disseminazione del contenuto. L'archivio all'interno del quale avviene la ricerca da parte dell'utente entra in comunicazione con l'altro sistema OAIS che risponde fornendo le informazioni richieste.

Ancora più interessante, se ben più complessa come procedura architettonica, è quella che vede la possibilità di creare una confederazione di archivi. In questo caso viene realizzata un'interfaccia condivisa da più archivi OAIS attraverso la quale è possibile accedere ad un catalogo comune per le richieste di accesso di quella che è potenzialmente una comunità globale.

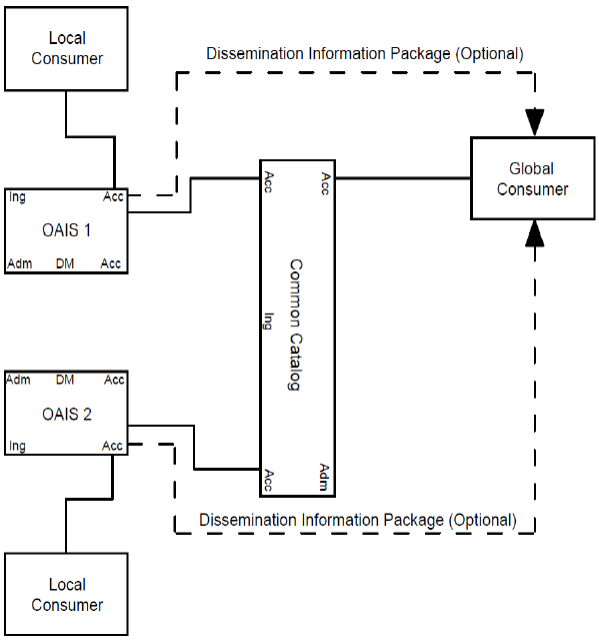
Per realizzare questa struttura è necessario che gli enti di conservazione condividano completamente metodi e linguaggi di descrizione dei pacchetti informativi.

La cosa interessante è che con la realizzazione di un catalogo comune è possibile implementare soluzioni basate su motori di ricerca semantici che risalgono al significato delle parole chiave utilizzate dalla query del ricercatore[416].

Questo implica che un utente se si interfaccia con la home page di una confederazione di archivi, il sistema, per come è costruito, consentirà di andare a ricercare in tutti gli archivi confederati i risultati relativi alla parola chiave utilizzata e in seguito a questo processo fornirà le risposte all'utente utilizzando la stessa interfaccia da cui si è partiti per la ricerca.

415 S. Pigliapoco, *Progetto archivio digitale. Metodologia, sistemi professionalità*, Civita Editoriale, Lucca, 2018 p. 167

416 Ibidem



Pur correndo il rischio di risultare ridondante non posso che rimanere affascinato nel riflette sulle implicazioni che tali modelli di gestione dell'informazione digitale possono avere nella ricerca storiografica. Un unico punto di accesso, semplici parole chiave e la possibilità di confrontarsi con un immenso patrimonio documentale senza doversi spostare dal proprio spazio di lavoro, davvero incredibili le possibilità telepistemiche date da questa struttura, e non di meno affascinante da un punto di vista epistemologico il ruolo della sovrascrittura, metadatazione, della fonte.

5.2 Il web semantico

Gli standard di archiviazione e il conseguente utilizzo di metadati sono oggi prassi consolidate nel mondo dell'archivistica e dell'archivistica informatica, la standardizzazione di specifici processi è avvenuta dopo lunghi anni di discussione e oggi è una realtà affermata.
Storici e storiche ancora fanno fatica a confrontarsi con il concetto di metadati e con le implicazioni epistemologiche che questi apportano nell'analisi della fonti. I Sistemi Informativi, se pur ormai molto strutturati e capaci di importanti azioni di interoperabilità tra le fonti, sono strumenti ancora poco usati nonostante, come spero di essere riuscito ad argomentare, si manifestino come sistemi di grande utilità per la ricerca e la contestualizzazione delle fonti.
La codifica di uno standard descrittivo con la flessibilità e la diffusione di Dublin Core, unita allo sviluppo di un sistema di gestione come OAIS, hanno *de facto* importato nel mondo della gestione documentale paradigmi concettuali ontologici per la rappresentazione dell'informazione, che sempre di più manifesta la sua essenza nell'aggregazione di dati al dato iniziale, dati che poi le macchine riescono a leggere in modo da renderli rappresentabili sia a noi che a loro.
Se già questo ci sembra, come lo è, incredibilmente innovativo e ci da indicazione del valore e delle possibilità di importare una semantica comprensibile a uomini e macchine per sviluppare un sistema di archiviazione, ordinamento e ricerca delle fonti, quello di cui proverò a parlare nelle prossime pagine, ci mostrerà come tutto ciò possa apparire già obsoleto, o meglio, superato, soprattutto nelle dinamiche di costruzione infrastrutturale.
Il Web semantico è innanzi tutto un concetto, un'idea, è la volontà di riuscire ad integrare nel Web per come lo conosciamo oggi, un livello di semantica comprensibile alle macchine, ciò vuol dire che si sta provando a far comprendere alle macchine il significato delle parole che usiamo sul Web. È bene intendere che le macchine non avranno mai la capacità di discernere ciò che è buono da ciò che è cattivo, salvo che non si sia stati noi a dare queste accezioni al concetto che la macchina sta elaborando, ma a differenza di prima, quando parlavamo di capacità metasintattiche, ciò che cambia è che grazie ad un processo all'apparenza semplice, ma in sostanza tutt'altro, utilizzando dei grafi a tripla, composti da soggetto, predicato e oggetto, le macchine saranno capaci di contestualizzare le informazioni e conseguentemente di procedere nella disambiguazione di un significato rispetto ad un altro; e ancor più interessante, saranno capaci di fornirci, in seguito ad una semplice ricerca, un numero elevatissimo di informazioni collegate in modo coerente alla richiesta primigenia che abbiamo espresso.

I processi del Web Semantico, già in atto da qualche anno ma ancor in una fase di diffusione tutto sommato molto limitata e ancora in fase di revisione, mostrano come primo determinate elemento la necessità di spostare l'attenzione dai file ai dati. Questo implica che i file intesi come elementi infomazionali univoci vengano destrutturati in atomi e ogni atomo sarà portatore di informazione.
Aver affrontato in modo abbastanza organico il significato, i principi e le funzioni dei metadati ci consente ora di provare ad osservare ciò che forse si manifesta come la più importante innovazione nei processi di funzionamento del Web, ciò che Berners Lee già nel 1999 descrisse come un sogno in cui

> Le macchine diventano capaci di analizzare tutti i dati nel Web, il contenuto, i link e le transazioni tra persone e computer. La Rete Semantica che dovrebbe renderlo possibile deve ancora nascere, ma quando l'avremo i meccanismi quotidiani di commercio, burocrazia e vita saranno gestiti da macchine che parleranno a macchine, lasciando che gli uomini pensino soltanto a fornire l'ispirazione e l'intuito[417].

Pensiamo a come Floridi ha descritto le tecnologie di terzo ordine di cui abbiamo parlato nel primo capitolo, una tecnologia che media tra altre tecnologie, questo è oggi il Web Semantico o Web dei dati. E ancora una volta, unitimene ad altre tecnologie, ciò che consente il funzionamento di quello che viene chiamato Web 3.0 è l'utilizzo di metadati.
Essendo la nostra ricerca incentrata sulle ricadute della rivoluzione digitale sul metodo di ordinamento analisi e diffusione delle fonti, il nostro punto di osservazione rispetto al Web semantico sarà legato alla possibilità di utilizzare questa rete semantica, capace di collegare tra di loro oggetti digitali e di rendere questi collegamenti comprensibili anche alle macchine attraverso l'attribuzione di significati ai dati, specificatemene per realizzare ontologie digitali all'interno degli archivi. É importante sottolineare che ad oggi l'ambizione del Web Semantico come rete semantica che si stende sopra a tutto il Web è un'idea, per quanto incredibilmente affascinante, ad oggi ancora di difficile realizzazione. Ciò che possiamo già osservare sono delle manifestazioni, realizzate in "ambienti" circoscritti, di ciò che potenzialmente un giorno potrà essere.
Il Web dei dati nasce con l'intenzione di legare ad una rete semantica tutti i contenti del Web, spostando i processi di collegamento dai link agli oggetti digitali stessi, che saranno identificati in modo univoco con delgi URI.
Nel 2001 Berners Lee propose un'estensione del Web che avrebbe permesso alle persone di sfruttare i computer per trattare le informazioni in una modalità più efficace. Chiamò la soluzione Web semantico, perché questo web avrebbe aggiunto un'infrastruttura con significato concettuale

417 T. Berners Lee, *L'architettura del nuovo Web: dall'inventore della rete il progetto di una comunicazione democratica, interattiva e intercreativa*, Feltrinelli, Milano 2001, p. 139

definito, la semantica, al di sopra di un'infrastruttura sintattica di HTML per il trattamento automatico dei documenti nel web[418].

In seguito, la soluzione proposta da Tim Berners-Lee nel 1994 fu quella di aggiungere valore semantico alle pagine web, in modo che il loro significato fosse in qualche misura accessibile anche dai software; l'idea discendeva direttamente, come ha evidenziato Floridi, «dal sogno di Leibniz di disegnare una *lingua characteristica* (un linguaggio nel quale tutta la conoscenza potesse essere formulata inequivocabilmente) e un *calculu ratiocinator* (un calcolo di ragionamento, un motore semantico inferenziale) cosi da migliorare la comunicazione e risolvere più facilmente le discordanze linguistiche»[419].

Il gruppo di ricerca sul Web semantico del World Wide Web Consortium definisce un framework comune in cui condividere e ritualizzare i dati, superando i confini tra applicazioni, enti e comunità differenti.

Il web semantico diventa un movimento di collaborazione, guidato dal World Wide Web Consortium, organismo internazionale e di normalizzazione. Il W3C sviluppa tecnologie che garantiscono l'interoperabilità per portare il World Wide Web al massimo del suo potenziale, agendo da forum di informazioni, comunicazioni e attività comuni.

Il Web Semantico si propone di convertire il web attuale caratterizzato da documenti non strutturati e semi strutturati, in un web di dati, incoraggiando l'inserimento di contenuto semantico nelle pagine web. L'idea di Web semantico può essere affrontata anzitutto come un fenomeno linguistico: il metodo che permette l'integrazione coerente di dati diversi. Esso si presenta come una lingua per i dati, una lingua inventata dagli umani per comunicare informazioni e pensieri fondamentalmente umani e per l'uso umano, una lingua tuttavia, concepita per essere letta ed elaborata non da lettori ascoltatori umani, bensì da computer secondo algoritmi ben definiti.[420]

Ancora una volta perché si riesca a dare forma a questi processi di innovazione nella gestione e nell'interazione con le informazioni digitali, in questo caso a tal punto da renderle comprensibili per le macchine, si assiste ad un lavoro di sovrascrittura.

Ferraris in *Documentalità* afferma controintuitivamente, che ancora oggi è la scrittura, se ben composta da linguaggi sempre nuovi, a rendere possibile i processi informativi. Nulla di ciò che abbiamo osservato nelle pagine precedenti esiste al di fuori della scrittura e questo non può che

418 M. Willer, G. Dunsire, Bibliographic information organization in the semantic web. Oxford, Cambridghe; New Delhi: Chandos Publishing, 2013

419 L. Floridi, *Web 2.0 vs. the Semantic Web. A Philosophical Assessment*, in «Episteme», 6 (2009), n. 1, doi: 10.3366/E174236000800052X , p. 30

420 M. Guerrini, T. Possemato, *Linked data per biblioteche archivi e musei,* Editrice Bibliografica, Milano 2015 p. 24

richiamare le nostre menti al significato di «oggetto sociale» per come lo abbiamo affrontato nel secondo capitolo.
Se nulla di sociale esiste fuori dal testo e se in seguito alla digitalizzazione ogni azione è registrazione, è chiaro che le aspirazioni del Web semantico portano a voler attribuire significato ad ognuno di questi atti di registrazione, sviluppando un processo di atomizzazione che sposta il significato dal documento al dato e per questo lo rende comprensibile alle macchine.
Ed è proprio grazie all'atomizzazione del dato che sta alla base del funzionamento del Web dei dati, che le pagine HTML, che hanno caratterizzato il Web fino ad oggi, evolvono a favore dei dati singoli, ai quali, grazie alle ontologie e ai vocabolari controllati viene attribuito significato.
Per provare a comprendere meglio ciò di ci stiamo parlando proviamo a tornare alla figura relativa alla costruzione dell'oggetto informativo in OAIS, lì possiamo notare come tutto parta dal dato, che attraverso un processo di metadatazione prima e di rappresentazione e interpretazione dell'informazione poi, diviene un oggetto informativo.
Vorrei rapidamente tornare all'ottavo paragrafo del primo capitolo, «da dato a informazione. Da evento a fonte». In quelle righe abbiamo osservato la teorizzazione della piramide dell'informatizzane, vedendo come alla base ci fosse il dato, che avevamo definito come una entità pura, non interpretata, e che da una prospettiva diaforica è qualcosa che esiste nel manifestarsi come differenza. Al livello subito superiore del dato troviamo l'informazione. L'informazione non si realizza solo attraverso una sovrapposizione di dati, ma è necessario che questi assumano del significato che possa essere compreso; ed è proprio nella fase di rappresentazione e interpretazione, osservabile nello schema di OAIS, che si trovano gli elementi capaci di trasformare il dato in informazione, sono i significati che attribuiamo a quel dato che ci permettono di elaborarlo in modo intelleggibile.
Se vogliamo contestualizzare all'interno delle riflessioni sul Web semantico l'intuizione di Topolski, secondo cui la produzione di una fonte è rappresentazione di un evento storico, dobbiamo fare lo sforzo di decostruire il concetto di fonte come struttura unitaria e volgere il nostro sguardo verso i dati che la compongono e ricercare nella produzione del dato l'evento stesso. Questo ragionamento mi pare che non si discosti dalle riflessioni di Ferraris rispetto all'importanza dell'iscrizione come azione propedeutica a quella informativa e dunque come elemento determinante per comprendere i processi di interazione digitale. Il primo atto dell'iscrizione è indubbiamente la produzione di un dato, a prescindere dalla complessità informativa che si sta generando.
Ciò che a noi interessa, da una prospettiva che vede nella dematerializzazione, metadatazione e risemantizzazione dei dati una possibilità per creare nuove forme di storiografia, è provare a

comprendere quali sono i sistemi che consentono di costruire una relazione tra queste fonti che giustamente vengo definite «fonti di dati»[421].

Il Web semantico è composto dalle seguenti proprietà:

- può contenere qualsiasi tipo di dato
- i dati possono essere pubblicati da chiunque
- il dato è autoconsistente, cioè porta in se la propria descrizione: nel caso un'applicazione incontrasse dati descritti con un vocabolario sconosciuto, può deferenziare l'URI (dunque estrapolarlo dal un processo referenziale tipico dei database tradizionali) che identifica ciascun termine del vocabolario e arrivare alla sua definizione.
- Le entità collegate da link RDF (Resource description Framework)[422] creano un grafo globale che si estende e ingloba fonti di dati differenti e consente alle applicazioni la scoperta di nuove fonti informative.
- Chi pubblica i dati nel Web non è obbligato ad utilizzare specifici vocabolari per rappresentare i propri dati
- Il Web Semantico è aperto: le applicazioni possono esplorare e scoprire nuovi dati seguendo i link del reticolo.

Come riportato da Guerrin e Possemato questo elenco di proprietà sottintende una serie di concetti che stanno alla base del funzionamento del Semantic Web.

- La correttezza formale e sintattica di una frase pubblicata nel Web non implica un'equivalente correttezza semantica: le asserzioni possono essere formalmente corrette e, quindi valide, ma non necessariamente vere; per esempio, l'asserzione "Immanuel kant era italiano" è formalmente corretta, ma non è vera.
- Il concetto di dato autoesplicante, che rimanda al concetto di URI dereferenziabili: l'oggetto identificato con l'URI porta con se una rappresentazione di se stesso e, dunque, diventa compressibile all'agente che lo incontra nel proprio percorso di esplorazione
- L'esplorazione del Web semantico difficilmente nasce e muore all'interno del medesimo insieme di dati strutturati (dataset): i dati pubblicati in modalità aperta e collegati con relazioni qualificate creano un reticolo che consente di utilizzarli in modo inatteso, non previsto all'inizio dell'esperienza di ricerca, e consentono quei processi inferenziali che costituiscono il valore aggiunto del Web semantico. Ciò vuol dire che a partire da un'asserzione è possibile dedurre nuova informazione, senza che essa sia espressamente dichiarata dal dato[423].

421 Ivi, p. 27
422 https://www.w3.org/RDF/ ultima consultzione 22/1/2022
423 M. Guerrini, T. Possemato, *Linked data per biblioteche archivi e musei,* Editrice Bibliografica, Milano 2015 p. 29

Nel Web Semantico i linked data sono il mezzo, la struttura, attraverso la quale è possibile costruire una rete di connessioni; utilizzando dati strutturati ed un insieme di regole di inferenza, si rende possibile ad una macchina di compiere "ragionamenti automatici".

Il concetto di inferenza è importante per capire cosa si intende quando si fa riferimento alla capacità di una macchina di compiere ragionamenti automatici. Con inferenza si intende il processo con il quale si passa da una proposizione accettata come vera ad una seconda la cui verità è dedotta grazie al contenuto della prima.

Facciamo un esempio, date queste 3 asserzioni:

Marco è figlio di Luigi

Bruna è figlia di Luigi

Luigi è figlio di Francesca

Possiamo dedurre che:

Marco e Bruna sono nipoti di Francesca

Marco è fratello di Bruna

Bruna è sorella di Marco

In pratica avendo i dati e delle regole ben definite è possibile, per una macchina, dedurre nuova conoscenza, cioè è possibile ricavare informazioni che non sono state direttamente inserite nel sistema.

Ai dati disseminati in documenti web come semplice testo e quindi non elaborabili se non con tecniche di information retrieval, possiamo dare una struttura, possono essere cioè associati a dei metadati descrittivi. In questo caso i metadati descrittivi strutturano il testo in qualche modo analogalmente a ciò che avviene in una tabella del modello relazionale dei data base, dove il nome della colonna attributo esplicita il significato del valore in una riga. In realtà l'effetto è molto maggiore poiché tali metadati associati a linguaggi logici consentono di definire concetti e relazioni tra i concetti associati, ad esempio di equivalenza: donna con almeno un figlio equivale a madre. Inoltre identificando concetti e relazioni mediante URI, si facilita la lettura e la condivisione dei dati tra computer ma anche l'elaborazione automatica da parte di sistemi di ragionamento o agenti dotati di qualche forma di autonomia e pertanto è possibile delegare alla macchina l' esecuzione di compiti complessi[424], di natura non unicamente quantitativa.

Ciò che abbiamo imparato a conoscere come oggetti digitali e che con DCMI divengono oggetti informativi, all'interno del Web semantico si manifestano come atomi informativi che divengono

424 T. Di Noia, R. De Virgilio, E. Sciascio, Francesco M. Donini, *Semantic Web, tra ontologie e open data*, Maggioli Editore, Santarcangelo di Romagna, 2018 p. 17

intellegibili grazie a metadati descrittivi che riescono ad entrare i relazione tra di loro e ad assumere significato attraverso l'utilizzo di concetti.
Nel Web semantico i collegamenti tra i dati hanno un loro significato specifico, non sono solo link. Ancora una volta credo che questo non possa che far riflettere storiche e storici e chiedersi quale il loro possibile ruolo in questo processo di decostruzione dell'informazione in un'ottica di realizzazione di una rete semantica capace di mettere in connessione ogni singola parte. Da canto mio non ho risposte, ma penso che anche noi, in quanto scienziati umanisti coinvolti nella costruzione del discorso sociale, non si possa stare fermi a guardare, è necessario che si provi ad entrare nel dibattito e ci si confronti sempre di più con informatici, linguisti, scienziati dell'informazione e della comunicazione, semiologi, archivisti.
In quanto public historian, l'esperienza sul campo nella realizzazione di prodotti di Public History anche molto diversi tra loro e spesso coprogettati dal basso con il coinvolgimento di piccole e medie comunità, mi ha portato a capire che utilizzare concetti trasversali - che hanno dunque un valore nel presente e nel passato e che presumibilmente lo avranno anche nel futuro - per clusterizzare differenti fonti da collegare attraverso significati specifici è un modo altamente funzionale per l'analisi e la diffusione del contenuto storico. Non sto proponendo nessun tipo di soluzione specifica, ma difronte ai cambiamenti in atto è necessario provare a comprendere come immaginiamo di sviluppare la ricerca in un futuro prossimo e forse ancora più necessario comprendere come già oggi, si iniziano a creare e a organizzare le informazioni che domani saranno fonti.
Il 27 luglio 2006 Berners-Lee pubblica un documento, dove descrive così i criteri per realizzare il Web Semantico:

- Utilizzare URI per identificare e dare un nome alle cose
- Utilizzare URI seguendo il protocollo HTTP affinché possano essere consultati
- Fornire informazioni utili quando qualcuno cerca le URI utilizzando standard come RDF e SPARQL
- Includere collegamenti ad altre URI per permettere di trovare più risorse[425]

Per descrivere le informazioni in maniera strutturata all'interno del Web dei dati, è stato sviluppato un modello di codifica e rappresentazione dei dati e concetti astratti, RDF (*Resource Description Framework*), che traduce il dato in asserzioni logiche composte da soggetto, predicato e oggetto, le

425 https://www.w3.org/DesignIssues/LinkedData.html

cosiddette triple. Il funzionamento di RDF si basa sull'associare ogni elemento ad un URI (autodescrittivo) «in modo che chiunque vi si possa riferire o reperire una rappresentazione»[426].
RDF attualmente rappresenta lo standard proposto dal W3C per la codifica, lo scambio e il riutilizzo di metadati strutturati, esso inquadra le strutture di dati secondo tassonomie gerarchiche rendendo possibile l'esecuzione di procedure inferenziali. In questo modo, si consentirebbe ai calcolatori di utilizzare i dati per eseguire ragionamenti automatici, le inferenze, e si porrebbero le basi per una sorta di indicizzazione concettuale, semantica, per l'appunto.
I linked data utilizzano RDF perché tipizza e, dunque, categorizza meglio di altri linguaggi i legami che intercorrono tra le risorse.
RDF prende in considerazioni solo relazioni binarie, le più semplici, nelle quali un soggetto e un oggetto sono messi in una certa relazione tra loro. Questa relazione è anche detta predicato, in quanto dal punto di vista linguistico, spesso la relazione coinvolge un verbo, e dal punto di vista logico, la relazione si può vedere anche come un predicato a due argomenti. La sequenza soggetto predicato oggetto viene denominata tripla.

RDF si pone quindi a cavallo tra le discipline della linguistica, della logica e dell'informatica. Intuitivamente, dalla linguistica trae le forme delle espressioni, dalla logica una loro semantica formale, e dall'informatica una rappresentazione suscettibile di essere elaborata automaticamente.[427]

Soggetto, predicato e oggetto sono a loro volta degli URI, questo implica che hanno altre relazioni e sono connessi ad altre informazioni.

In una base di conoscenza (costituita dal livello intenzionale e dal livello estensionale dell'ontologia considerati unitariamente) non ci si trova più di fronte un oggetto immediatamente isolabile dal contesto informativo, ma nella descrizione della risorsa trova posto il dato oggetto più tutti gli "oggetti" (in senso lato) con cui questo risulta in relazione.[428]

L'utilizzo di triple consente dunque alla macchina di essere in grado di contestualizzare l'informazione, non si tratta più solo di di leggere in modo metasintattico il termine utilizzato per la ricerca, ma appunto di inserirlo in un contesto che lo carichi di significato.

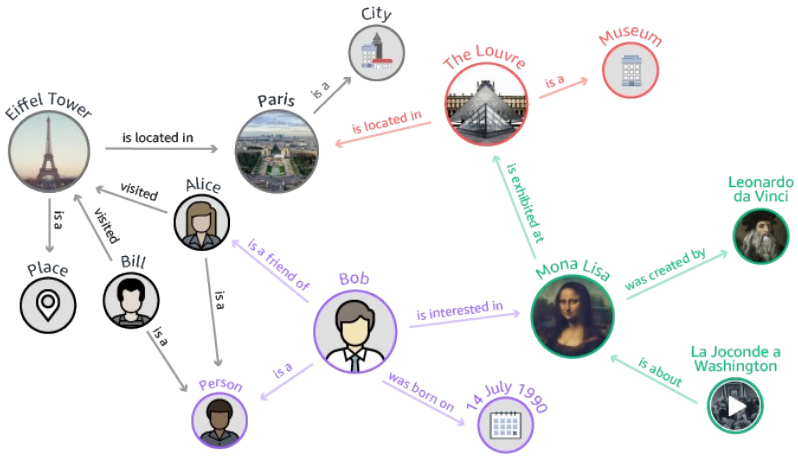
426 K. Weller, *Knowledge Rapresentation in the Social Semantic Web*, De Gruyter Saur, Berlin-New York 2010 p. 60
427 T. Di Noia, R. De Virgilio, E. Sciascio, Francesco M. Donini, *Semantic Web, tra ontologie e open data*, Maggioli Editore, Santarcangelo di Romagna, 2018 pp. 29,30
428 S. Di Fazio, *Le ontologie,* in G. di Marcantonio, F. Valacchi, a cura di, *Descrivere gli archivi al tempo di RIC-CM,* EUM, Macerata 2018 p. 92

Le triple di RDF formano ragnatele di informazioni su cose che sono in relazione. Poiché RDF utilizza queste informazioni per codificare le informazioni in un documento, l'URI assicura che questi concetti non siano solo parole in un documento ma che siano legati ad una definizione unica che chiunque può trovare sul Web.

Facciamo un esempio, se noi creiamo una tripla come: Mussolini ha fondato il fascismo, con Mussolini soggetto, ha fondato predicato, e fascismo, oggetto, la macchina grazie alla rete semantica consentita dai linked data e da RDF leggerà contestualmente la tripla "comprendendone" il significato e restituendoci, come nel caso di Dbpedia[429], l'accesso alle informazioni che stanno in relazione con il soggetto e l'oggetto, è chiaro che questo richieda però anche un lavoro di disambiguazione dei termini e chiaramente un attento lavoro di metadatazione. L'utilizzo di una tripla come quella proposta nell'esempio fa sì che il soggetto Mussolini rimarrà legato al predicato e all'oggetto, che a loro volta però saranno legati ad altri "significati", dunque facendo una ricerca semplice, partendo dal soggetto/dato "Mussolini" il sistema ci restituirà collegamenti con altri predicati e altri oggetti come per esempio "ha istituito" e "le leggi razziali". L'interoperabilità delle fonti farà sì che se presenti all'interno del Web come fonti digitalizzate, o presenti all'interno dell'archivio digitale, il sistema restituisca tutte le fonti collegate ad uno degli URI facenti parte della nostra tripla, dunque per esempio, in relazione a "le leggi razziali" potremmo avere accesso al medium-documento della legge, così come, sempre per fare un esempio, alla rivista «La difesa della razza».

Questi dati sono inseriti in una rete semantica che sarà capace di restituire tutti i possibili significati legati agli URI del soggetto, del predicato o dell'oggetto.

Questo un schema che sintetizza il funzionamento di triple e RDF



429 DBpedia è un progetto nato nel 2007 con lo scopo di estrarre informazioni strutturate da Wikipedia e pubblicarle sul Web come Liinked Open Data in formato RDF. https://www.dbpedia.org/

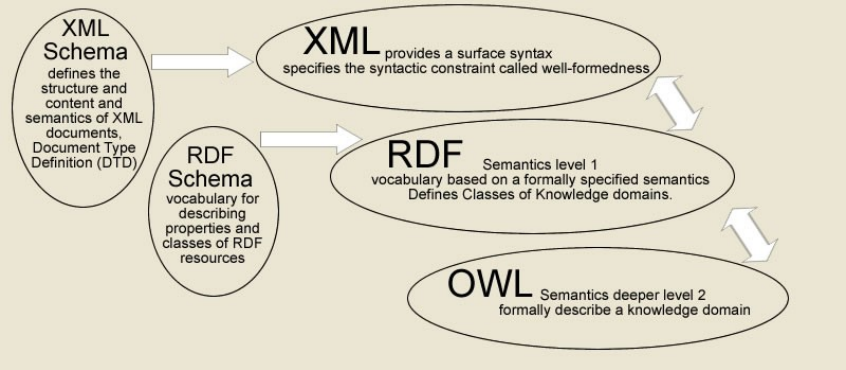
Da questa immagina si comprende bene il funzionamento delle relazioni all'interno di una tripla e di come ogni dato, soggetto, predicato e oggetto, sia collegato ad altri dati che è possibile navigare grazie a delle relazioni di natura semantica. Ricordiamoci che uno dei principi del Web semantico è che è valido con qualsiasi tipo di dato.
Per comprendere meglio ciò di cui stiamo parlando affronteremo, in modo sintetico, il funzionamento dell'RDF Schema

RDF Schema OWL e ontologie.

XML nella visione del Web semantico sancirebbe l'ingresso della semantica nel Web. Questo linguaggio permette di aggiungere contenuto, non solo *machine readable,* interpretabile in maniera differente a seconda di come viene istruito il sistema che lo interpreta. Per far ciò sarebbe sufficiente dichiarare l'esistenza di blocchi semantici minimi, i metadati, e definire una grammatica con delle regole sintattiche volta ad assicurare, proprio attraverso i metadati, la correttezza formale nell'uso degli oggetti costruiti.
Uno dei vantaggi dell'utilizzo di XML come (meta)linguaggio di *markup* universale risiede nella possibilità di accedere a informazioni provenienti da più fonti; tecnicamente infatti, un documento XML può utilizzare anche più di una grammatica. Tale caratteristica rappresenta sicuramente un grande vantaggio, ma, proprio in virtù di ciò essa porta con sé anche una serie di inconvenienti, uno tra tutti, il problema della disambiguazione del significato o polisemia. I documenti strutturati, nel momento in cui vengono confrontati o integrati, creano inevitabilmente conflitti tra termini. Ciò dipende essenzialmente dal fatto di essere sviluppati in maniera indipendente rispetto alla grammatica. Il problema che si viene a creare risiederebbe nella necessità di trovare un modo per identificare esattamente l'ambito di ciascun elemento, conciliare la presenza di elementi definiti in uno di più vocabolari, e soprattutto conciliare la presenza di elementi definiti con lo stesso nome in più vocabolari diversi.
XML fornisce un metodo per identificare in maniera non ambigua la semantica del dato, identificando anche l'Authority che gestisce il vocabolario. Una delle Autorithiy più usate è Dublin Core.
Si può quindi superare il problema facendo sì che XML punti alla risorsa Dublin Core che ne definisce la semantica.
Una volta attuato questo processo di disambiguazione l'RDF Data Model permette di definire un modello semplice per descrivere le relazioni tra le risorse.

RDF Schema permette di definire dei vocabolari, quindi l'insieme delle proprietà semantiche individuate da una particolare comunità. RDF Schema permette di definire significato, caratteristiche e relazioni di un insieme di proprietà, compresi eventuali vincoli sul dominio e sui valori delle singole proprietà. Inoltre, implementando il concetto (transitivo) di classe e sottoclasse, consente di definire gerarchie di classi, con il conseguente vantaggio che gli agenti software possono utilizzare queste relazioni per svolgere i loro compiti[430].

Sia il contenuto che gli assiomi, le regole logiche, sono pertanto nodi e archi che compongono il grafo, che insieme descrivono e si autodescrivono facendo perno sul meccanismo della tripla. La grande scommessa del W3C è stata di ipotizzare che attraverso la combinazione di triple, si potesse riuscire a dire qualunque cosa di qualsiasi cosa.
Un grafo è definibile come un insieme costituito da nodi e da connessioni (archi), e nell'ambito della teoria informatica costituisce, come anche l'albero e il nodo stesso, una delle possibili forme di rappresentazione della conoscenza.
Perché tutto ciò prenda forma in modo ordinato c'è bisogno di utilizzare dei linguaggi capaci di creare delle ontologie, dunque di leggere e comprendere in modo automatizzato le informazioni legate ad RDF.



431

OWL (Ontology Web Language) è un linguaggio utilizzato nel Web Semantico, in cui la tripla continua a giocare il proprio ruolo cardine, ma attraverso OWL è possibile esprimere delle vere e proprie teorie logiche che contemplano congiuntamente dati e norme che regolano i dati, dando vita a documenti formali, le ontologie, che costituiscono al contempo strumenti di rappresentazione di

430 A. Dorati, S. Costantini, *Approcci al Web semantico*, http://www.websemantico.org/articoli/approcciwebsemantico.php#inizio ultima consultazione 22/1/2022
431 L'immagine è stata presa da https://instructionaldesign.com.au/the-semantic-web/ ultima consultazione 12/2/2022

una data conoscenza e l'insieme delle regole che rendono possibile dedurre – tramite ragionamento automatico – da tale rappresentazione, ulteriore conoscenza.
Nel 2014, l'Agenzia per l'Italia Digitale (AgID) ha fornito una definizione di ontologia informatica particolarmente apprezzabile per chiarezza ed essenzialità; secondo questa definizione, «una ontologia è una rappresentazione formale e condivisa dei concetti e delle mutue relazioni che caratterizzano un certo dominio di conoscenza»[432]

OWL, è costruito su una logica che concatena particelle elementari di informazione, frammenti di conoscenza formalmente espressa, che si uniscono a catena irradiandosi e propagandosi in molteplici direzioni, generando quello che viene definito come grafo di conoscenza.

La tripla costituisce dunque una sorta di grado zero del linguaggio, che ne individua con grande immediatezza gli elementi essenziali (quasi esclusivamente ricorrendo a URI), e il ruolo che ciascuno di tali elementi gioca in essa. Una risorsa soggetto viene messa in relazione attraverso un predicato (graficamente corrispondente ad un arco orientato ed etichettato) con un complemento, che può essere un'ulteriore risorsa o un valore testuale[433]. La tripla è quindi il cardine di una rappresentazione semantica, ma costituisce di per sé una informazione frammentaria che va concettualizzata e contestualizzata.
Ogni risorsa informativa descritta da una ontologia è pienamente autoconsistente, radunando,
"a grappolo" intorno a sé, tutta l'informazione che la descrive complessivamente (relazioni in partenza) ma risultando in connessione anche con l'informazione alla cui rappresentazione partecipa con la funzione di complemento (relazioni in arrivo).
Si tratta di una visione oggetto-centrica della descrizione che si attua riportando sulla singola risorsa – come attributo esplicito veicolato da specifiche asserzioni – tutte le caratteristiche che ad essa sono riferibili, sia che le siano proprie (caratteristiche di *domain*) sia che le derivino dai vari contesti a cui risulta correlata. Cambia in un certo senso la prospettiva con cui si guarda all'oggetto della descrizione e se ne definiscono le caratteristiche, e in questo modo si modifica sensibilmente anche il concetto stesso di risorsa.

L'autoconsistenza dell'informazione consentita dal ricorso ad un formalismo come OWL si estende peraltro anche agli aspetti più prettamente tecnologici della *digital preservation*. Il modello a grafo consente infatti di

432 L'AgID pubblica ogni anno una versione aggiornata delle proprie linee guida. Agenzia per l'Italia Digitale, «Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico», 2014, 14, <http://www.agid.gov.it/sites/default/files/linee_guida/patrimoniopubblicolg2014_v0.7finale.pdf>.
433 S. Di Fazio, *Le ontologie,* in G. di Marcantonio, F. Valacchi, a cura di, *Descrivere gli archivi al tempo di RIC-CM,* EUM, Macerata 2018 p. 86

riportare sulla singola risorsa tutta l'informazione ad essa pertinente, compresa quella relativa ai diversi *Environment* di produzione, gestione e fruizione. L'adozione di un simile ambiente di formalizzazione permette quindi di produrre Information Packages (IP, secondo la terminologia resa ormai di riferimento per il dominio dal modello OAIS) realmente svincolati e indipendenti dalla tecnologia e dagli ambienti software di provenienza, e quindi al riparo dalla conseguente obsolescenza dei formati e dei sistemi di validazione esterni al sistema di conservazione[434].

Se XML e RDF si occupano di garantire un ordine, una struttura per le singole informazioni, le ontologie agiscono invece a un livello più alto. Esse si occupano di ordinare insiemi di dati strutturati. Un'ontologia permette di specificare i concetti e le relazioni che caratterizzano un certo dominio di conoscenza. Un dominio di conoscenza altro non è che un insieme, coeso e organico rispetto a un certo concetto, di dati singolarmente strutturati.
Le ontologie, allora, seppur a un altro livello rispetto ai vocabolari RDF, svolgono grossomodo la stessa funzione di questi ultimi: strutturano le informazioni.
Secondo Di Fazio, in una ontologia OWL vuota potrebbe essere rappresentata l'intera realtà: da un punto di partenza dove tutto è possibile, si procede nella compilazione in modo interattivo, introducendo progressivamente vincoli, e stabilendo esplicitamente cosa non è possibile e cosa è escluso dalla realtà che si vuole rappresentare, se ben ricordiamo anche Michetti parlando del funzionamento di OAIS in relazione ai processi di interpretazione dell'informazione attraverso una concatenazione ricorsiva di metadati, disse che si potrebbe arrivare a costruire «un'iterazione che condurrebbe paradossalmente alla conoscenza del mondo»[435]

L'ipotesi di mondo aperto (OWA, che è tipica delle ontologie e del linguaggio OWL) si oppone all'ipotesi di mondo chiuso (CWA) tipica invece dei database relazionali. In un sistema OWA l'incompletezza della base di conoscenza di riferimento è assunta come postulato, per cui ciò che non è incluso in essa è semplicemente "non conosciuto"; al contrario, nei sistemi CWA, tutto ciò che non è verificabile come vero attraverso i dati in esso inclusi, è falso[436].

Questa insistenza sul dato autoconsistente e autoesplicativo che grazie ad una rete semantica viene messo in relazione con altri dati autoconsistenti e autointerpretati, non può non farci riflettere riguardo a come sia stata strutturata l'informazione all'interno degli archivi digitali per come l'abbiamo descritta nei paragrafi precedenti.

434 Ivi, p. 93
435 G. Michetti, *Il modello OAIS*, in «Digitalia», Anno III, n° 1, 2008 p. 42
436 H.Sack, *Knowledge Engineering with Semantic Web Technologies*, Lecture 3: Ontologies and Logic. 3.9 DLs and the Open World Assumption, 2015, https://open.hpi.de/files/8e5a3cf1-09e6-4c26-9afa-c1e5b967e11f ultima consultazione 4/2/2022

A che pro utilizzare uno standard come ISAD(G) se i sistemi di interazione tra le informazioni non avvengono più in modo gerarchico, ma in modo ramificato e le relazioni tra le unità informative non son più multilivellari, ma multidimensionali. Non vi è più la necessità di costruire un contesto articolato per come lo abbiamo imparato ad osservare negli standard; sono i dati legati alle triple che trovano di volta in volta nuovi e espansi sistemi di contestualizzazione.

Gli archivi, per come abbiamo più volte ripetuto, hanno bisogno di un ordinamento preciso per essere fruiti, e sebbene la digitalizzazione avesse già imposto un ordinamento logico e non più fisico, pare che questo concetto scompaia a favore di processi semantici atti a collegare tra di loro informazioni completamente "libere" all'interno del Web.

Durante la IV Conferenza nazionale dell'Associazione Italiana di Public History, all'interno del panel *«Copy, transform, combine»: Gli archivi digitali e il loro uso partecipato*[437] coordinato da Lorenzo Pezzica e che vedeva come relatori oltre a me, Enrica Salvatori dell'Università di Pisa e Vittorio Iervese dell'Università di Modena e Reggio Emilia, si è aperto un interessante dibattito con il pubblico presente, tra le molte cose che sono state dette mi hanno particolarmente colpito le parole di Pierluigi Feliciati, che in seguito al mio intervento dal titolo *Oggetti digitali. La sovrascrizione delle fonti per una nuova accessibilità*, ha sottolineato come ISAD, per quanto standard fondamentale nello strutturare dei processi di omologazione dell'indicazione archivistica, sia oggi uno strumento da considerarsi novecentesco, non in grado, con la sua rigidità e la sua struttura multilivellare, di rispondere alle reali necessità di organizzazione dell'informazione digitale. È nel dato autocostituente e nella possibilità di organizzare l'informazione in un senso multidimensionale che anche Feliciati individua un processo di innovazione capace di modificare l'interazione tra i dati, l'informazione e i possibili fruitori.

Per provare a dare alcuni spunti che possano aiutarci a individuare delle possibili risposte è di grandissimo interesse il lavoro svolto per la redazione del testo *Descrivere gli archivi al tempo di RIC-CM*, a cura di Giorgia di Marcantonio e Federico Valacchi.

RiC-CM è un modello concettuale di descrizione degli archivi, elaborato dall'Expert Group on Archival Description (EGAD) dell'International Council on Archives (ICA), che si propone di integrare gli standard descrittivi esistenti, aggiornandone la prospettiva sulla base degli sviluppi che hanno caratterizzato negli ultimi tempi le riflessioni e la pratica archivistica e fornendo, al tempo stesso, i presupposti per un'applicazione al mondo archivistico delle tecnologie del web semantico.[438]

437 https://aiph.hypotheses.org/10695 ultima consultazione 13/07/2022

438 RiC-CM (Records in Contexts - A Conceptual Model for Archival Description) (beniculturali.it) ultima consultazione 4/2/2022

Con il modello concettuale RIC-CM sarà possibile restituire una descrizione dell'archivio multidimensionale. In altre parole non si andranno ad identificare ed esplicitare solo le singole entità descrittive ma anche le sue proprietà e relazioni cosicché si possa disegnare una vera e propria mappa dell'archivio[439].

L'avvento del Web semantico ha messo in "allerta" parte del mondo archivistico che ha visto in questa atomizzazione dell'informazione, con l'identificazione attraverso URI di singoli dati, il rischio che l'informazione nel Web diventasse ancora più liquida e difficilmente organizzabile all'interno di processi di ordinamento.
Questi timori mi trovano in parte in accordo, soprattutto considerando quello che più volte ho ribadito nel testo, dunque la possibilità di osservare i sistemi informativi e gli archivi digitali, come strumenti per arginare i processi apparentemente inarrestabili della misinfomation, garantire dei porti sicuri in cui far approdare chiunque in modo sempre più semplice e diretto.
L'esistenza del SAN, solo per citare uno dei diversi strumenti che abbiamo imparato a conoscere, garantisce una base di accuratezza delle informazioni in esso archiviate, non solo perché frutto di un'analisi preventiva da parte di archivisti, ma perché inserite in contesti standardizzati che offrono l'opportunità di trovare risorse collegate all'informazione che si sta cercando, consentendo così anche un'analisi comparativa ordinata a partire da un unico punto di accesso ad un enorme patrimonio documentale.
L'atomizzazione dell'informazione in qualche modo pare non consentire questo processo, ma, grazie a RIC-CM è possibile costruire una costellazione con gli standard già in utilizzo.
Perché ciò sia possibili Alfier ci invita ad osservare gli standard sempre e di nuovo come modelli concettuali che è possibile modificare, ampliare e come sapiamo rendere interoperabili. L'errore sta nel pensare che gli standard debbano essere superati con l'introduzione del Web semantico, quando in realtà quello che deve essere superato è immaginare un unico modo di costruire la comunicazione di questi standard.

A voler essere onesti poi, come già sottolineato in precedenza, il monolite di ISAD non è tanto nello standard in sé, ma in come lo si è applicato. Infatti a partire da indicazioni su quali punti siano essenziali in una descrizione archivistica lo si è trasformato in precetti da seguire, vincolati in campi di database che devono essere necessariamente quelli e che addirittura debbono essere visualizzati all'utente in quel modo. Il problema quindi non è tanto ISAD quanto gli isadiani verrebbe da dire. Il passaggio a un reticolo descrittivo

439 G. di Marcantonio, F. Valacchi, *Descrivere gli archivi al tempo di RIC-CM*, EUM, Macerata 2018 p. 9

che possa permettere la coesistenza di visioni (contesti appunto) diverse che si integrano fino a formare una tela di relazioni e conoscenza.[440]

Dello stesso avviso è anche Stella Di Fazio, che scrive

Si è parlato apertamente, da quel momento in poi, di «crisi del modello gerarchico di rappresentazione della struttura degli archivi», e della opportunità di allineare i modi e le forme della rappresentazione della descrizione archivistica a nuovi paradigmi, concettuali e tecnologici, a cui affidare l'incombenza teorica e operativa non tanto di sostituire quel modello tradizionale, ma più convenientemente di affiancarlo, potenziandolo e accrescendone la portata informativa attraverso il riconoscimento e l'esplicitazione della natura pluridimensionale delle entità in esso presenti.[441]

Dobbiamo considerare che con lo sviluppo sempre più avanzato del concetto di oggetto informativo, oggi si fa riferimento ad una serie di informazioni incapsulate nella fonte.
E' come se si stesse osservando un processo in cui gli archivi, da essere i contenitori delle fonti, attraverso i metadati, diventino contenuti nelle fonti.
Si tratta infatti di un trapasso per nulla ininfluente e che rimanda alla soluzione tecnica dell'incapsulamento. Un unico e coerente oggetto digitale con all'interno un nucleo, il file o i file che compongono il contenuto rappresentativo del documento digitale, e attorno ad esso un'accumulazione progressiva e strutturata di strati informativi, i metadati, che raccontano, tracciandola, la "storia" digitale del documento stesso, il suo passaggio tra molteplici ambiti applicativi e contesti d'uso.
I metadati divengono sempre di più gli elementi strutturali dell'organizzazione dell'informazione e contemporaneamente alfieri della dinamicità del documento per come si produce e archivia all'interno delle società iperstoriche.

Si tratta di una visione oggetto-centrica della descrizione che si attua riportando sulla singola risorsa – come attributo esplicito veicolato da specifiche asserzioni – tutte le caratteristiche che ad essa sono riferibili, sia che le siano proprie (caratteristiche di dominio) sia che le derivino dai vari contesti a cui risulta correlata. Cambia in un certo senso la prospettiva con cui si guarda all'oggetto della descrizione e se ne definiscono le caratteristiche, e in questo modo si modifica sensibilmente anche il concetto stesso di risorsa[442]

440 S. Vassallo, *La descrizione archivistica oggi*, in G. di Marcantonio, F. Valacchi, a cura di, *Descrivere gli archivi al tempo di RIC-CM*, EUM, Macerata 2018 p. 42
441 S. Di Fazio, *Le ontologie*, in G. di Marcantonio, F. Valacchi, a cura di, *Descrivere gli archivi al tempo di RIC-CM*, EUM, Macerata 2018 p. 82
442 Ivi, pp. 92, 93

Con i Metadati la fonte non solo si manifesta al fruitore ma consegna a questo ogni informazione sul dato, rendendo possibile un'interazione incredibilmente ramificata tra informazioni che garantiscono per se stesse.
In un universo non più ipotetico in cui le fonti non solo sono rappresentazione di se stesse, ma di ogni informazione a loro correlata, Alfier pone una questione davvero interessante che credo dovrebbe coinvolgere anche l'agire di storici e di storiche, quanto meno, come partecipanti ad una discussione.

Sull'archivista si concentra ancora una volta un ruolo di demiurgo, declinato però in modi radicalmente nuovi: assicurare preventivamente che questo processo di accumulo progressivo di tracce si realizzi compiutamente senza iati e lacune, a dispetto delle molte possibili fratture che potrebbero contrassegnare i passaggi del documento digitale tra i "sistemi complessi" del paradigma ICT, il suo approdo da questi al *locus credibilis* digitale e il suo proiettarsi verso nuovi contesti d'uso a partire dalla custodia offerta dall'archivio[443].

Non è mai stato compito dello storico quello di descrivere record di archivi, ma come già ho riflettuto quando abbiamo affrontato il funzionamento di OAIS e di DCMI, alla storica credo sia richiesto lo sforzo di immaginare forme di storiografia dove sistemi complessi e in qualche modo intelligenti, attraverso collegamenti semantici siano in grado di costruire percorsi narrativi in cui non solo le fonti mostrano se stesse, rendendosi intellegibili all'utente, ma mostrano all'utente anche tutte le informazioni e le correlazioni della loro natura storica e archivistica. Questi processi però non prendono forma in modo automatizzato, o meglio divengo automatizzati solo dopo che esseri umani hanno attribuito loro i dati atti a produrre informazioni e costruito i legami concettuali e semantici che le possano legare.
Immaginare che storici e storiche lavorino all'interno di questo processo predeterminando la comunicazione semantica automatizzata, credo che possa essere davvero molto interessante, se non necessario, per far sì che le competenze critiche e analitiche che sono loro proprie non vengano soppiantate completamente dal lavoro di altre professionalità.
Penso possa essere utile osservare RESTORE un interessante progetto iniziato nel giugno 2020 coordinato dall'Istituto Opera del Vocabolario Italiano del CNR, in partenariato con l'Archivio di Stato, il Museo di Palazzo Pretorio di Prato e la Soprintendenza Archivistica e Bibliografica della Toscana, con la partecipazione dell'azienda SPACE SpA nel ruolo di provider tecnologico.

443 A. Alfier, *La descrizione archivistica all'interno di sistemi complessi*, in G. di Marcantonio, F. Valacchi, a cura di, *Descrivere gli archivi al tempo di RIC-CM*, EUM, Macerata 2018 p. 72

Il progetto - cofinanziato dalla Regione Toscana - ha come scopo principale il recupero, l'integrazione e l'accessibilità di dati e oggetti digitali che sono stati prodotti negli ultimi vent'anni dai partner coinvolti, al fine di costituire una base di conoscenza riguardante la storia della città di Prato e delle sue istituzioni, lo sviluppo del suo tessuto economico e imprenditoriale e, non da ultimo, il ruolo centrale giocato dalle donne nella gestione e creazione di scambi d'affari, di network e di contatti che favorissero l'ambiente cittadino generando occasioni di ulteriore ampliamento, sviluppo e benessere non solo economico, ma anche socio-culturale per la città di Prato.

A partire dalla figura del mercante Francesco di Marco Datini, della sua famiglia e del suo *entourage* e muovendo da una dimensione solo apparentemente locale, sarà possibile ricostruire una parte rilevante della storia delle città d'Europa e del Mediterraneo del XIV secolo. Questo contributo si propone di offrire una panoramica della mappatura realizzata per le risorse archivistiche e museali fornite dai partner (modellate sugli standard EAD, EAC-CPF, ICCD, TEI) verso il modello concettuale CIDOC, ontologia scelta da RESTORE come "linguaggio comune" per l'integrazione semantica dei dati[444].

le risorse documentali trattate sono:

- il fondo Datini, costituito da 150.000 lettere e circa 600 registri da cui si possono ricavare informazioni sulle persone coinvolte, sui costi e le tipologie delle merci, sui luoghi deputati agli scambi ecc., è ad oggi l'archivio mercantile per il Medioevo più grande al mondo;
- il fondo dell'Ospedale Misericordia e Dolce, con le sue 7000 unità archivistiche, presenta tutte le articolazioni delle funzioni di un ente di assistenza: dal sostegno al viandante, alla cura del povero e dell'ammalato, fino all'accoglienza dei gettatelli, bambini abbandonati e allevati, grazie allo stesso Spedale, dall'intera Comunità pratese;
- una selezione di opere tratte dalla collezione del Museo di Palazzo Pretorio, catalogate secondo lo standard ICCD (schede cd. "OA" - Opera d'Arte)

Possiamo senza dubbio dire che già nei suoi propositi il progetto è ambizioso, vuole riuscire a rendere interoperabili e consultabili attraverso un processo semantico, risorse appartenenti a diversi enti e custodite con standard differenti.

Questa discrepanza tra i diversi standard richiede un primo e complesso lavoro di omologazione o meglio di traduzione in standard che possano essere utili per l'applicazione dell'ontologia CIDOM.

444 F. Coradeschi Et. al, CIDOC CRM mapping for the integration of RESTORE project resources in «Umanistica Digitale» n° 11, 2021, DHs for society: e-quality, participation, rights and values in the Digital Age https://doi.org/10.6092/issn.2532-8816/13690

Tralasciando le complesse procedure tecniche che hanno visto prima la necessità di studiare i diversi standard e in seguito quella di individuare dei vocabolari controllati ai quali far riferimento per tutti gli URI, in modo tale da non rischiare di imbattersi in processi mancati di disambiguazione, ciò che maggiormente ci interessa è provare ad individuare quegli elementi che ci possano essere utili a comprendere anche da unn punto di vista metodologico i mutamenti che la digitlizzazione sta già importando nei processi di fruizione dei documenti.

Alcuni dei campi presi in analisi per la costruzione dell'ontologia, come *persone, ruoli, localizzazione geoamministrativa, dati tecnici, stato di conservazione, eventi, informazioni, creazione*, sono tutti elementi che si trovano all'interno di standard come EAD o EAC o meglio ISAD(G) o ISAAR e che spesso divengono oggetti informativi grazie alle procedure che abbiamo potuto osservare in OAIS. Ho tenuto a sottolineare questo aspetto, perché per quanto possa sembrare poco rilevante, va a confermare quanto affermato da Vassallo quando fa notare che il primo problema degli standard non è dato dallo standard stesso, ma da come questo venga utilizzato. Vassallo ci fa riflettere sull'eccessiva rigidità nell'utilizzo di strumenti che in realtà con sé portano praticamente tutta l'infrastruttura informazionale necessaria a comprendere, disambiguare e rendere interoperabili le fonti e le informazioni, così CIDOC, come RIC-O (l'ontologia nata dal processo concettuale RIC-CM) si mostrano strumenti tecnologici composti da una costellazione di standard. Con la differenza, questa sì tipica del Web semantico, che tutte queste informazioni stanno nell'involucro oggetto informativo.

Del valore infrastrutturale che ancora assolvono ISAAD e ISAAR vi è chiaro riferimento anche nelle parole di Carriero, Daquino e Tomasi

> Le ontologie sviluppate all'interno del dominio dei beni culturali hanno innanzitutto l'obiettivo di preservare la descrizione delle informazioni descritte tramite gli standard archivistici o catalografici, cercando di ricondurre, quando non traslare, i descrittori in definizioni di classi e proprietà. Tale scelta è dettata dal desiderio di non perdere informazioni nel passaggio al nuovo modello a grafo.[445]

Uno dei passaggi più interessanti rispetto al progetto RESTORE consiste nella volontà dei ricercatori e sviluppatori di andare ad individuare una differenza marcata tra il momento di creazione dell'opera e quello di ideazione.

- La nascita dell'opera in quanto "idea" (le fasi della sua commissione e della sua progettazione da parte dell'autore/artista);

445 V.A. Carriero, M. Daquino, F. Tomasi, *Convergenze semantiche tra musei, archivi e biblioteche. Ontologie per le relazioni interpersonali*, «Jlis.it» 10, 1, 2019 p. 77 DOI: 10.4403/jlis.it-12499 https://cris.unibo.it/retrieve/handle/11585/688514/473119/12499-23397-1-PB.pdf ultima consultazione 4/2/2022

• La manifestazione fisica, tangibile dell'opera (quale elemento facente parte di una collezione fisicamente individuabile e, pertanto, collocato, musealizzato o archiviato, tutelato e anche custodito da Enti e "persone" a vario titolo nel corso della sua storia).

Il primo livello designa l'aspetto immateriale (intangibile) dell'opera, dal momento della sua creazione (anche solo concettuale), mentre il secondo livello ne descrive la produzione e trasposizione su di un supporto fisico (tangibile). La distinzione dei due aspetti è importante perché si tratta di punti di vista differenti che concorrono ugualmente alla realizzazione dell'oggetto (aristotelicamente: ente come sinolo di materia e forma) nel suo insieme e che, ai fini della mappatura dei dati, impongono una riflessione sull'informazione da circoscrivere e descrivere e ne richiedono una interpretazione corretta e concreta. Ad es., colui che identificheremo come "padre dell'idea" non necessariamente corrisponde a colui che ne crea la veste fisica o a colui che, infine, ne diventa il possessore o che ne detiene la custodia. Tutti questi passaggi necessitano di essere documentati[446].

Questo passaggio ci mette di fronte ad un lavoro di elaborazione concettuale che sottostà alla creazione di una ontologia, in questo caso gli sviluppatori per dare più elementi possibili di comprensione alla macchina e all'utente, hanno compiuto un processo di astrazione dividendo tra coloro che avevano pensato l'idea e coloro che l'idea l'avevano realizzata, creando così due categorie concettuali utili per l'utente sì, ma anche per la macchina a creare delle differenze nel processo di elaborazione e ricerca.
Dobbiamo chiederci però di chi sia la responsabilità intellettuale di un processo di questi tipo. Come si fa ad indagare un'idea? È chiaro che in un'analisi ermeneutica della storia o della fonte vi è lo spazio perché questo processo possa prendere forma, ma a chi dare questo ruolo d'analisi? Dal momento che questo sdoppiamento sul concetto di opera - da una parte il pensare l'opera, dall'altra il realizzare l'opera - si mostra come determinate nei processi di ricerca che poi svolgerà l'utente finale, penso che perché possa prendere forma debba passare da un attento dibattito. Non credo che all'interno del team di sviluppo del progetto manchino competenze adeguate, ma mi chiedo se si possa immaginare di rendere come buona prassi, che ogni qual volta si avvia un processo di costruzione di ontologie digitali all'interno di archivi il team di lavoro debba essere necessariamente interdisciplinare e che debba coinvolgere storici e linguisti oltre a informatici e archivisti.

446 F. Coradeschi Et. al, CIDOC CRM mapping for the integration of RESTORE project resources in «Umanistica Digitale» n° 11, 2021, DHs for society: e-quality, participation, rights and values in the Digital Age https://doi.org/10.6092/issn.2532-8816/13690

Il tema del Web semantico oltre ad intrecciarsi con le questioni legate all'intelligenza artificiale, dal momento che utilizzando linguaggi *machine readble* stiamo istruendo una macchina quanto meno ad orientarsi tra contesti generati dalle triple, ci riporta a dover ragionare sul tema dei big data per come lo abbiamo affrontato nel primo capitolo.
In quella circostanza, seguendo anche il ragionamento di Floridi ci chiedevamo se i processi computazionali per elaborare i big data dovessero essere di natura prevalentemente quantitativa. Già in quelle pagine siamo arrivati a capire che la quantità è tale che non è possibile gestirla, a meno che non si individuino delle strategie per dare anche un senso qualitativo alla mole di dati, così da poter capire non solo quanto ma cosa.
Riguardo il ruolo di storici e storiche nell'elaborazione di informazione all'interno di sistemi complessi, in queste pagine si è proposta la possibilità di individuare nell'utilizzo di concetti da attribuire alle fonti una strada utile per elaborare dei cluster che aiutino in senso qualitativo a gestire la quantità. Non di meno come abbiamo visto i processi del Web semantico si basano su relazioni di natura concettuale e non solo causale. Ed è proprio a partire da questa ridondanza della necessità dei concetti all'interno di questi sistemi che sono arrivato a pensare che il lavoro di storici e storiche, già oggi auspicabile all'interno dei processi di metadatazione e di costruzione di connessioni semantiche, possa essere, se sviluppato organicamente, uno strumento molto utile allo storico del domandi per capire come potersi muovere in questo diluvio di dati anche da un punto di vista qualitativo.
Di Fazio giustamente sottolinea come nei processi del Web semantico e in questo caso di RIC-CM, descrizione e rappresentazione siano termini inscindibili e ciò non può che farci ripensare a quanto affermato da diversi archivisti che sostengono come gli archivi debbano essere sistemi di conservazione e di comunicazione, ripensiamo quindi al ruolo di EAD o di EAC. Se già osservando questi sistemi e la stratificazione della metadatazione avevamo ipotizzato l'importanza di storici e storiche nel discutere specifici processi, ancor di più credo sia necessario farlo nel Web semantico.

la multidimensionalità informativa tradotta a livello applicativo nell'esposizione dei diversi contesti rispetto a cui si trovano programmaticamente ricondotti i materiali documentari, giocherebbe sul Web una funzione strategica di instradamento per l'utente. Non più quindi informazione complementare resa disponibile ad un'utenza di iniziati alla disciplina comunque in grado di muoversi in contesti di uso molto specializzati, ma vero e proprio orientamento concettuale alla fruizione, capace di costituire canali di divulgazione e di favorire l'accesso a fonti archivistiche destinate nella maggior parte dei casi a restare ignorate, inaccessibili e irraggiungibili all'utente non specialista.[447]

447 S. Di Fazio, *Le ontologie,* in G. di Marcantonio, F. Valacchi, a cura di, *Descrivere gli archivi al tempo di RIC-CM,* EUM, Macerata 2018 p. 95

Da public historian o da storici e storiche convinti della necessità di democratizzare l'accesso alle informazioni anche complesse, e in questo specifico caso di natura storica, non possiamo che rimanere affascinati da ciò che scrive Di Fazio, è indubbio che tentativo della Public History sia proprio quello di liberare la storia dalla fruizione e dalla discussione solo di "iniziati" capaci di trovare e criticare le fonti, abbiamo ben chiari i concetti di shared auothority e di citizen science, ma in questo contesto non possiamo non portare la nostra attenzione al termine "instradamento" che ha come oggetto l'utente e come soggetto la multidimensionalità informativa che si ottiene grazie all'utilizzo di metadati.

Come possiamo interpretare quel temine da un punto di vista metastorico, quale la sua influenza nella costruzione e nella diffusione di conoscenza storica. Il ruolo di storiche e storiche è di saper leggere e criticare le fonti al fine di restituire una narrazione comprensibile e verificabile del contenuto storico. Perché si arrivi a ciò prima di tutto storici e storiche danno forma ad un complesso groviglio di informazioni, prendendo fonti all'apparenza anche distanti tra di loro e collegandole attraverso rapporti certamente causali, ma indubitamente anche concettuali e semantici, questo groviglio di norma rimane nella testa e negli appunti di storiche che poi attraverso un processo di narrazione e spiegazione riescono a costruire un discorso fluido che instrada il lettore o l'ascoltatore verso una conoscenza della storia.

Ma se questi processi divengono automatizzati, dunque il groviglio si compone di significati e contesti attribuiti alle fonti attraverso metadati descrittivi, i quali trovano connessione grazie a linguaggi come RDF o SPARQL e un contesto tramite ontologie come OWL o RIC-O, quale il ruolo dello storico o della storica? Le fonti esistono a prescindere dal lavoro degli storici, esse sono rappresentazione di un evento che appartiene al passato, queste rappresentazioni fino a pochi anni fa erano reperibili solo all'interno di sistemi archivistici complessi e di natura fisica, la dematerializzazione e la rivoluzione digitale portano queste fonti in un nuovo ambiente non più fisico e facilmente accessibile da chiunque.

Già i Sistemi Informativi garantiscono, senza una intermediazione sincronica di uno storico o di un archivista, non solo di indagare le fonti ma di creare connessioni grazie ai metadati che consentono interoperabilità; ma che succede se ai dati e a queste connessioni viene data la possibilità di assumere significato, e di generare una sorta di spiegazione tramite connessioni concettuali in una neo-intermediazione digitale?

Mi si dirà, e non vi è possibilità di confutazione, che il lavoro dello storico e della storica si compone anche di un lungo percorso di studio e di conoscenza della disciplina, percorso che consente di comprendere al meglio le fonti e di portare la ricerca lì dove non è semplice arrivare

senza competenze specifiche. Questo è assolutamente vero, il lavoro compiuto da uno storico sarà sempre più articolato e composito di un processo di narrativizzazione delle fonti attraverso un'ontologia, dal momento che, come ben ci ricorda Topolsky «La selezione e la gerarchizzazione che costituiscono il livello informativo del racconto storico, tramite la loro stessa strutturazione, trasmettono anche il contenuto non informativo (ideologia, politica, morale)[448]».

Ma senza doversi inoltrare in un dibattito critico intorno al ruolo decisivo dello storico o della storica, nel produrre contenuto informativo non direttamente o meglio non unicamente legato al contenuto informativo presente nelle fonti, credo che la domanda da porsi sia in quanti oggi, per ottenere informazioni, si affidano al lavoro realizzato da professionisti della ricerca e in quanti invece cerchino informazioni nel Web in modo a volte eccessivamente acritico.

Chiaramente la mia è una domanda retorica, è indubbio che la maggior parte delle persone per trovare una risposta a domande anche complesse riguardo la storia, ma non solo, si rivolga a Google prima che ad un testo di storiografia. Ma proprio per questo, se crediamo nell'importanza della diffusione sociale della storia dobbiamo andare ad intercettare coloro che non sono avvezzi alla storiografia e sfruttando le possibilità della rivoluzione digitale, costruire una storiografia basata su metadati, concetti e possibili risultati ontologici.

448 J. Topolski, *Narrare la storia. Nuovi principi di metodologia storica*, Bruno Mondadori, Milano, 1997 p. 70

Bibliografia

AA. VV., *La stampa in Italia. 1998-2001*, Federazione Italiana Editori e Giornali, 2001 disponibile online *https://www.fieg.it/upload/studi_allegati/studio19982001.pdf*

AA. VV., *Toward Descriptive Standards. Report and Recommendations of the Canadian Working Group on Archival Descriptive Standards*, Ottawa, Bureau of Canadian Archivists, 1985

AA.VV., *Consultation restreinte de l'Unesco sur l'etablissement d'un plan d'action international à long terme pour l'élaboration de normes de description pour les archives*, Paris (UNESCO), France, 1-2 décembre 1989, in «Bulletin du Conseil international des Archives», 34, juin 1990

AA.VV., G. Frezza, a cura di, *L'arca futura. Archivi mediali digitai, audiovisivi, web,* Meltemi Editore, Roma 2008

AA.VV., *Planning committee on descriptive standards, Comité de planification sur les normes de description, Rules for Archival Description, Regles pour la description des documents d'archives*, Ottawa (Canada), Bureau of Canadian Archivists, 1990

AA.VV., *Toward International Descriptive Standards for Archives. Papers presented at the ICA Invitational Meeting of Experts on Descriptive Standards*. National Archives of Canada, Ottawa, 4-7 October 1988, K.G. Saur, Munchen, New Providence, London, Paris, 1993

Agamben G., *Requiem per gli studenti*, originariamente pubblicato il 23/05/2020 Per l'Istituto Italiano per gli Studi Filosofici https://www.iisf.it/index.php/attivita/pubblicazioni-e-archivi/diario-della-crisi/giorgio-agamben-requiem-per-gli-studenti.html

Agosti M., Crestani F., Pasi G. (a cura di), *Lectures on Information Retrieval: Third European Summer-School, ESSIR 2000, Varenna, Italy, September 11-15, 2000*, Springer Verlag, Heidelberg 2001

Allegrezza S., Gorgolini L., *Gli archivi di persona nell'era digitale*, Il Mulino, Bologna 2016

Archivio di Stato di Perugia, *I modelli dati XML per la codifica elettronica dei dati d'archivio: EAD e EAC-CPF*, 201*3*

Bacci M.L., *Storia minima della popolazione mondiale*, Il Mulino, Bologna 2005

Baeza-Yates R. e Ribeiro-Neto B., *Modern infomation retrieval*, Addison Wesley, Reading (US) 1999

Baldacci C., *Archivi impossibili. un'ossessione dell'arte contemporanea*, Johan e Levi, Monza 2016

Barbanti I., *La codifica informatica dei dati di archivio: EAD ed EAC,* «Regesta» Scrinia, 2/3 2005 https://www.regesta.com/wp-content/uploads/documentazione/Barbanti-EADEAC2005.pdf

Bateson G., *Verso un'ecologia della mente*, Adelphi, Milano

Becattini E., Scanagatta M., *La nuova grande guerra, la prima guerra mondiale nel cinema contemporaneo*, G. Gozzini, P. Masciullo (a cura di) *La guerra delle immagini nel XXI Secolo,* «Cinema e Storia», Rubbettino Editore, Catanzaro 2020

Benasayag M., *Il cervello aumentato l'uomo diminuito*, Erickson, Trento 2015

Benigno F., *L'età moderna. Dalla scoperta dell'America alla restaurazione*, Editori Laterza, Roma Bari, 2006

Benjami W., *Sul concetto di storia*, Eiunaudi, Torino 1997

Berners Lee T., *L'architettura del nuovo Web: dall'inventore della rete il progetto di una comunicazione democratica, interattiva e intercreativa*, Feltrinelli, Milano 2001

Berners-Lee T., *Metadata Architecture: documents, metadata and links, axiomof Web architecture 3. Cambridge, Mass.:Mit, 1997*, https://www.w3.org/DesignIssues/old/Metadata.html

Berners-Lee T., *Network Working Group*, 1998, https://www.ietf.org/rfc/rfc2396.txt

Bertucelli L., Bertella Farneti P., Botti A., *Public History. Discussioni e pratiche*, Mimesis Edizioni, Milano Udine, 2017

Biancato L., Tonioli D., *101 idee per una didattica digitale integrata*, Erickson, Trento, 2021

Bloch M., A*pologia della storia. O mestiere di storico,* Giulio Einaudi Editore, Torino 1998

Bo C., *Letteratura come vita,* Venezia, Marsilio 2003

Bogdan R.J., *Grounds for cognition: how goal-guided behavior shapes the mind*. Hillsdale: Erlbaum, 1994

Bolter J.D, Grusin R., *Rimediation. Competizione e integrazione tra media vecchi e nuovi*, Guerini, Milano 2002

Bonfiglio Dosio G., Pigliapoco S., a cura di, *Formazione, gestione e conservazione degli archivi digitali*, Eum, Macerata 2015

Braudel F. *Scritti sulla storia,* Bompiani, Milano 2016

Braudel F., *Storia misura del mondo*, Il Mulino, Bologna, 1998

Buckland M. K., *Information and information systems*. New York: Praeger, 1991

Bunnell A., *Review of Encoded Archival Description Tag Library – Version EAD3,* «Journal of Western Archives» 7: 1-4 Jodi. 2016

Burke P. *Testimoni oculari. Il significato storico delle immagini*, Carocci editore, Roma 2020

Burke P., *Una rivoluzione storiografica*, Laterza, Roma-Bari, 2017

C. Ginzburg, *Raporti di forza. Storia retorica prova*, Feltrinelli, Milano 2014

Cambridge Modern History, Cambridge 1902

Caplan Priscilla, *Capire PREMIS*, 2021, *https://www.loc.gov/standards/premis/understandingPREMIS_italian_2021.pdf*

Carr E., *Sei lezioni sulla storia*, Einaudi, Torino 2000

Carr N., *Internet ci rende stupidi? Come la rete sta cambiando il nostro cervello,* Cortina Raffaello, Milano 2001

Carriero V.A., Daquino M., Tomasi F., *Convergenze semantiche tra musei, archivi e biblioteche. Ontologie per le relazioni interpersonali,* «Jlis.it» 10, 1, 2019 p. 77 DOI: 10.4403/jlis.it-12499 *https://cris.unibo.it/retrieve/handle/11585/688514/473119/12499-23397-1-PB.pdf*

Cilandri P., *Lo zen e l'arte della manutenzione archivistica. Ibridi e resilienti,* jlis.it, 9, 2018, 3

Cinelli M., De Francisci Morales G., Galeazzi A., Quattrociocchi W., Starnini M., *The echo chamber effect on social media,* PNAS 2021

Cinelli M., Quattrociocchi W., Galeazzi A., *et al. The COVID-19 social media infodemic. Sci Rep* 10, 16598, 2020

Collingwood R. G., *Il concetto della storia*, Fabbri Editori, Milano 1966

Colombi V., Greppi C., Manera E., Olmoti G., Roda R., *I linguaggio della contemporaneità*. *Una didattica digitale per la storia,* Il Mulino, Milano 2018

Comitato MAG, *manuale utente*, 2006 *https://www.iccu.sbn.it/export/sites/iccu/documenti/manuale.html#-b3*

Cook M., Grant K., *A Manual of Archival Description*, Liverpool, Society of Archivists, 1985

Cooley M., *Architect or bee: the human price of technology*. London: Hogarth, 1987

Coradeschi F. Et. al, CIDOC CRM mapping for the integration of RESTORE project resources in «Umanistica Digitale» n° 11, 2021, DHs for society: e-quality, participation, rights and values in the Digital Age

Croce B., *Teoria e storia della storiografia,* Adelphi, Milano 2001

D'Autilia G., *L'indizio e la prova, la storia nella fotografia*, Mondadori, Milano 2005

d'Orsi A., *Picc*olo m*anuale di s*toriografi*a*, Mond*adori*, Milano 2002

Davidson D., *Soggettivo, intersoggettivo, oggettivo,* Raffaello Cortina Editore, Miano 2003

Day M. *Metadata. Mapping between metadata formats.* Uklon: The UK Office For Library and Information Networking, *http://www.ukoln.ac.uk/metadata/*

De Biase L., Homo Pluralis, *Esseri umani nell'era tecnologica*, Codice Edizioni, Torino 20016

de Certeau M., *La scrittura della storia*, Il Pensiero Scientifico Editore, Roma 1977

De Felice R., *L'archivio contemporaneo. Titolario e classificazione sistematica di competenza nei moderni archivi di enti pubblici e provati,* Roma, La Nuova Italia Scientifica, 1988

De Luna G., *La passione e la ragione,* Mondadori, Milano 2004

Dellavalle C., *Economia e industria nella guerra. Le fonti e gli archivi in* P*iemonte*, Franco Angeli, 1987

Detti T., Lauricella G., *Le origini di internet*, Mondadori, Milano 2013

di Marcantonio G., F. Valacchi, a cura di, *Descrivere gli archivi al tempo di RIC-CM,* EUM, Macerata 2018

Di Noia T., De Virgilio R., E. Sciascio, F.M. Donini, *Semantic Web. Tra ontologia e open data*, Maggioli Editori, Santarcangelo di Romagna 2018

Di Nunzio G.M., *Sistemi di elaborazione delle informazioni*, Università degli studi di Padova, 2014, *http://www.dei.unipd.it/~dinunzio/fdi-2014-2015/04_dati_informazioni.pdf*

Dilthey W., *Esperienza vissuta e poesia*, Il nuovo melangolo, Genova 1999

Dilthey, W., *Per la fondazione delle scienze dello spirito,* Franco Angeli, Milnao 1985

Dinoi M., *Lo sguardo e l'evento*. *I media, la memoria, il cinema*, Le Lettere, Firenze 2012

Dorati A., S. Costantini, *Approcci al Web semantico,* http://www.websemantico.org/articoli/approcciwebsemantico.php#inizio

Duranti L., *Il documento archivistico,* in a cura di, L. Giuva, M. Guercio, *Archivistica. Teorie, metodi, pratiche,* Carocci, Roma 2014

Durate M., *Potere computazionale. l'impatto delle ICT su diritto società, sapere*, Meltemi, Milano, 2019

Eco U., *I limiti dell'interpretazione*, La nave di Teseo, Milano 2016

Erpanet Traing Seminar, *Metadata in digital preservation – final report,* Marburg3-5- settembre, 2003, *https://www.yumpu.com/en/document/read/30354260/metadata-in-digital-preservation-erpanet*

Evans M.J., *Authority Control: an Alternative to the Record Group Concept*, in « American Archivist », XLIX, 1986

Fano R.M., *Joseph Carl Robnett Licklider,* in National Research Council Biographical memories V.75, https://www.nap.edu/read/9649/chapter/12#191

Fano R.M., *Joseph Carl Robnett Licklider,* in National Research Council Biographical memories V.75, https://www.nap.edu/read/9649/chapter/12#191

Feliciati P., *I metadati nel ciclo di vita dell'archivio digitale e l'adozione del modello PREMIS nel contesto applicativo nazionale*, UniMc, Pubblicazioni aperte digitali, 2015

Ferraris M., *Documentalità. Perché è necesario lasciare tracce*, Laterza, Bari Roma 2019

Ferraris M., Scienza nuova. *Ontologia della trasformazione digitale*, Rosenberg e Sellier, Torino 2018

Finley M.I., *Uso e abuso della storia : il significato, lo studio e la comprensione del passato,* Einaudi, Torino 1981

Flores M., *Il secolo mondo,* il Mulino, 2002

Floridi L., L*a quarta rivoluzione. Come l'infosfera sta trasformando il mondo,* Raffaello Cortina editorie, Milano, 2017

Floridi L., *Pensare l'infosfera, La filosofia come design concettuale,* Raffaello Cortina Editore, Milano, 2020

Floridi L., The Philosophy of information, Oxford University Press, Oxford 2011

Floridi L., *Web 2.0 vs. the Semantic Web. A Philosophical Assessment*, in «Episteme», 6 (2009), n. 1

Foskett D. J., *Libraries and information systems: a fruitful partnership*. In: *Proceedings of the Annual Conference of CAIS*. Canadian Association for Information Science, 1995, p. 13 *https://journals.library.ualberta.ca/ojs.cais-acsi.ca/index.php/cais-asci/article/view/405/353*

Foucault M., *L'archeologia del sapere. Una metodologia per la storia della cultura*, BUR, Milano 2016

Foucault M., *Le parole e le cose*, BUR, Milano 2020

Geake J. G., *Il cervello a scuola. Neuroscenza e educazione tra verità e falsi miti*, Erickson, Trento 2009

Giacomini G., *Potere Digitale*, Milano 2018

Giannacchi G., *Archiviare Tutto, una mappatura del quotidiano*, Treccani, 2021

Ginsberg A., *Howl,* Il saggiatore, Roma 2015

Ginzburb C., *Il formaggio e I vermi, Adelphu, Milano, 2019*

Ginzburg G., *Il filo e le tracce.Vero e falso Mito*, Feltrinelli, Milano 2015

Girardet H., *Vedere, toccare, ascoltare. L'insegnamento della storia attraverso le fonti*, Carocci, Roma 2015

Goldberg K., (a cura di) *The robot in the garden: telerobotics and telepistemology in the age of the internt, MIT Press, Cambridge, MA-London*

Gorman M., *Metadata or Cataloging? A false Choice*, «Journalof Internet Catloging». 2/1

Guasti C., Commissione Cibreo, *Sul riordinamento degli archivi di stato,* 15 marzo 1870, p. 4 *https://www.icar.beniculturali.it/biblio/pdf/Studi/cibrario.pdf*

Guercio M., *Archivistica informatica. I documenti in ambiente digitale*, Carocci, Roma 2000

Guercio M., G. Marinelli, *La conservazione delle memorie digitali: un quadro di riferimento per un progetto applicativo,* in «Archivi & Computer», a. XVII, fasc. 1/07, 2007

Guercio M., *Gli archivi come depositi di memorie digitali*, «DigItalia. Rivista del digitale nei beni culturali», III/2 (2008)

Guerrini M., Possemato T., *Linked data per biblioteche archivi e musei,* Editrice Bibliografica, Milano 2015

Han B-C., *Nello sciame. Visioni de digitale*, Nottetempo, mIlano 2015

Harari N., *Homo deus. Breve storia del futuro*, Bompiani, Milano2018

Heery R., M. Patel., *Application profiles: mixing and matching metadata schemas*, http://www.ariadne.ac.uk/issue/25/app-profiles/

Hempel. C.G., Antiseri D., *Come lavora uno storico*, Armando Editore, Roma 2018

Hensen S.L., *Archives, Personal Papers, and Manuscripts: a Cataloguing Manual for Archival Repositories, Historical Societies, and Manuscripts Libraries*, Chicago, Society of American Archivist, 1989

Hilbert M., *Video animation on The World's Technological Capacity to Store, Communicate, and Compute Information from 1986 to 2010*

Hjørland B., *Information: objective or subjective/situational?*, «Journal of the american society for information science and technology», 58 (2007), n. 10

Hobsbawm E.J., *L'età degliimperi. 1875 1914*, Editori Laterza, Roma Bari, 2000

Hodgson K., *Metadata foundations, potential and applicationes*, Edmonton: School of Library and Infomation Studies, University of Alberta, 1998

International Council On Archives, ISAAR(CPF): International Standard Archival Authority Record for Corporate Bodies, Persons and Families, Prepared by the Ad Hoc Commission on Descriptive Standards, Paris, France, 15-20 November 1995, Ottawa, 1996

Kiesling K., *EAD as an Archival Descriptive Standard*, American Archivist 60/3

Koselleck R., *Futuro passato. Per una semantica dei tempi storici*, Clueb, Bologna 2007

Koselleck R., *Storia la formazione del concetto moderno*, CLUEB, Bologna 2009

Kuhn T. S., *La struttura delle rivoluzioni scientifiche*, Einuaudi, Torino 2009

Landes D.S., *Prometeo liberato. Trasformazioni tecnologiche e viluppo industriale nell'Europa occidentale dal 1750 ai giorni nostri*, Einaudi, Torino 2017

Landino C., *Archimista, nuove posisbilità per la descrizione e l'interoperabilità*, 2018 https://www.icar.beniculturali.it/materiali/Presentazione_Archimista_20180919.pdf

Lavoi F.B., *Il modello di riferimento per un Sistema informativoaperto per l'archiviazione*, Associazione Italiana Biblioteche, 2017, http://www.soprintendenzaarchivisticatoscana.beniculturali.it/fileadmin/risorse/materiali/Lavoie_-_OAIS.pdf

Le Goff J., *Il tempo continuo della storia*, La Terza, Bari-Roma 2022

Le Goff J., Vernant J.-P., *Dialogo sulla storia*, Laterza, Roma Bari 2020

Leonelli S., *The philosophy of data*

Mantovani G., G. Riva, *Buildinig a bridge between different scientific communities: On Sheridans eclectic ontology of presence*, in Presence: teleoperators and virtual Environments, 10 (5)

Marcuse H., *L'uomo a una dimensione. l'ideologa della società industriale avanzata*, Einaudi, Torino 1970

Mazzini U., *Correzioni critiche di alcune date del Regesto del Codice Pelavicino per il socio Ubaldo Mazzini*. Genova: Società ligure di storia patria, 1914

McLuhan M., *gli strumenti del comunicare*, Il Saggiatore, Milano 2015

Meschini F., *Frammenti e connessioni, il WorldWidWeb come ambiente informativo e il ruolo delle Digital Humanities*, in Offcine della storia, n° 21, Maggio 2020, *https://www.officinadellastoria.eu/it/2020/05/07/frammenti-e-connessioni-il-worldwideweb-come-ambiente-informativo-e-il-ruolo-delle-digital-humanities/*

Meschini F., *Reti memoria e narrazione, archivi e biblioteche digitali tra ricostruzione e racconto*, Settecità, Viterbo 2018

Michetti G., *Il linguaggio per la descrizione archivistica: analisi di un* caso, Archivi & Computer 1

Michetti G., *Il modello OAIS*, in «Digitalia», Anno III, n° 1, 2008

Minuti R., a cura di, AA.VV, *Il web e gli studi storici. Guida critica all'uso della rete*, Carocci Editori, 2017

Montani P., *Limmaginazione intermediale, Perlustrare rifigurare, testimoniare il mondo visibile*, Editori Laterza, Bari 2010

Noiret S., M. Scanagatta, D. Paci, M. Ravveduto, *La storia come bene comune, le nuove frontiere della public history digitale*, Passato e Presente, F. 113, Franco Angeli, Milano 2021

Noiret S., *Public History' e 'storia pubblica' nella rete*, Ricerche storiche, 39, num. 2-3, 2009

Nordau Max, *Il senso della storia*, Barion, Milano 1930

Paci D., a cura di, AA.VV., *Storia in Digitale. Teorie e metodologie*, Unicopli, Milano 2019

Paci D., *Conoscere è partecipare: digital public history, wiki e citizen humanities*, Umanistica Digitale

Pala F., *Lo standard EAD3 per la codifica dei dati archivistici: qualche novità e molte conferme*, JLIS.it 8, 3, DOI: 10.4403/jlis.it-12407, 2017 *file:///D:/backup%2028_10_18/Documenti/unimore/dottorato%2017-18/tesi/fonti%20digitali/Pala_%20EAD_2017_12407-22919-3-PB.pdf*

Palmini O., recensione a L. Floridi, *Pensare l'infosfera. La filosofia come design concettuale,* in «Pandora Rivista» 20-02-2020

Pezzica L., *Giorni di Storia. Festa dei lavoratori.* prodotto multimediale realizzato da DHLab http://memoria.lazio.it/iniziative/giorni-di-storia/festa-dei-lavoratori/

Pezzica L., *L'archivio liberato. Guida teoricoì-pratica ai fodni storici del novecento*, Editrice Bibliografica, Milano 2020

Pezzica L., Valacchi F., *Dimensioni archivistiche. Una piramide rovesciata,* Editrice Bibliografica, Milano 2021

Pierazo E., *Digital Scholarly Editing: Theories, Models and Methods*. Farnham, Surrey, England: Ashgate, 2015

Pierazzo E., *Edizioni digitali di documenti complessi*, in ciclo di seminari di informatica umnistica, *La tecnologia informatica applicata alle scienze filologiche e librarie*, Scuola di Alta Formazione A. Varvaro, Unina, *https://www.youtube.com/watch?v=IbIySwYOsgM&feature=share&fbclid=IwAR1nDi5RiaNDFeI4BlFNRfvpfiVmgd_D69oCvXga4EzGG0QC4jOWOWI5r_4*

Pigliapoco S., *Progetto archivio digitale. Metodologia, sistemi, professionalità*, Civita Editori, Torre del lago 2018

Pitti D., *Encoded Archival Description: The Development of an Encoding Standard for Archival Finding Aids*, The American Archivist 60

Pitti D., *Una ricchezza in comune verso uno standard che codifichi strumenti per la ricerca*. Archivi & Computer 2:154-62

Pomian K., *L'ordine del tempo*, Einaudi, Torino 1992

Popper K. R., *Miseria dello storicismo*, Giangiacomo Feltrinelli Editore, Milano 2019

Popper K.R., *Miseria dello storicismo*, Giangiacomo Feltrinelli Editore, Milano 2019

Rampazzi M., Tota A.L., a cura di, *Il linguaggio del passato, memoria colettiva mass media e discorso pubblico*, Carocci, Roma 2013

Ricci F., *Gli standard internazinali di descrizione archivistica: dalle origini alla seconda versione di Isad(G), in «*Rassegna degli Archivi di Stato», anno LXIII - n. 1, Roma, 2003

Ricci F., *Gli standard internazionali di descrizione archivistica: dalle origini alla seconda versione di Isad(G), in «*Rassegna degli Archivi di Stato», anno LXIII - n. 1, Roma, 2003

Ricoeur P., Castoriadis C., *Dialogo sulla storia e l'immaginario sociale*, Jaca Book, Milano 2017

Ricoeur P., *La memoria, la storia, l'oblio*, Raffaello Cortina Editore, 2003

Ridi R., *La piramide dell'informazione, una introduzione, https://aibstudi.aib.it/article/view/11903/11481#5*

Ridi R., *Livelli di irrealtà: oggettività e soggettività nell'organizzazione della conoscenza*, «Bibliotime», 18 (2015), n. 2, http://www.aib.it/aib/sezioni/emr/bibtime/num-xviii-2/ridi.htm, par. 2

Ridi R., *Piramidi delle informazioni e KO,* 10° incontro ISKO Italia, Università di Pavia, 14/05/2021, a questo link è possibile sentire il suo intervento *http://www.riccardoridi.it/video/isko2021.mp4*

Riva G.*, Tecnologie della Presenza. Concetti e applicazioni*, Mondo Digitale, 2009; (3)

Romano R., *Braudel e noi*, Donzelli Editore, Roma 1995

Romiti A., *Archivistica generale. Primi elementi*, Torre del Lago, Civita editoriale 2011

Rossi P., a cura di, *La teoria della storiografia oggi*, Il Saggiatore, Milano 1983

Rovelli C., *L'ordine del tempo*, Adelphi, Miano 2017

Sack H., *Knowledge Engineering with Semantic Web Technologies*, Lecture 3: Ontologies and Logic. 3.9 DLs and the Open World Assumption, 2015, *https://open.hpi.de/files/8e5a3cf1-09e6-4c26-9afa-c1e5b967e11f*

Salvatori E., R. R. Del Turco Et. al, *Il codice Pelavicino tra edizione digitale e Public History,* https://umanisticadigitale.unibo.it/article/view/7232/6991

Savoja M., *Descrizioni di archivi e soggetti produttori: ISAD(G) e ISAAR(CPF), in Gli standard internazionali per la descrizione archivistica: le regole ISAD(G).* Incontro seminariale sulle proposte di revisione elaborate dal gruppo di lavoro ANAI. Bologna, 11 febbraio 1998, Bologna, Archivio di Stato di Bologna - ANAI Sezione Emilia-Romagna, 1998

Scanagatta M., *E l'America creò gli hippie. Storia di una avanguardia*, Mimesis, Milano Udine 2013

Science 01 Apr 2011:Vol. 332, Issue 6025, pp. 60-65 DOI: 10.1126/science.1200970

Sebastiani M., *Identificatori persistenti per gli oggetti digitali*, «DigItalia», 2005

Spengler O., *Il tramonto dell'occidente*, Longanesi, Milano, 2012

Spina S., *Archivi nell'era delle digital humanities dei big data e della genetica,* Algra Editore, Catania 2020

Spitzer M., *Demenza digitale. Come la nuova tecnologia ci rende stupidi*, Corbaccio, Milano 2019

Stibbe H.L.P., *Commission Ad Hoc sur les norms de description (ICA/DDS)*, in « Bulletin du Conseil international des Archives », 46, 1996

Stibbe H.L.P., K. Hawoth, *ISAAR (CPF) Standard internazionale per i record d'autorità archivistici di enti, persone, famiglie*, International Council on Archive, II edizione, 2004

Tedeschini Lalli M., *Un secolo e mezzo di Storio di nuovo online*, 23 feb. 2021, *https://tedeschini.medium.com/un-secolo-e-mezzo-di-storia-di-nuovo-online-594a9c1e6358*

Tedeschini-Lalli M., *Tecnologia digitale obsoleta. Un secolo e mezzo di Storia a rischio*, 27 nov. 2020, *https://tedeschini.medium.com/tecnologia-digitale-obsoleta-un-secolo-e-mezzo-di-storia-a-rischio-1bb75bf68c2f*

The consultative committee for Space Data System Practices, *Reference model for an open archival infromation sistem*, Recommendend Practice, 2012, https://public.ccsds.org/pubs/650x0m2.pdf

Thom R., *Dalla morfogenesi alla struttura*. In: *Teoria dell'informazione*, atti del Seminario tenutosi a Venezia nella sede della Fondazione Giorgio Cini dal 28 maggio al 1 giugno 1973, a cura di Jacques Roger. Bologna: Il Mulino, 1974, p. 73-86: p. 83. Cfr, R. Ridi, *la piramide dell'informazione. Una introduzione*, https://aibstudi.aib.it/article/view/11903/11481#5

Topolski J., *Narrare la storia. Nuovi principi di metodologia storica*, Bruno Mondadori, Milano, 1997

Traverso E., *Il passato. Istruzioni per l'uso, Storia, memoria, politica*, Ombre Corte, Verona 2005

Valacchi F., *Gli archivi tra storia e uso futuro,* Editrice Bibliografica, Milano 2020

Valacchi F., *L'applicazione degli standard agli strumenti archivistici nel web*, UniMc, 2020 https://docenti.unimc.it/federico.valacchi/teaching/2020/22027/files/standard-di-descrizione-archivistica

Valacchi F., *L'applicazione degli standard agli strumenti archivistici nel Web,* Università i Macerata, 2020 https://docenti.unimc.it/federico.valacchi/teaching/2020/22027/files/standard-di-descrizione-archivistica

Valacchi F., *Pezzi di cose nel mondo. Il processo di integrazione delle descrizioni archivistiche nei sistemi interculturali,* «Jlis.it» 7,2 2016

Valacchi F., *Software di descrizione e sistemi informativi archivistici: aspetti teorici*, https://docenti.unimc.it/federico.valacchi/teaching/2019/19962/files/archivi-e-ict,

Vettore S., *Archivi digitali contemporanei. Arsenali di autorità o democrazia?*, in Archivi, 9 2014

Vitali S., *La descrizione degli archivi nell'epoca degli standard e dei sistemi,* in, a cara di, L. Giuva, M. Gercio, *Archivistica. Teoria, normativ e prassi,* San Miniato, Archilab, 2002

Vitali S., *Ordine e caos: Google e l'arte della memoria,* in A. Spaziante, a cura di, *Il futuro della memoria: la trasmissione del patrimonio culturale nell'era digitale*, CSI Piemonte, Torino

Vitali S., *Passato digitale. Le fonti dello storico nell'era del computer,* Bruno Mondadori, Milano 2004

Weinberger D., *La stanza intelligente. La conscenza come prorietà della rete*, Codice Edizioni, Torino 2020

Weller K., *Knowledge Rapresentation in the Social Semantic Web*, De Gruyter Saur, Berlin-New York 2010

Weston P.G., *L. Sardo, Metadati*, Associazione Italiana Biblioteche, Milano 2017

White H., *Forme di storia. Dalla realtà ala narrazione*, Carocci Editore, Roma 2006

White H., *Metahistory vol. I, vol. II*, Meltemi, Milano 2019

Willer M., Dunsire G., Bibliographic information organization in the semantic web. Oxford, Cambridghe; New Delhi: Chandos Publishing, 2013

Zanni Rosiello I., *Gli archivi nella società contemporanea,* Il Mulino, Bologna 2009

Sitografia

https://www.istat.it/it/archivio/255785

https://www.popularmechanics.com/adventure/sports/a5772/cameras-fouls-and-referees/#sidepanel

http://news.bbc.co.uk/sport2/hi/tennis/7305404.stm

https://www.hawkeyeinnovations.com/sports/tennis

https://www.jpl.nasa.gov/news/press_kits/mars_2020/launch/

https://mars.nasa.gov/technology/helicopter/#Milestones

https://www.apple.com/it/newsroom/2006/05/23Nike-and-Apple-Team-Up-to-Launch-Nike-iPod/

https://www.nike.com/it/help/a/apple-watch-nike

https://www.ted.it/nest-google/

https://evernote.com/intl/it

https://www.samsung.com/it/refrigerators/familyhub/

https://www.apple.com/it/airtag/

https://rfid.it/apple-punta-sulla-ultra-wideband-con-i-nuovi-airtag/

https://www.rfidglobal.it/tecnologia-rfid/

https://www.rfidglobal.it/case-history-gse/

http://dati.camera.it/it/linked-data/

https://www.agid.gov.it/it/agenzia/strategia-quadro-normativo/codice-amministrazione-digitale/

https://web.archive.org/web/20080913174125/http://www.web2con.com/web2con/.

https://www.domo.com/

https://www.statista.com/

https://www.raconteur.net/

https://kinsta.com/

https://www.socialinsider.io/;

https://www.cisco.com/c/en/us/index.html

https://www.gsmaintelligence.com/

https://www.gwi.com/

https://locowise.com/

https://www.appannie.com/en/

https://www.similarweb.com/

https://www.statista.com/statistics/871513/worldwide-data-created/

https://www.sba.unipi.it/it/risorse/banche-dati/statista

https://wearesocial.com/blog/2019/01/digital-2019-global-internet-use-accelerates
https://datareportal.com/reports/digital-2021-global-overview-report
https://web-assets.domo.com/blog/wp-content/uploads/2020/08/20-data-never-sleeps-8-final-01-Resize.jpg
https://advanced-hindsight.com/blog/new-big-data/
https://www.nsf.gov/cise/bigdata/
https://www.insuranceup.it/it/call4ideas/luciano-floridi-per-avere-successo-con-i-dati-ci-vuole-etica/
http://freedom.dicea.unifi.it/pub/pluto/files/journal/pj0207/Digital_Divide.html
https://www.istat.it/it/archivio/236920
https://avanzamentodigitale.italia.it/it/progetto/spid
https://innovazione.gov.it/progetti/identita-digitale-spid-cie/
http://www.icviacutigliano.it/sito/news_2018-2019/Manifesto_Storia
https://www.lingue.unich.it/sites/st07/files/04_-_codifica_informazioni.pdf
https://ispr.info/about-presence-2/about-presence/
https://it.wikipedia.org/wiki/Napster
https://www.buzzfeednews.com/article/alexkantrowitz/how-the-retweet-ruined-the-internet
https://www.unive.it/pag/14024/?tx_news_pi1%5Bnews%5D=8464&cHash=8fb0ead51f3e55bdbbe155b4779281bb
https://www.facebook.com/Autonomia-Bianconera-Juventus-Supporter-1399776340263923/
https://developers.google.com/search/docs/basics/how-search-works?hl=it
https://www.google.com/search/howsearchworks/?hl=it
https://siusa.archivi.beniculturali.it/cgi-bin/pagina.pl?TipoPag=profist&Chiave=238
http://www.militarystory.org/il-piano-emergenza-27-del-comitato-di-liberazione-nazionale/
https://www.linkedin.com/in/michele-giuliani-31185422/
http://www.metarchivi.it/
https://www.bl.uk/about-us/our-story/facts-and-figures-of-the-british-library#
https://www.youtube.com/watch?v=fU7Dk8ZZSLY
http://www.qgistutorials.com/it/docs/3/georeferencing_basics.html
https://ispr.info/
https://www.pandorarivista.it/articoli/pensare-infosfera-di-luciano-floridi/
https://www.camera.it/_dati/leg18/lavori/documentiparlamentari/indiceetesti/015/134/00000005.pdf
https://www.archivioluce.com/
https://www.gazzettaufficiale.it/eli/id/2017/08/14/17G00140/sg

http://www.ilmondodegliarchivi.org/component/content/article?id=546:la-liberalizzazione-delle-riproduzioni-digitali-con-mezzo-proprio-negli-archivi-di-stato-italiani
https://www.isontina.beniculturali.it/getFile.php?id=4034
https://luceperladidattica.com/2018/05/08/narrare-la-storia-ricerca-e-creativita-a-cura-di-manfredi-scanagatta-con-gli-studenti-della-vb-liceo-classico-cornelio-tacito-di-roma/
http://www.unesco.org/new/en/communication-and-information/memory-of-the-world/register/full-list-of-registered-heritage/registered-heritage-page-6/newsreels-and-photographs-of-istituto-nazionale-luce/
http://san.beniculturali.it/web/san/home;jsessionid=64BCBB8C1B69429976A5BA7F84AB157C.sanapp01_portal
https://sias.archivi.beniculturali.it/cgi-bin/pagina.pl
https://siusa.archivi.beniculturali.it/cgi-bin/siusa/pagina.pl
https://www.maxxi.art/centro-archivi/
http://www.archivi.beniculturali.it/index.php/abc-degli-archivi/glossario
http://www.ciscra.org/mat/mat/term/1854
http://blog.unimc.it/archivisticattiva/pagina-di-esempio/
https://www.limesonline.com/rubrica/riconoscimento-facciale-cina-huawei
https://aiucd2021.labcd.unipi.it/programma/
https://www.theguardian.com/news/series/cambridge-analytica-files
https://www.vaticannews.va/it/papa/news/2020-02/papa-francesco-intelligenza-artificiale-algor-etica-discorso.html
https://aws.amazon.com/it/s3/glacier/
https://www.crunchbase.com/organization/carpathia-hosting
https://www.cogentco.com/en/products-and-services/success-stories/717-cloudhelix
https://www.ovhcloud.com/it/bare-metal/?xtor=SEC-13-GOO-[it_int_2020_ovh_baremetal_undefinite_sale_acquisition_srch_defensive()]-[266448374853]-S-[ovhcloud]&gclid=Cj0KCQjwvaeJBhCvARIsABgTDM5hHOmhbYYn5UQrEXwpSI2xYmbmi0kcuvm22HMkTbXdEMky1iPlgG8aAlqyEALw_wcB
https://www.digitalrealty.com/data-centers/london/cloud-house-1-muirfield-crescenthttps://www.digitalrealty.com/data-centers/london/cloud-house-1-muirfield-crescent
https://engineering.fb.com/2015/05/04/core-data/under-the-hood-facebook-s-cold-storage-system/
http://www.labcd.unipi.it/progetti/codice-pelavicino-digitale/
https://tei-c.org/
http://dss.collections.imj.org.il/project

https://depero.it/archivio/

https://epistolariodegasperi.it/#/archivio_digitale/cerca?search=true

https://eac.staatsbibliothek-berlin.de/

https://dublincore.org/specifications/dublin-core/usageguide/glossary/

https://cdlib.org/resources/technologists/glossary-of-digital-library-terms/#D

https://cordis.europa.eu/article/id/7486-desire-development-of-a-european-service-for-information-on-research-and-education

https://mpeg.chiariglione.org/

https://www.iccu.sbn.it/export/sites/iccu/documenti/manuale.html

https://www.diglib.org/

http://ims.dei.unipd.it/data/san/met/documents/2010-05-20_oggetti-digitali-san20100425-2.pdf

http://www.labcd.unipi.it/ph-remix/

https://www.thinglink.com/

https://ndsr.nycdigital.org/tag/oais-reference-model/

https://www.w3.org/RDF/

https://www.w3.org/DesignIssues/LinkedData.html

https://instructionaldesign.com.au/the-semantic-web/

RiC-CM (Records in Contexts - A Conceptual Model for Archival Description) (beniculturali.it)