

# Prospettive per l'intermediazione bibliotecaria

**Dal documento ai dati**

# Cos'è un documento?

***«I documenti sono oggetti che portano segni»***

*Alfredo Serrai*

# Documenti e intermediazione bibliotecaria

- **La funzione della biblioteca è favorire l'incontro tra lettori e documenti**
- **Per raggiungere questo scopo, è necessario:**
  - **Per ogni documento produrre una scheda che lo descriva e identifichi in maniera inequivocabile (catalogazione)**
  - **Rendere il documento accessibile ordinandolo in base a dei criteri determinati (indicizzazione)**

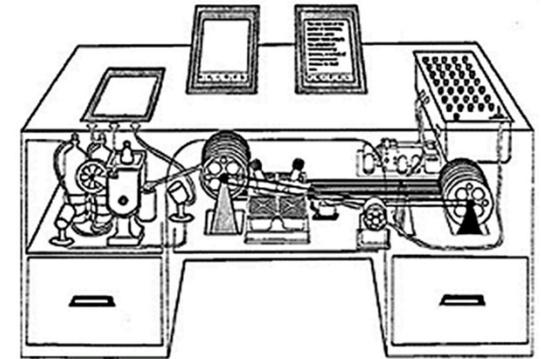
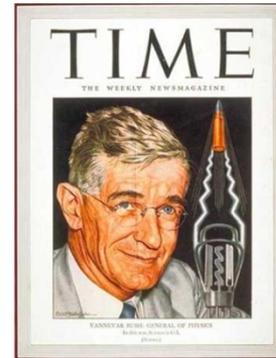
Il documento diventa  
“sottoprodotto”  
dell'evoluzione tecnologica

***“i documenti sono solo sottoprodotti  
occasionalisti dell'accesso all'informazione e non  
la sua principale incarnazione”***

**Lewis M. Branscomb**

# Il Memex di Vannevar Bush

- Il Memex (*memory expansion*) era un calcolatore analogico dotato di un sistema di archiviazione, ideato negli anni trenta e mai realizzato, considerato il precursore del computer e degli ipertesti
- Bush descrisse il memex nel suo saggio *As We May Think* (lett. "Come potremmo pensare"), pubblicato nel luglio 1945
- Vannevar Bush era partito da un'esigenza concreta, dalla ricerca della soluzione di un problema specifico:
  - come rendere più efficiente l'archiviazione ed il reperimento del sapere, dato che l'informazione, nelle università, organizzata in biblioteche risultava spesso di difficile o impossibile accesso



# Il documento come sottoprodotto dell'informazione

- Con l'evoluzione tecnologica si afferma una nuova tendenza di lungo periodo nella concettualizzazione e trattamento dei documenti
- Dal Memex di Vannevar Bush fino all'invenzione del Web acquisisce sempre più importanza il concetto di "Rete"
- la dimensione connettiva prevale nella concettualizzazione del sistema informativo
- Di conseguenza il documento è riposizionato sempre di più in subordine rispetto al concetto (tecnologicamente) più scattante e flessibile di informazione



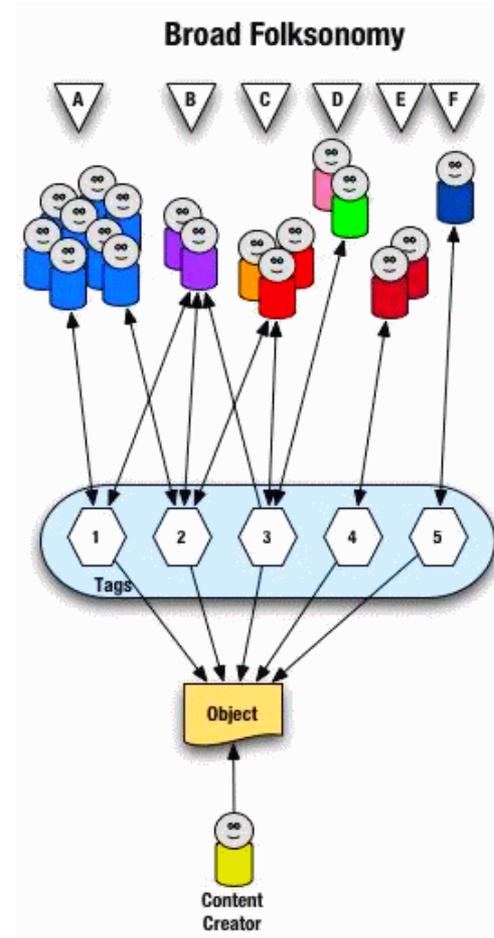
# La narrazione delle “*biblioteche obsolete*”...rispetto alla nuova organizzazione del sapere

- Da più parti si avanza l'immagine di una progressiva estinzione o riduzione del ruolo culturale e informativo delle biblioteche: risultato di un esito ineluttabile indotto dall'evoluzione tecnologica
  - Segno del declino: il fenomeno della **disintermediazione**
- Di conseguenza il modello di organizzazione del sapere proposto dalle biblioteche appare superato
  - perché orientato verso un ideale di conoscenza unitaria e convergente
  - Mentre l'evoluzione tecnologica impone piuttosto un approccio performativo, rivolto alla capacità di mantenere delle connessioni piuttosto che dei contenuti



# La narrazione delle *"biblioteche obsolete"...*incapaci di catalogare il *NUOVO* sapere

- Accusate di essere rigide, ingessate nella pervicace applicazione di tassonomie calate dall'alto
- Incapaci di far emergere il nuovo potenziale informativo interagendo con la Rete, vale a dire:
  - Tag
    - Lo standard di fatto per la classificazione dei contenuti in Internet
  - Folksonomia
    - l'operazione di categorizzare informazioni compiuta dagli utenti mediante l'utilizzo di parole chiave (o tag) scelte liberamente



# Ma, la profezia della morte del documento si sta davvero avverando..?

***Perché, al contrario, si sono cominciate a cercare soluzioni per garantire, anche nel contesto digitale, quelle condizioni di validazione autoriale, di datazione, di integrità ereditate dal mondo dei documenti analogici?***



# Siamo comunque nell'era dei Dati "Caotici"

- Secondo stime, circa l'80% dei dati attualmente gestiti dalle organizzazioni non è strutturato in colonne e righe ben definite. Si tratta di una valanga confusa di messaggi di posta elettronica, feed di social media, immagini da satellite, segnali GPS, log di server e altri file non strutturati e non relazionali.
- Secondo altre stime, soltanto il 5% di tutti i dati digitali sarebbe strutturato, cioè organizzato in maniera da inserirsi agevolmente in un database tradizionale
- Avendo ingenti quantità di dati di diversi tipi e vari gradi di qualità, difficilmente questa massa di informazione si può inquadrare in categorie precise – come accadeva nell'era dei "small data"



# La prima informatizzazione dei documenti: il formato MARC

- **l'informatizzazione delle schede catalografiche**
- **Un successo dovuto alla combinazione di due fattori:**
  - **capacità di ottimizzare la scarsa (e costosa) memoria di sistema del tempo**
  - **gestione non “troppo complicata” delle informazioni trasferite su supporto informatico**
- **formato molto strutturato – in etichette da tre cifre – ma soprattutto rigidamente chiuso**

## MARC Record is a Machine Readable Cataloging record

---

```
Leader          01041cam  2200265 a 4500
Control No.     001 ###89048230 /AC/r91
Control No. ID  003 DLC
DTLT           005 19911106082810.9
Fixed Data      008 891101s1990   maua   j           001 0 eng
LCCN            010 ## $a ###89048230 /AC/r91
ISBN           020 ## $a 0316107514 :
                $c $12.95
ISBN           020 ## $a 0316107506 (pbk.) :
                $c $5.95 ($6.95 Can.)
Cat. Source     040 ## $a DLC
                $c DLC
                $d DLC
```

---

# MARC must die: quel muro che separa il catalogo bibliografico dal Web

- **MARC must die**
  - E' il 2002, quanto il bibliotecario-tecnologo Roy Tennant pubblica il famoso articolo: "[Marc Must Die](#)".
  - Da quel momento, per il mondo delle biblioteche diventa chiaro
  - che è necessario un cambiamento, in quanto il MARC è
    - Un Formato strutturato e chiuso
    - I records bibliografici sono isolati rispetto al Web

# Il record bibliografico: verso la decostruzione

- **Fare cadere il muro che separa il catalogo bibliografico dal Web**
- **Dalle descrizioni: ISBD**
- **Alle relazioni: FRBR**
- **Passando poi a linee guida e istruzioni per la formulazione dei dati per la scoperta delle risorse nel mondo digitale : RDA**
- **E per il modello dati: BIBFRAME**
- **fino alla definitiva trasformazione dei documenti in dati: Web semantico e Linked Data**

# Le descrizioni: ISBD

(International Standard Bibliographic Description)

- **Nato dalla necessita di uniformare su scala internazionale la descrizione bibliografica (IFLA)**
- **E' basato su una descrizione organizzata che stabilisce:**
  - **Il contenuto ( quali informazioni vanno inserite)**
  - **L'ordine (in quale successione)**
  - **Il modo di presentare i dati (da quali fonti, in che lingua, la punteggiatura ecc.)**

# Le relazioni: FRBR

(Functional Requirements for Bibliographic Records)

**Schema sviluppato dalla International Federation of Library Associations and Institutions (IFLA)**

- **Realizzato tramite modello concettuale entità-relazioni**
  - **le entità sono cose specifiche di interesse per l'utente (opera, espressione, manifestazione, item, persona, ente ecc.)**
- **le relazioni sono tra entità**
  - **Le relazioni consentono all'utente, mediante una ricerca, di individuare i collegamenti tra le entità e di navigare tra i vari record**

# Le relazioni: FRBR

(Functional Requirements for Bibliographic Records)

**Questo modello concettuale ha cambiato il modo di concepire il record bibliografico**

**Spostando l'approccio dall'oggetto al contenuto della risorsa catalogata andando oltre la materialità dei documenti**

**Aprondo due questioni di primaria importanza:**

- **regole di catalogazione inadeguate al mutamento dell'universo bibliografico**
  - caratterizzato da risorse estremamente diversificate nei contenuti, nei supporti, nelle modalità di diffusione e di accesso
- **cataloghi elettronici inadeguati**
  - non in grado di rappresentare l'universo bibliografico in modo ottimizzato
    - ovvero avvantaggiandosi delle enormi potenzialità offerte dai sistemi informatici in termini di efficacia e di economia

# FRBR: navigare tra i dati

- **Necessità di Navigare tra i dati bibliografici**
- **«Significa muoversi liberamente in una base dati bibliografica, cioè trovare opere correlate a una determinata opera tramite generalizzazione, associazione o aggregazione; trovare attributi collegati tramite equivalenza, associazione, gerarchia» (Elaine Svenonius)**
- **Per realizzare la funzione «navigare» i dati che rappresentano gli attributi delle identità FRBR devono essere connessi tra loro (Linked) in un reticolo**
- **Per la soddisfazione degli utenti il reticolo dei dati bibliografici deve estendersi fuori dall'ambito bibliografico e aprirsi al Web**

# Gestire le risorse del mondo digitale: RDA (Resource Description and Access)

- **Un nuovo codice per la descrizione e l'accesso alle risorse progettato per il mondo digitale e rilasciato nel 2010**
- **Rispondere alla necessità di uno strumento condiviso da tutte le comunità del mondo che si occupano di organizzare e gestire le collezioni più varie**
- **Gli obiettivi di RDA nel contesto delineato da FRBR:**
  - **L'utente al centro della scoperta delle risorse**
  - **Adattare i cataloghi delle biblioteche al Web**
  - **Metadati di qualità per l'integrazione tra biblioteche e altre comunità (editori, aggregatori, librai ecc.)**
  - **Facilitare la ricerca di informazioni nel contesto delle attuali tecnologie**

# L'utente al centro della scoperta delle risorse

RDA fornisce un insieme di linee guida e istruzioni per la registrazione dei dati, per favorire la scoperta delle risorse. I dati creati utilizzando RDA per la descrizione una risorsa sono progettati per assistere gli utenti che svolgono le seguenti funzioni:

***trovare*** — cioè trovare risorse che corrispondono ai criteri di ricerca stabiliti dall'utente ***identificare*** — cioè confermare che la risorsa descritta corrisponde alla risorsa cercata, o distinguere tra due o più risorse con caratteristiche simili

***selezionare*** — cioè scegliere una risorsa appropriata ai bisogni dell'utente

***ottenere*** — cioè acquisire o avere accesso alla risorsa descritta. I dati creati utilizzando RDA per descrivere un'entità associata a una risorsa (una persona, famiglia, ente, concetto, etc.) sono progettati per assistere gli utenti che svolgono le seguenti funzioni:<sup>2</sup>

***trovare*** — cioè trovare informazioni su quell'entità e sulle risorse a essa associate ***identificare*** — cioè confermare che l'entità descritta corrisponde a quella cercata, o distinguere tra due o più entità con nomi simili

***chiarire*** — cioè chiarire le relazioni tra due o più entità simili, o chiarire le relazioni tra l'entità descritta e un nome con il quale essa è conosciuta

***comprendere*** — cioè comprendere perché un particolare nome o titolo, o una particolare forma del nome o del titolo, è stato scelto come nome o titolo preferito per quell'entità.

RDA fornisce un insieme completo di linee guida e istruzioni che copre ogni tipo di contenuto e di media.

# FRBR – RDA: rapporto strettissimo

**Il trattamento delle risorse RDA comprende tutte le entità del Gruppo 1 di FRBR:**

- **Opera**
- **Espressione**
- **Manifestazione**
- **Item**

**Le 2 funzioni principali di RDA sono gli attributi di un'entità:**

- **Identificare**
- **Registrare**

# RDA: Standard di contenuto

**Approccio aperto alla tecnologia per integrare il mondo delle biblioteche con il Web**

- **Per raggiungere questo obiettivo, RDA è stato progettato come**
  - **Standard di contenuto**
    - **mantenere lo stesso significato anche cambiando formato**
    - **Infatti, non si occupa della visualizzazione dei dati come facevano i codici precedenti**
    - **Fornisce invece istruzione su come identificare i dati**
    - **Ma non spiega come e dove presentare i dati individuati, selezionati e raccolti in base alle linee guida**

# RDA: la porta di accesso al Web semantico

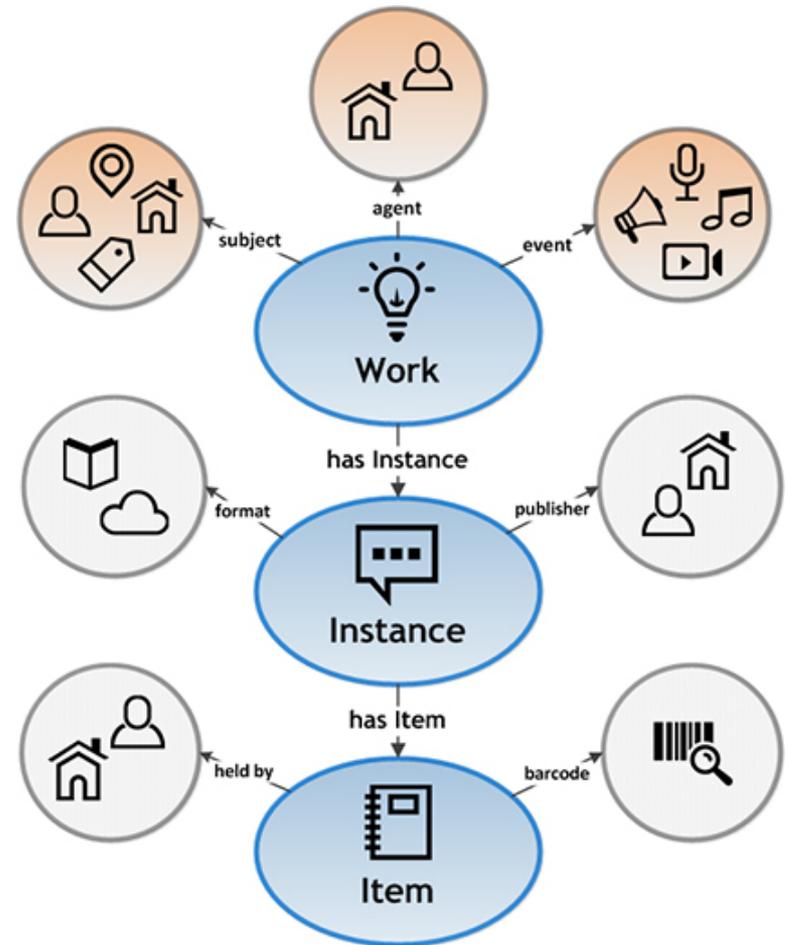
*«Il lavoro da svolgere di qui in avanti per far crescere RDA e seguirlo nella sua evoluzione consiste da una parte collaborare alla costruzione degli strumenti necessari, primi fra tutti vocabolari e ontologie, dall'altra nella conversione dei record che si trovano nei cataloghi delle biblioteche, archivi e musei in formati compatibili con RDA che rappresenta la porta di accesso al web semantico».*

**Gordon Dunsire**

**Responsabile del comitato direttivo per lo sviluppo di RDA  
(postfazione, Introduzione a RDA)**

# BIBFRAME: un progetto di modello dati per una comunità bibliotecaria aperta al Web

- La Library of Congress lo ha lanciato nel 2012
- È un modello dati per le descrizioni bibliografiche progettato per sostituire i formati MARC e utilizzare i dati collegati
- Principi:
  - Distinzione contenuto concettuale / manifestazione fisica di un'opera
  - Rappresentazione delle relazioni esistenti tra entità per favorirne l'uso in termini di navigabilità



# BIBFRAME: un progetto di modello dati per una comunità bibliotecaria aperta al Web

**Offre anche un modello alternativo per metadati dal basso**

- **Non solo creati da professionisti e sistemi di automazione**
- **Ma creati anche da utenti e derivanti da annotazioni, recensioni e informazioni sull'utilizzo delle risorse**

# roadmap BIBFRAME

**Ex Libris, società globale per l'automazione delle biblioteche**

**Supporta le seguenti funzioni per il [passaggio a BIBFRAME](#):**

- **Trasformazione da MARC a BIBFRAME**
- **Pubblicazione di set di dati in formato BIBFRAME**
- **Supporto per le URI BIBFRAME**
- **La trasformazione MARC a BIBFRAME, prima fase della roadmap pianificata, sarà realizzata in stretta collaborazione con la biblioteca di Harvard**