

## Digital curation: new skin for the old ceremony Strumenti OCLC per il Born-digital

(Gabriele Lunati, Ifnet Srl, OCLC development partner Italia)



Nel 1974 Leonard Cohen pubblicò il suo quarto album in studio dal titolo “New skin for the old ceremony” (*Pelle nuova per l'antico rito*); questo titolo mi è venuto in mente trovandomi a distanza di anni ad interessarmi di digitale - grazie all'invito di Ornella Foglieni - ed in particolare di ciò che viene chiamato “Digital curation”. Precisamente per riferirvi su ciò che OCLC, che di digitale si occupa da sempre, ha prodotto e sta producendo in questi ultimi tempi.

Perché Leonard Cohen? Nella nostra professione siamo piuttosto abili ad inventare nuove sigle e nomi più calzanti per fare una specie di “restyling” o, se preferite l'italiano, di “rivisitazione” di problemi, tecniche, metodologie tipiche della professione.

Così mi sembra di poter parlare di rivisitazione per la *Digital curation* che fa da titolo collettivo a una serie di operazioni, di scelte e di “cure” che finora erano considerate separatamente e purtroppo spesso, almeno da noi, non viste in connessione e con quella consequenzialità che ho sempre ritenuto e sostenuto essere necessaria.

Tuttavia una peculiarità della *Digital curation* esiste e due anni fa un importante contributo di Laura Testoni ci chiariva le idee su cosa significasse *Digital curation* e *Content curation*.<sup>1</sup>

Prescindo dalla elencazione delle otto fasi che Laura Testoni riporta dalla lista del DCC (Digital Curation Center) per notare solo che, oltre alle ovvie operazioni di scelta, creazione, ecc. vi sono le fasi finali di accesso, uso, riutilizzo e trasformazione.

**... “la content curation allude (...) alla necessità di organizzare i contenuti (...) in un contesto di proliferazione esponenziale dell'informazione.”**

Sempre Laura Testoni sottolinea la ... **“Inquietante familiarità della Content curation”**; ...

**“il concetto di Content curation è davvero familiare per un bibliotecario e ha evidenti somiglianze con molte attività abituali per la professione, soprattutto per chi presidia i servizi informativi ed il reference (disseminazione selettiva, abstracting, indexing)”**.

Insomma un problema complesso e urgente. Non si tratta infatti solo di quantità dell'informazione ma anche della sua qualità e della sua inevitabile obsolescenza causata anche dagli strumenti e dai supporti utilizzati per la sua produzione.



Quella che vedete è la sede centrale di OCLC: il Kilgour Building così chiamato in onore del suo fondatore Frederick Kilgour. Si trova a Dublin Ohio, e proprio nella sua sede “*nel marzo 1995 si è tenuta una conferenza alla quale i partecipanti – bibliotecari, archivisti, editori, ricercatori e sviluppatori di software, oltre ad alcuni membri dell'IETF (Internet Engineering Task Force) – hanno convenuto sulla necessità di creare un insieme di*

*strumenti condivisi per l'accesso alle risorse digitali.*”<sup>ii</sup>

Nacque così il **Dublin Core**, il primo schema di metadati da quel momento riferimento per ogni schema successivo; da allora (or sono vent'anni) OCLC non ha più abbandonato le ricerche e l'impegno per lo studio del digitale in ogni sua forma.

OCLC è più noto per il suo immenso catalogo collettivo ricco di ogni possibile tipo di record, proveniente da ogni paese immaginabile (Italia compresa) e in ogni possibile lingua. Certamente una cosa unica nel suo genere. Ma secondo me il vero valore aggiunto di questa istituzione è il suo **Centro di Ricerche**,<sup>iii</sup> una struttura in cui lavorano una quarantina di ricercatori tra cui nomi assai noti (Tennant, Lavoie, Van der Werf), coordinati dal più noto di tutti il britannico Lorcan Dempsey.



Tra i ricercatori in prima linea sul digitale abbiamo la signora che vedete nella foto che si chiama **Ricky Erway**; dalle sue attività sono partite alcune delle ricerche più interessanti e delle quali vorrei mettervi a parte e che riguardano principalmente il materiale Born-Digital, cioè digitale nativo.

Il trattamento di questo materiale costituisce un problema serio anche per aspetti legati ai media su cui esso viene prodotto e reso disponibile.

Nel 2010 la signora Erway ha prodotto un breve documento col titolo ***Defining born digital***, in cui elencava e dava una definizione precisa delle varie tipologie di materiale Born Digital, individuandone anche i maggiori rischi di perdita nella obsolescenza di hw e sw, nel degradarsi delle informazioni, e indicando come primo passo per la sua

salvaguardia l'invito a stabilire delle politiche comuni di trattamento ed una nuova consapevolezza del problema.<sup>iv</sup>

Per inciso, val la pena di ricordare che nel 2006 è nato il **Museum of Obsolete media**, utile ausilio a comprendere meglio su che razza di spazzatura informatica sono immagazzinati molti dei nostri dati.<sup>v</sup>

Nel 2012 Ricky Erway pubblica due nuovi e sempre sintetici documenti (rispettivamente di 5 e 7 pagine), che iniziano a dare delle indicazioni metodologiche e di lavoro molto concrete.

Il primo, dal titolo piuttosto significativo, **You've Got to Walk Before You Can Run** si occupa dei primi passi da compiere; non ci si aspettino elaborazioni teoriche e discorsi di principio ma indicazioni molto pragmatiche come per esempio:

- passate in rassegna ed inventariate ogni tipo di supporto presente;
- memorizzatene la collocazione rispetto alla raccolta;
- riportate tutto ciò che potete circa l'hw con cui è stato prodotto;
- scegliete il materiale definendone le priorità in termini di importanza rispetto alla sua conservazione.

E così via fino alla indicazione di documentare tutte le operazioni compiute.

Chiude il documento una tabella di undici punti, **Technical steps for readable media**, e la dichiarazione che “dopotutto lo scopo non è la conservazione fine a se stessa ma quello di garantire l'accesso all'informazione.”<sup>vi</sup>

La seconda pubblicazione, dal titolo intraducibile, **Swatting the Long Tail of Digital Media**, ha però un sottotitolo molto esplicito **“Un appello alla collaborazione”**. La parola swatting deriva da una nuova sigla che la signora Erway introduce: **SWAT (software and workstations for antiquated technology)** che sta alla base dei cosiddetti S.W.A.T. sites che lei propone.

In sostanza Ricky Erway parte dalla considerazione che non tutte le strutture hanno gli strumenti e le competenze tecniche necessarie per trattare i materiali Born-Digital su media obsoleti, e che invece possono esserci delle istituzioni (università o centri specializzati) che potrebbero fornire queste competenze come un servizio condiviso, mettendo a disposizione la propria esperienza e soprattutto i propri macchinari. Alcuni di questi servizi potrebbero essere tariffati in modo ragionevole.

Delegare il recupero dei dati a tali centri permetterebbe a chi usa il servizio di concentrarsi sulla scelta e sulla gestione dei contenuti piuttosto che sul loro complesso ed oneroso recupero.

Anche questa pubblicazione si chiude con una tabella finale di domande utili all'analisi, all'uso ed alla eventuale fondazione di uno SWAT site.<sup>vii</sup>

Nel 2013 un breve articolo dal titolo **Demystifying Born Digital**, apparso sulla rivista di OCLC **NextSpace**, dà conto di due indagini condotte negli Stati Uniti, nel Regno Unito ed in Irlanda presso svariate decine di importanti istituti, in cui oltre l'80% dei rispondenti sottolineava l'esigenza di avere maggiore formazione sul trattamento del materiale Born-digital.<sup>viii</sup>



In estrema sintesi, le ricerche si concludevano con un grido di dolore:

- **Under-collected**
- **Under-counted**
- **Under-managed**
- **Un-preserved**
- **Inaccessible**

Questa la situazione rilevata e denunciata.

Arrivo così alla ricerca ed allo strumento più ampio elaborato sull'argomento, **Walk this way**, un documento di 35 pagine che si concentra sui passi da compiere per il trattamento del materiale Born-Digital. I quattro capitoli di cui si compone si occupano di:

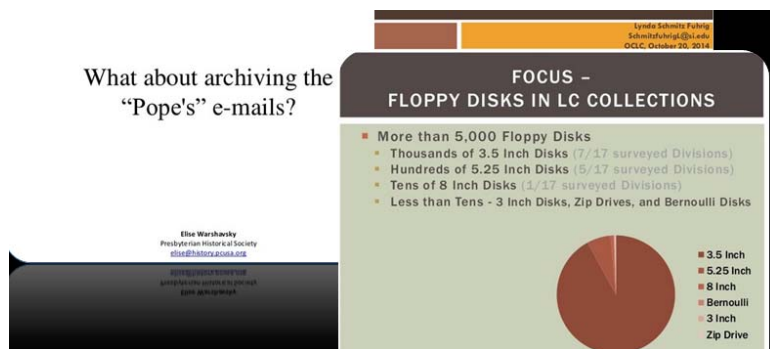
- preparazione delle stazioni di lavoro;
- trasferimento dei dati;
- gestione della sicurezza (virus, ecc.);
- cura dei contenuti (per esempio la deduplicazione dei file).

Segue una serie di esempi di flussi di lavoro sperimentati e soprattutto, laddove necessario, si danno indicazioni su strumenti disponibili in rete, per esempio per la deduplicazione sono indicati ben sei diversi sw adatti allo scopo.

Conclude un'ampia e selezionata rassegna di siti di approfondimento.

L'ultimo documento che vorrei suggerire (anche l'ultimo in ordine di apparizione nel 2014), dal titolo **Agreement Elements for Outsourcing Transfer of Born Digital Content**, è una breve e concentrata rassegna di punti da tenere presente nel caso in cui le attività di conservazione venissero richieste in outsourcing. E' un documento a più mani che si è avvalso della consulenza di ben sei esperti di varia estrazione.<sup>ix</sup>

A supporto del dibattito sul problema del Born-Digital, nell'ottobre dello scorso anno 2014, si sono riuniti presso OCLC, in un convegno dal titolo **Innovative solutions for dealing with Born digital contents in obsolete formats**, otto personalità di varia provenienza (dalla Smithsonian Istitution, alla Library of Congress, dalla Presbyterian Historical society alla Pennsylvania State University) per conoscere e condividere le scelte operate per il trattamento del materiale digitale nativo.



Guardando le slide disponibili seguendo il Webinar online, ci rendiamo conto della complessità della materia. C'è chi si chiede come dovrebbero essere trattate le mail del Papa e qualcun altro cosa fare dei 5000 floppy rintracciati alla Library of Congress.

E così via.<sup>x</sup>

Rimando alla lettura dei documenti citati, liberamente accessibili, di cui in nota trovate tutti gli URL.

### Concludendo...



È interessante osservare come il problema della conservazione della memoria stia superando i confini strettamente professionali per approdare nel sentire culturale comune, al punto che non stupisce trovare su un magazine popolare (il supplemento Sette del Corriere della sera) una paginetta a firma Roberto Cotroneo in cui leggiamo:

*Quante cose dobbiamo conservare? E di quante cose abbiamo bisogno per ricordare? E quante devono essere di continuo a portata di mano? L'era di Internet ci ha obbligato a tenere tutto*

...

*Dentro il nostro piccolo universo abbiamo preteso di metterne uno infinito che non sappiamo di cosa sia fatto. Ci siamo illusi che conservare equivalga a ricordare. E abbiamo dimenticato cosa significhi davvero saper ricordare.<sup>xi</sup>*

In ogni caso, pur rispettando le opinioni e i dubbi di Cotroneo, il nostro compito è urgente come ci ricordava nel citato convegno Linda Smitz Fuhrig, archivista della Smithsonian Institution Archive: **AS SOON AS POSSIBLE IF POSSIBLE**

---

<sup>i</sup> **Laura Testoni:** *Digital curation e content curation: due risposte alla complessità dell'infosfera digitale che ci circonda, due sfide per i bibliotecari* *Bibliotime*, marzo 2013  
<http://www.aib.it/aib/sezioni/emr/bibtime/numxvi1/testoni.htm>

<sup>ii</sup> **Dublin Core:** [http://it.wikipedia.org/wiki/Dublin\\_Core](http://it.wikipedia.org/wiki/Dublin_Core)

<sup>iii</sup> **OCLC Research:** <http://www.oclc.org/research.html>

<sup>iv</sup> **Ricky Erway:** *Defining "Born-Digital"*  
<http://www.oclc.org/content/dam/research/activities/hiddencollections/borndigital.pdf?urlm=168879>

<sup>v</sup> **Museum of obsolete media:** <http://www.obsoletemedia.org/>

<sup>vi</sup> **Ricky Erway:** *You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media*  
<http://www.oclc.org/content/dam/research/publications/library/2012/2012-06.pdf?urlm=168601>

<sup>vii</sup> **Ricky Erway:** *Swatting the Long Tail of Digital Media: A Call for Collaboration*  
<http://www.oclc.org/content/dam/research/publications/library/2012/2012-08.pdf?urlm=168716>

<sup>viii</sup> **Ricky Erway:** *Demystifying Born-Digital* in *NextSpace Issue 22* (dicembre 2013)  
<http://www.oclc.org/publications/nextspace/articles/issue22/demystifyingborndigital.en.html>

<sup>ix</sup> **Ricky Erway, Ben Goldman, Matthew McKinley:** *Agreement Elements for Outsourcing Transfer of Born Digital Content*  
<http://www.oclc.org/content/dam/research/publications/library/2014/oclcresearch-born-digital-content-transfer-2014-a4.pdf>

<sup>x</sup> **Innovative solutions for dealing with Born digital contents in obsolete formats**  
Webinar: <http://www.oclc.org/research/events/2014/10-20.html>

<sup>xi</sup> **Roberto Cotroneo:** *Vuoi salvare qualcosa? Non fidarti di Internet*. In: *Sette*, 27 febbraio 2015, p. 9.