

UNIVERSITÀ DEGLI STUDI DI TORINO

DIPARTIMENTO DI STUDI UMANISTICI

Corso di Laurea Specialistica in Storia e Critica delle Culture e dei Beni Musicali



**La catalogazione degli strumenti musicali
tra *open knowledge* e *semantic web***

Studio per un'implementazione in XML Schema
della Scheda Guizzi

Relatore
Ilario Meandri

Candidato
Giorgio Bevilacqua

Anno Accademico 2015 - 2016

In ricordo del prof. Febo Guizzi

Dedicata ai miei genitori
che mi hanno sempre sostenuto in questo percorso

Indice generale

Introduzione

Le 250 schede del Museo del Paesaggio Sonoro.....	5
---	---

Prima parte - *Open knowledge* e interoperabilità: tecnologie e politiche

Lineamenti di cultura *open*

I concetti di apertura e interoperabilità.....	14
--	----

<i>Free software</i> e <i>open source</i> : etica, mercato e licenze d'uso.....	21
---	----

La cattedrale e il <i>bazaar</i>	26
--	----

Accesso libero alla letteratura scientifica.....	30
--	----

La centralità del formato XML

Diffusione e pregi.....	37
-------------------------	----

Documenti XML.....	39
--------------------	----

Schemi XML.....	43
-----------------	----

I LOD (*Linked Open Data*) e il ruolo delle istituzioni culturali nel *semantic web*

Web di documenti e web di dati.....	46
-------------------------------------	----

URI e RDF per l'identificazione dei concetti e il collegamento dei dati.....	49
--	----

Vocabolari e ontologie.....	55
-----------------------------	----

Musei, archivi, biblioteche. La questione dell'affidabilità delle fonti.....	58
--	----

Seconda parte - MIMO. Nuove tecnologie per la catalogazione di strumenti musicali

Il progetto MIMO

Obiettivi e strategie.....	68
----------------------------	----

La disseminazione online.....	72
-------------------------------	----

La piattaforma tecnica e i protocolli di trasmissione.....	75
--	----

LIDO: lo schema per la strutturazione dei metadati

Il modello e l'applicazione all'ambito degli strumenti musicali.....	79
--	----

<i>Unwrapping</i> lido-v1.0.xsd.....	83
--------------------------------------	----

MIMO data model organization.....	85
-----------------------------------	----

Verbosità e granularità.....	101
------------------------------	-----

MIMO, Europea e i LOD	
<i>Thesauri</i> e classificazioni.....	104
Il <i>metadata enrichment</i>	108
CIDOC-CRM: un'ontologia <i>event-based</i>	
Il Conceptual Reference Model.....	111
Il modello <i>object-oriented</i>	116
L'impostazione evento-centrica.....	121
Il <i>top-level</i> del CRM.....	123
Aspetti critici e modelli alternativi.....	127

Terza parte - La Scheda Guizzi e il nuovo contesto tecnologico

Il modello di una teoria organologica	
Storia, funzioni, adattamenti.....	136
Il concetto di <i>oggetto sonoro</i>	139
La precarietà dell'incontro con l'oggetto: tre motivi.....	142
La digitalizzazione della scheda	
La prima implementazione in XML Schema (versione 1.0).....	145
La mappatura Guizzi-MIMO.....	152
Linee-guida per successive implementazioni.....	154
Il problema delle <i>keywords</i> e della sistematica Hornbostel-Sachs.....	162

Conclusioni e sviluppi futuri.....	169
---	------------

Appendice

1. La Scheda Guizzi.....	174
2. scheda guizzi 1_0.xsd.....	185
3. Istruzioni per una mappatura Guizzi-MIMO.....	195
4. Esempio di record Guizzi (versione 1.0).....	200
5. Esempio di record MIMO.....	202

Bibliografia.....	214
--------------------------	------------

Introduzione

Le 250 schede del Museo del Paesaggio Sonoro

Il presente lavoro nasce dall'intenzione di valorizzare la collezione del Museo del Paesaggio Sonoro di Riva presso Chieri¹; essa è il frutto del lavoro di ricerca su campo condotto da Domenico Torta, musicista, compositore, insegnante e ricercatore di origine rivese, che a partire dagli anni '80 ha esplorato con costanza, profondità e curiosità il paesaggio sonoro dei suoi luoghi d'origine. La collezione che ne risulta è un insieme multiforme di *oggetti sonori*: richiami da caccia, strumenti giocattolo, strumenti propriamente detti e dispositivi acustici in generale, offrono una ricca panoramica della cultura sonora rivese. Essa dimostra l'esistenza di competenze culturali diffuse relative alle proprietà acustiche dei materiali, alle tecniche costruttive degli strumenti, alle tecniche esecutive e all'organizzazione formale e significativa dei suoni; inoltre lo studio condotto da Torta sugli usi e i repertori connessi a tali strumenti evidenzia non solo che un'*altra musica*² è presente al di fuori della cornice formale del "fare musicale", ma che gli scambi culturali tra i due ambiti (la competenza diffusa del paesaggio sonoro e quella specialistica del "fare musicale") sono più intensi di quanto possa apparire a prima vista.

Presentata ripetutamente solo in occasione di mostre temporanee, la collezione ha trovato una sistemazione stabile nel 2011, nell'ambito dell'istituzione di un museo permanente nelle sale di Palazzo Grosso a Riva presso Chieri. Questo risultato è stato raggiunto anche grazie all'interessamento dell'Università di Torino³, e di Febo Guizzi in particolare, che ha trovato nel paesaggio sonoro rivese importanti conferme ma anche stimolanti integrazioni alle ricerche da lui condotte sullo strumentario popolare italiano e soprattutto sul cosiddetto "strumentario minore"⁴.

Una parte della collezione (circa 250 su un totale approssimativo di 1000 strumenti) è stata catalogata nel 2007, nell'ambito di un progetto finanziato dalla Regione Piemonte; per la compilazione delle schede catalografiche è stato adottato il modello ideato da Febo Guizzi e già

1 www.museopaesaggiosonoro.org [ultimo accesso: 16 settembre 2016; si consideri che tutti gli URL riportati nella tesi sono stati consultati l'ultima volta il 16 settembre 2016].

2 Il concetto di *altra musica* è elaborato in Leydi (2008); l'edizione del testo di Leydi, uscito in prima edizione nel 1991 e poi nel 2008 in una edizione curata da Febo Guizzi, è preceduta da un'*Introduzione* ad opera dello stesso Guizzi, della quale riportiamo un passo: «[...]l'alterità è anche dentro noi, attraversa territori apparentemente ben protetti dalle frontiere della cultura scritta della musica d'arte delle classi dominanti, dell'apparato dell'industria culturale in tutta la sua potenza, dalle espressioni più elitarie a quelle della cosiddetta cultura di consumo» (Leydi, 2008, p. 9).

3 Il progetto del Museo è stato elaborato da Guido Raschieri e Domenico Torta.

4 Cfr. La Vena (1996, 2001) e Guizzi (2002).

sperimentato “sul campo” nel lavoro di catalogazione svolto per diversi musei e collezioni private. Le 250 schede del Museo del Paesaggio Sonoro sono state originariamente compilate sotto forma di semplici file di testo. Questo lavoro è dedicato alla digitalizzazione del modello elaborato da Guizzi, in funzione di un suo aggiornamento alle più moderne tecnologie catalografiche.

La ricerca si è da subito orientata verso l’implementazione della scheda nel linguaggio XML: esso rappresenta il formato digitale stabile⁵ più utilizzato nella gestione dell’informazione strutturata, ed è quindi particolarmente adatto per i progetti di metadattazione. Ma la centralità di XML va ben oltre la sua capacità di ordinare i metadati in strutture ad albero funzionali ed agili: esso si è rivelato il formato più diffuso nella serializzazione di altri formati utilizzati dalle tecnologie del *semantic web*, di cui costituisce il sub-strato sintattico.

Un’analisi approfondita dei linguaggi informatici è fuori dallo scopo di questo lavoro⁶. La ricerca ha individuato MIMO (Musical Instrument Museums Online⁷) come il progetto internazionale di maggior rilievo e visibilità che utilizza XML per la gestione e pubblicazione dei metadati relativi agli strumenti musicali.

I caratteri più innovativi delle politiche di gestione dei metadati adottate da MIMO riguardano l’implementazione di tecnologie relative a *Linked Open Data* (LOD) e *semantic web*⁸. L’interesse suscitato da queste tecnologie nell’ambito delle cosiddette “istituzioni della memoria” (musei, archivi e biblioteche)⁹ è cresciuto esponenzialmente negli ultimi anni; l’aspetto più interessante è legato al fatto che, per una realizzazione efficace del *semantic web*, il mondo dell’informazione digitale dichiara di aver bisogno delle competenze e delle informazioni strutturate generate dalle istituzioni della memoria, alle quali è espressamente richiesta una partecipazione attiva e costruttiva che prefigura un’estensione del ruolo delle attività di catalogazione e metadattazione.

Il confronto tra la prassi consolidata della catalogazione organologica e i metodi di sviluppo dei LOD e del *semantic web* evidenzia però l’esistenza di alcuni importanti scarti metodologici; essi vanno ricondotti alla presenza, nei più recenti sviluppi tecnologici, di caratteristiche che discendono

5 Il linguaggio XML 1.0 è divenuto una raccomandazione del W3C nel 1998; da allora la versione 1.0 ha subito lievi revisioni che hanno portato alla quinta edizione, pubblicata nel 2008. Una versione semplificata di XML (la 1.1) è stata pubblicata nel 2004, ma il suo utilizzo è poco diffuso.

6 Alcuni cenni tecnici al linguaggio XML sono contenuti nel capitolo “La centralità del formato XML”. Esso presenta un breve inquadramento storico del formato e l’insieme delle scelte terminologiche adottate nel presente lavoro. Per un’analisi approfondita delle caratteristiche del linguaggio si rimanda alla letteratura specifica.

7 MIMO (Musical Instrument Museums Online) - <http://www.mimo-international.com/MIMO/>

8 Tale implementazione è stata resa possibile anche grazie al supporto di Europeana, la biblioteca digitale europea su cui MIMO ha caricato i propri dati. - <http://europeana.eu/>

9 Il termine “memory institutions” è stato introdotto da Dempsey (2000) per sottolineare la convergenza di obiettivi e strategie che accomuna musei, archivi e biblioteche nelle politiche di digitalizzazione dei contenuti. Il suo uso si è diffuso sia nella documentazione relativa alla progettazione sia nella letteratura scientifica dedicata alla riflessione sulle nuove tecnologie di catalogazione e sul *semantic web*.

dal più generale movimento *open source* e *open content*. Il riconoscimento di questo passaggio è essenziale per una corretta definizione del ruolo delle istituzioni culturali nel nuovo scenario tecnologico.

La tesi è articolata in tre parti più una serie di appendici. La prima parte, ***Open knowledge e interoperabilità: tecnologie e politiche***, è dedicata all'individuazione del concetto di *apertura* e all'analisi della sua applicazione nei diversi contesti del movimento *open*.

Il concetto di apertura è strettamente connesso al processo di digitalizzazione dei contenuti: un requisito fondamentale per stabilire se un contenuto è aperto, è la sua codifica in un formato digitale. Di conseguenza tutto il mondo *open* manifesta un profondo legame con alcune correnti culturali della scienza informatica (quali ad esempio il *free software* e l'*open source*), il cui paradigma è stato di fondamentale ispirazione nella definizione stessa del concetto generale di *open knowledge* (cfr. *infra* cap. "I concetti di apertura e interoperabilità").

Tali constatazioni rendono più evidente il fatto che anche le particolari novità strategiche e metodologiche introdotte nel mondo dell'*open content* (sia quelle motivate da un principio etico di accessibilità dei contenuti, sia quelle funzionali volte a un generale incremento di efficienza) hanno la loro principale origine nel movimento del *free software* e dell'*open source*. Qui infatti sono state messe in atto politiche innovative di *circolazione* dei contenuti, in particolare in merito ai modelli di sostenibilità economica e alla definizione di licenze d'uso connesse a un diverso modo di intendere la proprietà intellettuale. Inoltre sono stati sperimentati con successo nuovi metodi di *elaborazione* dei contenuti: la provvisorietà e la rapidità di diffusione delle versioni dei software che caratterizzano il *bazaar* dell'*open source* ha generato un sistema di programmazione collettiva che non ha eguali, quanto meno nell'ambito dell'*open content*¹⁰.

Le prospettive e le strategie introdotte dall'*open source* si stanno pertanto diffondendo in tutto il mondo dell'*open content*, in virtù del comune riferimento allo stesso paradigma informatico; ma il fatto che l'applicazione di tali strategie nell'ambito del software sia ormai prassi consolidata, non implica che simili travasi metodologici possano avvenire sempre con facilità.

Nonostante le frequenti dichiarazioni istituzionali di adesione programmatica, il movimento *open access* incontra difficoltà e resistenze nel proporre le sue soluzioni innovative. I nuovi modi di intendere la proprietà intellettuale, la valutazione bibliometrica e scientometrica ed il rapporto con forme di pubblicazione inusuali o innovative (*pre-print*, *post-print* ecc.) devono scontrarsi con

¹⁰ Va inoltre aggiunto che l'efficacia di tali politiche ha il suo fondamento in una generale cornice di interoperabilità e condivisione delle tecniche e dei saperi.

l'inerzia del sistema di produzione della letteratura scientifica che, se da una parte è motivata da osservazioni di carattere scientifico (per esempio in relazione alle garanzie di qualità delle pubblicazioni *open access*), dall'altra è il risultato dell'intreccio di dinamiche che hanno profonde radici storiche; a questo proposito, bisogna considerare che l'influenza di forti interessi economici contribuisce a rallentare l'evoluzione delle abitudini di pubblicazione.

Nella tecnologia dei *Linked Open Data* il paradigma informatico dell'apertura si evolve qualitativamente ed acquista una valenza ancora maggiore. Negli *open data* esso è applicato esclusivamente al tema dell'*accessibilità* dei dati: è solo grazie alla digitalizzazione dei contenuti e all'istituzione di una rete telematica di collegamento che essi possono essere liberamente¹¹ consultati, aggregati, modificati e ridistribuiti. I "*Linked*" *open data* estendono il paradigma informatico all'ambito dei *collegamenti* tra dati, i quali assumono una connotazione semantica, nel senso che il loro significato, espresso in un formato *machine-readable*, diventa "comprensibile dalla macchina".

Il *semantic web*, di cui i LOD sono il presupposto tecnico, implica il passaggio da una concezione tradizionale della rete (il *web di documenti*) al cosiddetto *web di dati*: un ambiente in cui l'informazione, atomizzata in concetti minimi (i dati) e liberamente riaggregabili *on the flow*, è strutturata in relazioni semanticamente connotate. Se nel *web di documenti* solo l'operatore umano sceglie quale collegamento ipertestuale aprire e come interpretare il documento recuperato, nel *web di dati* si apre un nuovo scenario in cui anche la macchina è in grado di interpretare i dati e fornire risposte che inferiscono nuova conoscenza.

Del *semantic web* si fornisce una presentazione teorica e tecnica nei capitoli dedicati; la loro lettura è importante per avere la misura di molti degli argomenti che vengono discussi successivamente. Qui ci preme sottolineare che esso estende l'applicazione del paradigma informatico contestualmente all'espansione del concetto stesso di web; originariamente inteso come una rete di collegamento tra dispositivi elettronici, ora il *semantic web* comprende la rete di collegamenti semantici che c'è tra le cose della vita reale, la cui rappresentazione è affidata a descrizioni fornite da aggregazioni di dati in rete. A musei, archivi e biblioteche è richiesto di assumere un ruolo di primo piano nella definizione del *semantic web* e nel controllo dell'affidabilità delle informazioni: essi possono fornire le proprie competenze per operare una corretta atomizzazione delle informazioni, individuare ed ordinare i concetti fondamentali (i vocabolari) ed elaborare i modelli di

¹¹ Anche il fatto che ciò possa avvenire *liberamente* deriva dal modello di sviluppo dei software, in particolare dall'applicazione di specifiche licenze d'uso basate su una concezione meno restrittiva della proprietà intellettuale.

rappresentazione (le ontologie) del dominio di pertinenza. Per definire con chiarezza il proprio ruolo in un contesto tecnologico così innovativo le istituzioni della memoria devono saper riconoscere che le tecnologie e le politiche di LOD e *semantic web* vanno anch'esse commisurate al paradigma informatico dell'apertura: molti comportamenti, dettati da esigenze, strategie ed obiettivi particolari, risulterebbero altrimenti inspiegati.

L'adozione di questo paradigma, sia nel campo della ricerca accademica che di quella più legata al patrimonio delle istituzioni museali, richiede la revisione se non l'abbandono di prassi consolidate: i repentini slittamenti di responsabilità generano "vuoti di potere" e una conseguente sensazione di perdita di controllo (dei percorsi, degli strumenti, delle garanzie ecc.) a cui si può far fronte anche aderendo a politiche culturali coraggiose, sperimentali e dall'esito imprevedibile.

La seconda parte, **MIMO. Nuove tecnologie per la catalogazione di strumenti musicali**, è un'analisi del progetto MIMO e delle principali tecnologie di cui esso si avvale. La ricerca è stata svolta tramite lo studio della documentazione ufficiale prodotta da MIMO, la consultazione dei canali di disseminazione (soprattutto i siti web) e la partecipazione ad una *training school* organizzata da MIMO e WoodMusick/COST presso l'Istituto Fryderyk Chopin di Varsavia nell'aprile 2016. L'attenzione si è concentrata sulla gestione dei metadati e sull'implementazione delle tecnologie dei *Linked Open Data*.

Va innanzitutto segnalato che la rapidità con cui MIMO ha pubblicato i propri dati è conseguenza anche dell'adozione di metodi tipici del paradigma informatico¹²: i dati prodotti in seguito all'acquisizione digitale dei contenuti (fotografie, registrazioni audiovisive e record catalografici di strumenti musicali) sono collocati nel *framework* concettuale di riferimento e pubblicati in rete in tempi brevi, non per forza in una forma definitiva, dal momento che sia essi che il *framework* possono essere sottoposti ad ulteriori raffinamenti. In altre parole, i dati sono pubblicati in rete non solo in funzione di una "presentazione al pubblico" dei risultati raggiunti, ma anche come momento conoscitivo nell'ambito di un preciso *iter* progettuale: solo in seguito all'aggregazione dei dati e alla raccolta del *feedback* generato dagli utenti è possibile individuare le carenze e i punti di forza della struttura, ed agire di conseguenza per un suo rafforzamento¹³.

12 La fase iniziale del progetto MIMO, finanziata in parte dal programma *eContentplus* dell'Unione Europea, si è svolta dal 1 settembre 2009 al 31 agosto 2011; in questo periodo MIMO ha pubblicato materiale (metadati, registrazioni audiovisive e almeno un'immagine per strumento) relativo a circa 50000 strumenti musicali provenienti da 21 collezioni europee. Il progetto è tuttora attivo, ed è possibile aderirvi versando un contributo economico.

13 Anche se può risultare un'osservazione secondaria e ovvia, segnaliamo che se i dati non fossero stati pubblici, una ricerca come quella presentata in questo lavoro (che intende *anche* fornire un *feedback* sul progetto) non avrebbe avuto a disposizione materiale essenziale su cui ragionare.

Il progetto MIMO è analizzato dettagliatamente: si individuano attori coinvolti, obiettivi, strategie e politiche. In seguito l'attenzione si concentra sulle tecnologie adottate. La sezione dedicata a LIDO, lo schema per la strutturazione dei metadati utilizzato da MIMO, è particolarmente ricca di osservazioni tecniche; il capitolo "MIMO data model organization" si configura come una descrizione "ragionata" del modello di strutturazione dei metadati: la sua lettura può essere affiancata alla consultazione della documentazione ufficiale prodotta da MIMO¹⁴, le cui specifiche tecniche sono contestualizzate nell'ambito di una interpretazione del contesto tecnologico di riferimento.

Si forniscono quindi gli strumenti per la comprensione di quegli aspetti di LIDO che, a prima vista, generano un po' di sconcerto: la pronunciata granularità e verbosità dello schema dei metadati e soprattutto la sua impostazione evento-centrica. Essi dipendono sia da caratteristiche insite nel linguaggio adottato (XML), sia da una scelta ontologica di base (l'ontologia CIDOC-CRM) la cui complessità ha richiesto la trattazione in un capitolo a parte. Infine si analizzano i vocabolari e i *thesauri* elaborati da MIMO¹⁵ nell'ambito delle tecnologie LOD e del processo di *metadata enrichment* (le scelte operate da MIMO sono riprese e discusse da un punto di vista critico-organologico nella terza parte della tesi, nel capitolo "Il problema delle *keywords* e della sistematica Hornbostel-Sachs").

In conclusione si evince che il progetto MIMO si è incaricato di svolgere tutte le funzioni richieste alle istituzioni della memoria per lo sviluppo del *semantic web*: l'atomizzazione dell'informazione, l'individuazione e l'identificazione dei concetti, l'elaborazione dei vocabolari e, non per ultimo, la condivisione in rete di un ampio *dataset* di metadati formalmente strutturati. Il progetto adotta politiche inclusive e collaborative per il suo ulteriore sviluppo e la definizione dei contenuti: è stato previsto un periodo di sostenibilità a lungo termine¹⁶ che, al fine di agevolare l'adesione di nuove istituzioni, ha visto lo stanziamento di fondi per il mantenimento dell'infrastruttura informatica; è fornito supporto tecnico, anche tramite l'organizzazione di workshop, alle istituzioni interessate all'adesione; elaborazione, revisione ed aggiornamento di vocabolari e *thesauri* sono improntati a uno spirito collaborativo, e tutti i *data provider* sono invitati a fornire il proprio contributo. La disponibilità dimostrata non esclude quindi che alcune critiche costruttive possano essere rivolte al modello elaborato e alle scelte fatte nell'ambito del progetto.

14 Cfr. Bailly e Le Meur (2010) e Le Meur (2010). La descrizione "ragionata" va intesa come un'interpretazione del modello proposto da MIMO; seppur rigorosa, essa non intende (né può) in alcun modo sostituire la documentazione ufficiale.

15 MIMO ha redatto tre vocabolari relativi agli strumenti musicali: una lista controllata di Instrument Makers, una tassonomia per la sistematica Hornbostel-Sachs e un *thesaurus* multilingue di *keywords* relative agli strumenti.

16 Il periodo, della durata di 5 anni, è scaduto il 31 agosto 2016.

Nella terza parte, **La Scheda Guizzi e il nuovo contesto tecnologico**, i modelli di rappresentazione del dominio degli strumenti musicali esaminati in precedenza sono messi a confronto con quello elaborato da Febo Guizzi; la natura del confronto è reciproca.

In primo luogo, il progetto di digitalizzazione della Scheda Guizzi implica necessariamente una sua revisione e successiva normalizzazione in funzione del paradigma informatico: l'adozione di una nuova tecnologia infatti non si esaurisce nella mera mappatura del contenuto ma richiede di indagare ulteriormente la natura del modello originario e verificarne l'efficacia a confronto con problemi e, soprattutto, possibilità nuove.

In secondo luogo anche le tecnologie adottate sono sottoposte a verifica: da un punto di vista puramente organologico, quindi distaccandosi dal fermento generato dall'innovazione, il paradigma informatico va considerato per quello che è: un *modello* che deve dimostrarsi in grado di rappresentare coerentemente il proprio dominio di pertinenza. Nel caso degli strumenti musicali, esso è messo di fronte al compito di rappresentare il modello complesso elaborato da Guizzi, in una sorta di *mappatura* di contenuti da un "formato" all'altro.

La teoria organologica che sta alla base della Scheda Guizzi non va circoscritta all'interpretazione del dominio degli strumenti musicali: essa è espressione di una teoria antropologica *tout court* che lega l'uomo ai suoi oggetti in una cornice culturale in cui il suono e la dimensione del "fare" (culturale e sonoro) si caricano di specifiche attribuzioni di senso. Nell'individuare l'incontro con l'oggetto come il nodo dal quale scaturiscono i significati culturali, Guizzi manifesta un particolare interesse per quei territori liminali e precari (nel caso specifico rappresentati dall'incontro con gli *oggetti sonori* e gli strumenti effimeri) nei quali l'antropologia è solita misurare le teorie e le certezze acquisite. La neutralità del web, l'impostazione etica delle tecnologie *open* e l'aspirazione olistica del *semantic web* devono garantire l'accessibilità e una giusta collocazione a simili oggetti: è proprio nei casi-limite, nell'imprevedibilità delle questioni che essi generano, che si misura l'efficacia complessiva dei modelli e il raggiungimento degli obiettivi.

La teoria organologica di Guizzi è espressione di quella competenza di dominio ricercata dalle tecnologie del *semantic web* per una definizione efficace dei propri concetti. L'esperienza maturata da Guizzi, sia nell'ambito della ricerca che nel lavoro a stretto contatto con le collezioni di strumenti musicali, ha fatto sì che egli concepisse un modello di scheda catalografica molto rigoroso, sia dal punto di vista scientifico che da quello pratico e metodologico. Come la teoria organologica, anche la scheda è imperniata sull'importanza dell'incontro tra organologo e oggetto sonoro, e sulla moltitudine di questioni generate da tale incontro. Attraverso la scheda, Guizzi non intende solo definire i campi di una corretta catalogazione, ma anche consigliare e guidare il

catalogatore in una meticolosa raccolta dei dati: un progetto di digitalizzazione che intenda proporre tale modello come esempio di competenza di dominio non deve perdere l'integrità e la ricchezza del complesso di questioni che esso solleva. In questo contesto, l'importanza attribuita da Guizzi alle classificazioni come momento conoscitivo complessivo del dominio, induce a valutare criticamente alcune scelte operate da MIMO nell'elaborazione dei vocabolari delle *keywords* e della sistematica Hornbostel-Sachs.

Se il ruolo delle istituzioni della memoria può essere quello di contrafforte alla spinta propulsiva delle tecnologie dei LOD, questo non deve indurre a stabilizzarsi su posizioni acquisite; bisogna considerare la questione da un punto di vista dinamico e dialettico nella consapevolezza che, qualsiasi sia il paradigma di riferimento, si sta comunque lavorando su *modelli* di rappresentazione che possono influenzarsi vicendevolmente e sono suscettibili di variazioni.

I capitoli dedicati alla digitalizzazione della scheda e l'**Appendice** propongono soluzioni e spunti per la ricerca, nella convinzione che l'incontro tra i due paradigmi è possibile e reciprocamente vantaggioso. Se la lezione della cultura *open* è stata recepita in modo corretto, tali proposte aspettano di essere sottoposte al vaglio del paradigma informatico: la validità della teoria deve misurarsi nella concretezza dei progetti e dell'implementazione tecnologica.

Prima parte

Open knowledge e interoperabilità: tecnologie e politiche

Lineamenti di cultura *open*

I concetti di apertura e interoperabilità

«La conoscenza è aperta quando chiunque ha libertà di accesso, uso, modifica e condivisione ad essa – avendo al massimo come limite misure che ne preservino la provenienza e l'apertura»¹.

La definizione del concetto di “apertura” data dalla Open Knowledge International² include un insieme di elementi caratteristici di tutta la cultura *open*; essi possono essere ricondotti ad un duplice criterio: un contenuto è “aperto” se soddisfa un nucleo di requisiti relativi alla sua accessibilità e alle licenze d'uso con cui è distribuito.

La «libertà di accesso» è intesa sia da un punto di vista tecnico (se la conoscenza è “tecnicamente” accessibile a tutti) che economico (se è “economicamente” accessibile a tutti); questa ambivalenza è sottolineata dall'approfondimento del concetto di “apertura” che la Open Knowledge International fornisce nell'ambito specifico degli *open data*: «I dati devono essere disponibili nel loro complesso, per un prezzo non superiore ad un ragionevole costo di riproduzione, preferibilmente mediante scaricamento da Internet. I dati devono essere disponibili in un formato utile e modificabile»³.

Il resto della definizione (le libertà di «uso, modifica e condivisione» e le restrizioni finalizzate alla preservazione di «provenienza e [...] apertura») fa riferimento alle licenze d'uso⁴. La Open Definition è infatti una derivazione della Open Source Definition, che è sostanzialmente la definizione dei criteri che una licenza software deve rispettare perché possa essere considerata *open source*⁵.

1 <http://opendefinition.org/od/2.0/it/>

2 L'Open Knowledge International (OKI), conosciuta fino al 2014 come Open Knowledge Foundation (OKF), è un'associazione no-profit che promuove l'apertura e il rilascio dei dati e incentiva la diffusione di competenze per la condivisione della conoscenza. Fondata nel 2004 da Rufus Pollock, ha elaborato la “Open definition” e svolge attività di monitoraggio su uso e interpretazione del termine. <https://okfn.org/>

3 <http://opendatahandbook.org/guide/it/what-is-open-data/>

4 La preservazione di provenienza ed apertura è una funzione caratteristica delle *open licenses*; la provenienza è garantita dalle clausole di attribuzione di paternità (ad es. la clausola “Attribution” delle licenze Creative Commons”) mentre la preservazione dell'apertura rimanda al concetto di *copyleft* e alle clausole di distribuzione di opere derivate introdotte dalla famiglia di licenze del progetto GNU (GNU/GPL, LGPL, GFDL e AGPL). Dal concetto di *copyleft* discende anche la clausola “Share Alike” delle licenze Creative Commons (cfr. *infra* pp. 23-25).

5 Gli aspetti principali della Open Source Definition (OSD) sono stati definiti da Bruce Perens nell'ambito dell'Open Source Initiative. (cfr. *infra* p. 24).

I requisiti generali del concetto di apertura possono essere applicati a numerosi ambiti che prevedono la circolazione di idee, saperi, informazioni, opere della creatività e dell'ingegno:

- si parla di *open data* e *open government data* riguardo all'apertura di dati e informazioni pubblici;
- i *Linked Open Data* sono una particolare tecnologia di collegamento di dati "aperti" nel contesto tecnologico del *semantic web*;
- *open source* (insieme al concetto precursore di *free software*) si riferisce all'apertura del codice sorgente dei software;
- *open format* fa riferimento all'accessibilità delle specifiche tecniche e a politiche di interoperabilità nell'ambito dei formati di codifica digitale;
- il concetto di *open standard* definisce pratiche di standardizzazione trasparenti e neutrali finalizzate a interoperabilità e durevolezza degli standard;
- *open license* definisce i criteri delle licenze che promuovono e preservano l'apertura dei contenuti;
- *open access* fa specifico riferimento alla libera accessibilità alla letteratura scientifica
- il termine *open content*, coniato in relazione all'apertura delle opere creative in senso stretto, fa riferimento anche all'apertura di generici contenuti (ad esclusione del software).

Ognuna di queste declinazioni stabilisce un rapporto diverso con due temi fondamentali della cultura *open*: la motivazione etica e le implicazioni funzionali.

Il principio etico è facilmente individuabile tra le motivazioni che sostengono il modello dell'*open government*, il cui esercizio prevede politiche di apertura e accessibilità ai dati governativi, in un senso assimilabile al concetto di "amministrazione trasparente" della legislazione italiana.

La trasparenza è intesa come accessibilità totale dei dati e documenti detenuti dalle pubbliche amministrazioni, allo scopo di tutelare i diritti dei cittadini, promuovere la partecipazione degli interessati all'attività amministrativa e favorire forme diffuse di controllo sul perseguimento delle funzioni istituzionali e sull'utilizzo delle risorse pubbliche⁶.

L'uso combinato dei dati pubblicati dalle amministrazioni può consentire anche la realizzazione di progetti strutturati di monitoraggio dell'attività politica.

Nell'ambito della trasparenza, progetti come il Finlandese 'tax tree' (l'albero delle tasse) e il Britannico 'where does my money go' (dove vanno i miei soldi) permettono di identificare come i soldi delle tasse dei cittadini

6 Decreto legislativo n. 97 del 25 maggio 2016. <http://www.gazzettaufficiale.it/eli/id/2016/06/8/16G00108/sg>

sono impiegati dal governo. In Canada i dati aperti hanno fatto risparmiare 3.2 miliardi di dollari in un caso di frode fiscale legato alla beneficenza⁷.

Secondo un principio etico più generale, i dati prodotti dalle pubbliche amministrazioni sono un “bene comune” e devono quindi essere resi disponibili ai cittadini in forma aperta e in tempi brevi; essi sono una risorsa che può essere utilizzata per qualsiasi scopo, non solo per un esercizio di controllo dell’amministrazione pubblica⁸.

Nell’ambito generale dell’*open content* e dell’*open source*, il rilascio dei contenuti in modalità *open* ha importanti implicazioni “funzionali”, termine con cui si vuole indicare sia l’emergere di nuove funzioni connesse all’uso dei contenuti aperti, sia il delinearsi di nuovi approcci, strategie e tecniche utili ad esercitare tali funzioni. I contenuti aperti non solo sono tecnicamente accessibili, ma sono anche tendenzialmente gratuiti; inoltre sono modificabili e ridistribuibili, in quanto soggetti a politiche di proprietà intellettuale meno restrittive. Come conseguenza “funzionale”, l’apertura degli *open government data* dovrebbe favorire «la partecipazione dei cittadini [...] nella redazione e la revisione della legislazione e delle previsioni di bilancio»⁹. Tra gli esempi di partecipazione dei cittadini alla revisione dei dati pubblici (che non comporta però una loro inclusione nel processo decisionale e redazionale) va segnalato che la pubblicazione degli *open government data* ha incentivato la diffusione di iniziative spontanee di *civic hacking*, che utilizzano i dati pubblici aperti per acquisire nuove informazioni o trovare soluzioni a problemi di carattere pubblico¹⁰.

L’applicazione del concetto di apertura in altri ambiti ha avuto effetti “funzionali” ancor più radicali. Nel caso dell’*open access* per esempio, la gratuità di reperimento dei contenuti ha condotto alla sperimentazione di nuovi modelli di sostenibilità economica; la rapidità e libertà di pubblicazione dei risultati della ricerca ha indotto a introdurre nuovi criteri di valutazione qualitativa; l’allentamento dei vincoli d’uso legati alla proprietà intellettuale ha fatto nascere metodi sperimentali di collaborazione alla produzione dei testi¹¹.

7 <http://opendatahandbook.org/guide/it/why-open-data/>

8 «Il sito danese husetsweb.dk aiuta a trovare i modi migliori di risparmiare energia elettrica in casa, inclusa la pianificazione finanziaria e la possibilità di contattare gli artigiani che potranno eseguire il lavoro. Funziona grazie al riutilizzo di dati catastali, a informazioni sugli incentivi governativi e al registro delle imprese locali. Google Translate usa l’enorme volume di documenti dell’Unione Europea, disponibili in tutte le lingue d’Europa, per allenare gli algoritmi di traduzione automatica, aumentando la precisione del servizio offerto».
<http://opendatahandbook.org/guide/it/why-open-data/>

9 <http://opendatahandbook.org/glossary/it/terms/open-government/>

10 Il *civic hacking* può assumere anche la forma “goliardica” dell’*hackathon*, maratona in cui un gruppo di *hacker*, fissato un termine temporale, si sfida nella risoluzione di un quesito relativo all’impiego dei dati. Si veda la lista degli *hackathon* organizzati dalla comunità *Spaghetti Open Data* nel corso del raduno 2016:
<http://www.spaghettiopendata.org/term/tipo-di-attivita/civic-hackathons-07052016> .

11 Per un’analisi più approfondita cfr. *infra* cap. “Accesso libero alla letteratura scientifica”.

I provvedimenti legislativi preposti, sia a livello nazionale che europeo, allo sviluppo di politiche di *e-government* e *open government* convergono sull'importanza di un tema chiave per l'accessibilità ai contenuti: l'interoperabilità.

La definizione data dall'EIF (European Interoperability Framework) version 1.0¹² si concentra principalmente sull'aspetto tecnico del concetto di interoperabilità:

Interoperability means the ability of information and communication technology (ICT) systems and of the business processes they support to exchange data and to enable the sharing of information and knowledge.¹³

La successiva definizione fornita dall'EIF version 2¹⁴ estende il concetto di interoperabilità ad una dimensione più ampia che comprende

the ability of disparate and diverse organisations to interact towards mutually beneficial and agreed common goals, involving the sharing of information and knowledge between the organisations, through the business processes they support, by means of the exchange of data between their respective ICT systems.¹⁵

Sono quindi individuati quattro livelli di interoperabilità:

- Legal Interoperability (Legislative Alignment): aligned legislation so that exchanged data is accorded proper legal weight;
- Organisational Interoperability (Organisation and Process Alignment): coordinated processes in which different organisations achieve a previously agreed and mutually beneficial goal;
- Semantic Interoperability (Semantic Alignment): precise meaning of exchanged information which is preserved and understood by all parties;
- Technical Interoperability (Interaction and Transport): planning of technical issues involved in linking computer systems and services.¹⁶

L'interoperabilità tecnica è da considerarsi un requisito fondamentale per favorire l'accesso aperto ai contenuti e, come sostiene la Digital Agenda for Europe¹⁷, essa è generalmente ottenuta «through standards or shared technical specifications on file formats, API's and protocols»¹⁸.

12 L'EIF (European Interoperability Framework) version 1.0 è stata definita nel 2004 dal programma europeo IDABC (Interoperable Delivery of European eGovernment Services to public Administrations, Businesses and Citizens).

13 <http://ec.europa.eu/idabc/servlets/Docd552.pdf?id=19529>

14 L'EIF version 2 è stata definita dall'ISA (Interoperability Solutions for European Public Administrations) il programma europeo successore dell'IDABC.

15 http://ec.europa.eu/isa/documents/isa_annex_ii_eif_en.pdf

16 http://ec.europa.eu/isa/documents/isa_annex_ii_eif_en.pdf

17 <https://ec.europa.eu/digital-single-market/en> . La Digital Agenda for Europe (DAE) è la strategia per la digitalizzazione del programma Europe 2020. Il corrispettivo nazionale italiano è rappresentato dall'Agenzia per l'Italia Digitale (AgID), <http://www.agid.gov.it/> .

18 <https://ec.europa.eu/digital-single-market/en/glossary#i>

Anche il concetto di standard va sottoposto ad un'osservazione critica in relazione alle politiche di apertura e interoperabilità. Nel contesto informatico, l'individuazione di uno standard (sia *de jure* che *de facto*) rappresenta un momento di cristallizzazione di un modello relativo ad una data tecnologia, finalizzato all'ottenimento di un punto di riferimento comune per l'interoperabilità. Tale cristallizzazione deve evitare di diventare un freno all'innovazione: una volta che lo standard è ben radicato, si verifica spesso una sorta di inerzia tecnologica che rallenta l'adozione di nuove soluzioni, migliori ed innovative. Queste situazioni di stallo sono notoriamente sfruttate e ricercate da quelle aziende che, avendo assunto una posizione dominante in un settore del mercato, sfruttano commercialmente gli «effetti di rete» tipici della diffusione dei mezzi tecnologici.

Gli economisti definiscono “economie di rete” quelle forme di interdipendenza tecnologica, economica, giuridica e psicologica per le quali «l'utilità che un consumatore trae dal consumo di un bene dipende (in modo positivo o negativo) dal numero di altri individui che consumano lo stesso bene (o che lo abbiano acquistato).¹⁹» L'esempio più classico cui si può fare riferimento è quello del telefono, in cui l'utilità di possedere una linea e un apparecchio telefonico è direttamente proporzionale al numero di altri utenti che dispongono della stessa strumentazione. [...] In ambito tecnologico è molto facile che, in assenza di regole e di opportuni sistemi di monitoraggio, questi effetti di rete si trasformino in meccanismi di irrigidimento del mercato, fino ad arrivare a veri e propri casi di monopolio, passando per numerosi casi di abuso di posizione dominante. (Aliprandi, 2010, pp. 24-25)

Siccome i comitati tecnici degli enti di standardizzazione si avvalgono anche dell'importante competenza degli *stakeholder* di settore (tra cui quindi anche le aziende che hanno interessi economici legati all'impiego dello standard), si è ritenuto necessario definire i requisiti degli *open standard*, al fine di scongiurare il diffondersi di processi di standardizzazione opachi e abusivi²⁰.

Un *open standard* deve generalmente soddisfare requisiti di trasparenza del processo di standardizzazione, accessibilità alle specifiche che lo descrivono e garanzia di effettiva e durevole interoperabilità; il W3C²¹ ha elaborato una serie di requisiti che definiscono un *open standard*:

19 http://it.wikipedia.org/wiki/Economie_di_rete. [NdA]

20 Si veda in Aliprandi (2010, pp. 77-83) il resoconto del processo di standardizzazione ISO del formato OOXML avviato e concluso con successo da Microsoft. L'azienda di Redmond ha cercato in questo modo di arginare la diffusione di ODF, il formato documentale *open* (sviluppato da OASIS e Sun MicroSystem) che è divenuto norma ISO nel 2006. Sia ODF che OOXML sono basati su XML (le estensioni del formato OOXML sono: .docx, .xlsx e .pptx.).

21 Il W3C (World Wide Web Consortium) è un'organizzazione no-profit fondata da Tim Berners-Lee nel 1994, che ha lo scopo di sviluppare protocolli e linee-guida per lo sviluppo a lungo termine del Web. Il suo fondamentale impegno per l'interoperabilità è testimoniato dall'incessante attività di standardizzazione promossa. Nell'ambito dei linguaggi di *mark-up* il W3C ha definito i seguenti standard: HTML, XML, XHTML, XML Schema, XPath, XQuery e XSLT. In merito al *semantic web*: RDF, SKOS, OWL e SPARQL. Tra gli altri standard di interesse

- transparency (due process is public, and all technical discussions, meeting minutes, are archived and referencable in decision making);
- relevance (new standardization is started upon due analysis of the market needs, including requirements phase, e.g. accessibility, multi-linguism);
- openness (anybody can participate, and everybody does: industry, individual, public, government bodies, academia, on a worldwide scale);
- impartiality and consensus (guaranteed fairness by the process and the neutral hosting of the W3C organization, with equal weight for each participant);
- availability (free access to the standard text, both during development and at final stage, translations, and clear IPR rules for implementation, allowing open source development in the case of Internet/Web technologies);
- maintenance (ongoing process for testing, errata, revision, permanent access)²².

A prescindere dai processi di standardizzazione, che riguardano l'istituzione di standard *de jure*, i formati di codifica dei file sono centrali nel garantire un'infrastruttura sintattica per l'interoperabilità. La differenza che intercorre tra i formati proprietari e quelli *open* sta nel fatto che i primi non consentono l'accessibilità alle specifiche che illustrano il processo di decodifica dei file. In assenza di queste informazioni è necessario ricorrere a complicati e instabili procedimenti di *reverse engineering* per disegnare software in grado di decodificare il formato proprietario²³. Questa condizione rappresenta un forte freno all'interoperabilità.

Il terreno del software e dei formati proprietari è particolarmente adatto alla messa in atto di strategie commerciali che sfruttano in modo anti-concorrenziale gli "effetti di rete". La tendenza delle aziende che si trovano in posizione di monopolio di un settore del software è, spesso, quella di istituire un forte legame tra un formato "proprietario-nativo" e il software venduto. Quanto più il formato è "blindato", tanto più è forte il legame di fidelizzazione che unisce l'utente al prodotto e alle sue successive versioni. Questo effetto di traino formato-software genera situazioni di mercato in cui la diffusione del formato è talmente alta che esso è considerato (a torto) uno standard *de facto*; invece si tratta spesso solo del frutto di un "abuso di posizione dominante"²⁴.

generale: CSS e SVG. <https://www.w3.org/>

22 <https://www.w3.org/2005/09/dd-osd.html>

23 https://en.wikibooks.org/wiki/Reverse_Engineering/File_Formats . Questo spiega la difficoltà riscontrata talvolta nella decodifica di formati proprietari da parte di applicazioni *open source*.

24 Nel 1998 Eric Raymond, uno dei fondatori della Open Source Initiative, è entrato in possesso di tre documenti "Microsoft confidential", relazioni interne all'azienda sul fenomeno Linux e sulle possibili strategie per contrastare la diffusione dell'*open source*. I documenti, noti come *Halloween documents*, sono pubblicati e commentati sul sito di Raymond (<http://www.catb.org/~esr/halloween/index.html>).

Tra le strategie elaborate da Microsoft, Raymond individua le caratteristiche del motto "Embrace, Extend and Extinguish", espressione programmatica attribuita da più parti all'ambiente dirigenziale Microsoft; essa consiste nell'entrare nel mercato di prodotti basati su standard riconosciuti, estendere i formati tramite l'introduzione di

La differenza “tecnica” tra i formati proprietari e quelli *open* è espressione di una diversa concezione del mercato: alla logica del profitto basata sulla vendita, definita da condizioni unilaterali, di pacchetti chiusi in un contesto di segretezza industriale, il mondo del *free software* e dell'*open source* contrappone un modello sovversivo, che non rinuncia né al *business* né ai principi etici di condivisione e libertà.

caratteristiche proprietarie, e sfruttare la posizione dominante acquisita per rendere incompatibili le applicazioni concorrenti con il nuovo formato, divenuto uno standard *de facto* di proprietà Microsoft. Il motto è documentato da un rapporto del Dipartimento di Giustizia degli Stati Uniti sull'inchiesta della “guerra tra i *browser*” che ha visto lo scontro tra Microsoft e Netscape riguardo all'estensione del formato HTML.

<https://www.justice.gov/sites/default/files/atr/legacy/2006/06/01/V-A.pdf>

Free software e open source: etica, mercato e licenze d'uso

L'applicazione del concetto di apertura all'ambito generale dell'*open content* è un fenomeno relativamente recente: si tratta infatti di ottenere l'apertura di contenuti la cui gestione è stata, per tutto il Novecento, appannaggio semi-esclusivo di istituzioni o specifici gruppi sociali (come nel caso di *open data*, *open government* e *open access*).

Nel caso dell'*open source* (e quindi del suo contenuto, il “software”) questo discorso non è valido. È vero che il termine *open source* ha iniziato a diffondersi solo recentemente, in concomitanza con la fondazione nel 1998 della Open Source Initiative; ma questo evento rappresenta un momento (nel caso specifico, una scissione) nella storia di un più ampio movimento che, a partire dai primi anni '80, ha cercato di preservare i caratteri di “apertura” originari del software e di opporsi alle crescenti operazioni di “chiusura” messe in atto da alcune aziende impegnate nella sua commercializzazione²⁵.

Il software nasce, libero, nel corso degli anni '50 in stretta connessione con i primi calcolatori usati nelle università e nei centri di ricerca; il codice sorgente dei primi software distribuiti era accessibile, in modo che ogni *hacker*²⁶ potesse modificarlo per correggere *bug* o adattarlo alle proprie esigenze. Verso la fine degli anni '70, contestualmente all'emersione del concetto di *personal computer*, si sono verificati i primi casi di chiusura del codice sorgente, tramite l'applicazione di tutela del copyright²⁷. Il *boom* economico degli anni '80 ha portato i *personal computer* nelle case di una massa di persone inesperte e messe nell'impossibilità di prendere atto delle possibili alternative; il mercato del software proprietario ha fatto leva sulla diffusione di pacchetti standard difficilmente modificabili (in quanto non accompagnati dal codice sorgente) e sullo sfruttamento intensivo degli “effetti di rete” generati da politiche di segretezza industriale.

25 Dal punto di vista tecnico, il termine *open source* denota l'effettiva possibilità di accedere al “codice sorgente” del software. Il codice sorgente è il testo, scritto in un linguaggio di programmazione *human-readable*, che rappresenta gli algoritmi costitutivi di un software; essendo *human-readable*, il codice sorgente, così come è stato creato da un programmatore umano, può essere modificato. I software proprietari sono invece distribuiti solo come “codice oggetto”, ovvero istruzioni binarie *machine-readable* non comprensibili né modificabili direttamente dall'uomo.

26 L'etimologia del termine *hacker* è complessa, ma va comunque fatta risalire al termine *hack* con il quale, nel gergo del MIT, si indicavano «gli scherzi elaborati che gli studenti del MIT s'inventavano regolarmente» (Levy, 1984, citato in Aliprandi, 2005, p. 26). Per essere definite tali le imprese degli *hacker* (dai primi *relè* per modellini ferroviari alla programmazione di software) dovevano dimostrare innovazione, stile e virtuosismo tecnico. Lungi dal riferirsi alla ormai diffusa accezione peggiorativa del termine, qui si intende con *hacker* «un esperto di informatica a cui piace programmare, che lo fa non con intenti di profitto ma per una sorta di irrefrenabile passione, quasi per vocazione» (Aliprandi, 2005, p. 15).

27 Il 1980 è la data emblematica di emanazione del *Software Copyright Act* da parte del legislatore statunitense, con cui si dichiara che il software va tutelato in base al *Copyright Act* del 1976. Anche in Europa il software è tutelato dalle leggi nazionali sul diritto d'autore (in Italia, legge n. 633 del 1941 e successive modifiche), essendo esso generalmente assimilato ad un “testo” espressivo di carattere creativo.

Questa evoluzione del mercato ha preso le mosse da una base tecnologica messa a disposizione da decenni di programmazione collettiva a cui hanno contribuito le prime generazioni di *hacker*, il cui intento era diffondere la cultura informatica al di fuori dei centri di ricerca. Nel 1983 si pongono le basi per un movimento di “controrivoluzione”: Richard Stallman, *hacker* delle prime generazioni, abbandona il MIT e fonda il Progetto GNU²⁸, con l’intento di definire un sistema operativo svincolato dalle logiche di tutela del copyright e quindi liberamente distribuibile; nel 1985 nasce la Free Software Foundation (FSF)²⁹, organizzazione no-profit finalizzata alla raccolta di fondi per il progetto GNU che svolge un’intensa attività di sensibilizzazione sostenuta da un’inflexibile vocazione etica.

Nell’enunciare i principi che guidano l’attività della FSF, Stallman sottolinea l’importanza del termine “free”, nella sua valenza semantica di “libertà”:

“Free software” is a matter of liberty, not price. To understand the concept, you should think of “free” as in “free speech,” not as in “free beer.” Free software is a matter of the users’ freedom to run, copy, distribute, study, change, and improve the software. More precisely, it refers to four kinds of freedom, for the users of the software:

- Freedom 0: The freedom to run the program, for any purpose.
- Freedom 1: The freedom to study how the program works, and adapt it to your needs. (Access to the source code is a precondition for this.)
- Freedom 2: The freedom to redistribute copies so you can help your neighbor.
- Freedom 3: The freedom to improve the program, and release your improvements to the public, so that the whole community benefits. (Gay, 2002, p. 43)

L’insistenza con cui Stallman cita i principi etici di “libertà” e “solidarietà”³⁰ non esclude il fatto che il software possa divenire una fonte di guadagno per i suoi sviluppatori.

Many people believe that the spirit of the GNU project is that you should not charge money for distributing copies of software, or that you should charge as little as possible—just enough to cover the cost. Actually, we encourage people who redistribute free software to charge as much as they wish or can.

Specifically, it means that a user is free to run the program, change the program, and redistribute the program with or without changes.

Free programs are sometimes distributed gratis, and sometimes for a substantial price. Often the same program is available in both ways from different places. The program is free regardless of the price, because users have freedom in using it. Non-free programs are usually sold for a high price, but sometimes a store will give you a

28 GNU’s Not Unix (GNU) - <https://www.gnu.org/>

29 Free Software Foundation (FSF) - <http://www.fsf.org/>

30 Stallman torna più volte sull’argomento nella sua dilagante attività propagandistica; l’attenzione per il principio di “libertà” si è maggiormente rinvigorita in seguito alla “scissione” che ha dato origine, nel 1998, alla Open Source Initiative. Si veda per esempio il capitolo *Why “Free Software” is Better than “Open Source”* in Gay (2002, p. 57).

copy at no charge. That doesn't make it free software, though. Price or no price, the program is non-free because users don't have freedom. (Gay, 2002, p. 65)

Non c'è una corrispondenza necessaria tra il concetto di *free/libero* e *free-gratis*: un software, pur essendo *free-software* (quindi modificabile, riproducibile e distribuibile legalmente), può avere un costo. Un software proprietario può essere acquisito gratuitamente (per esempio, tramite download), ma ciò non comporta che si possa liberamente modificare secondo le proprie esigenze o per creare un nuovo prodotto da immettere sul mercato.

Il principio di apertura del *free software* e dell'*open source* propone un nuovo modello di *business* nel mercato del software: non si guadagna più sulla vendita di pacchetti “chiusi”, ma su una gamma di servizi collegati al software (manutenzione, aggiornamento, customizzazione, corsi di formazione) che comportano guadagni meno elevati, ma ottimizzano la filiera produttiva e la qualità dei prodotti: si lavora su una piattaforma di software interoperabili, la cui stabilità è garantita da una comunità interessata al suo mantenimento a lungo termine; il risparmio ottenuto dalla disponibilità dei frutti del lavoro collettivo può essere investito in una migliore customizzazione dei prodotti.

Trovandosi ad agire in un contesto dominante di “chiusura” determinato dalla logica del copyright, il sistema produttivo del progetto GNU necessitava la definizione di strumenti legislativi appropriati che ne garantissero la sopravvivenza. Dichiarare il dominio pubblico dei software avrebbe comportato dei rischi:

Pubblico dominio [...] non vuol dire 'proprietà di tutti' intendendo 'tutti' come 'comunità organizzata', bensì intendendo 'tutti' come 'ciascuno'. Di conseguenza, ciascuno potrebbe anche attribuirsi arbitrariamente i diritti esclusivi di tutela e iniziare a distribuire il software da lui modificato come se fosse software proprietario, criptando il sorgente e misconoscendo la provenienza pubblica del software, senza che nessuno possa agire nei suoi confronti. (Aliprandi, 2005, p. 35)

Stallman ha quindi escogitato un sistema, denominato *copyleft*, che garantisce la preservazione della libertà del software, sfruttando gli stessi meccanismi del copyright dal quale cerca di svincolarsi.

Sia in ambito di *common law* anglo-americana che di diritto continentale, le leggi sul copyright e sul diritto d'autore fanno riferimento a un diritto “esclusivo” che conferisce all'autore

una possibilità d'azione (nei confronti dell'opera) preclusa a qualunque altro soggetto: si cita spesso, per mostrarne la portata, l'espressione latina 'ius excludendi alios', cioè il diritto di escludere gli altri da una

determinata sfera d'azione, la possibilità da parte del titolare di tali diritti di vietare condotte da lui non autorizzate. (*Ivi*, p. 33)³¹

Il sistema di *copyleft* consiste nel dichiarare il software sotto la tutela del copyright, per poi definire, tramite una licenza d'uso, le libertà che l'utente può esercitare e le condizioni di distribuzione che egli deve rispettare affinché la licenza sia valida³². Libertà e condizioni d'uso sono contenute nella licenza GNU General Public License, elaborata nel 1998 e giunta nel 2007 alla sua terza edizione³³. Il *copyleft* si avvale quindi di una duplice azione: da una parte il “diritto” si occupa della tutela dell'opera, dall'altra la “licenza” svolge una funzione autorizzativa e esplicita i termini e le condizioni di distribuzione:

To copyleft a program, we first state that it is copyrighted; then we add distribution terms, which are a legal instrument that gives everyone the rights to use, modify, and redistribute the program's code or any program derived from it but only if the distribution terms are unchanged. Thus, the code and the freedoms become legally inseparable. (Gay, 2002, p. 91)

Il riferimento di Stallman all'invariabilità dei termini di distribuzione è esplicitato in una fondamentale clausola della GNU/GPL, che rappresenta l'essenza del concetto di *copyleft*: chi intende ridistribuire il programma originario o una sua versione modificata è obbligato a rispettare i termini della licenza originale e quindi a garantire ai successivi utenti le stesse condizioni di utilizzo della GPL. Quindi uno sviluppatore che crea un software utilizzando parti di codice appartenenti a un software distribuito secondo licenza GPL, è tenuto a distribuire il nuovo software usando la stessa licenza; in caso contrario la GPL non è più valida ed egli può essere accusato di violazione delle norme sul copyright (che non consentono di modificare l'opera né tantomeno di ridistribuirla). Ogni software contenente codice licenziato tramite GPL deve essere ridistribuito con licenza GPL: la “viralità” del *copyleft*, se da un lato testimonia la tenacia con cui la FSF persegue il proprio scopo, dall'altra rappresenta un ostacolo per chi volesse essere “libero” di combinare *free software* con software proprietario. L'intransigenza che Stallman e la FSF dimostrano su questo ed altri punti ha causato una “scissione” all'interno del movimento; nel 1998 Bruce Perens ed Eric Raymond, due

31 «L'autore ha il diritto esclusivo di pubblicare l'opera. Ha altresì il diritto esclusivo di utilizzare economicamente l'opera in ogni forma e modo originale o derivato, nei limiti fissati da questa legge, ed in particolare con l'esercizio dei diritti esclusivi indicati negli articoli seguenti». Legge n. 633 del 1941, art. 12. [NdR]

32 Il termine *copyleft* contiene un doppio gioco di parole, che dipende dal significato attribuito al termine *left*: «un'idea di 'ribaltamento' degli stereotipi del copy-right tradizionale, se s'intende 'left' (sinistra) come contrario di 'right' (destra); ma anche e soprattutto un'idea di 'libertà d'azione' dato che 'left' è il participio passato di 'leave', cioè 'lasciare', 'permettere'» (Aliprandi, 2005, p. 35). La prima sfumatura linguistica è enfatizzata dalla dicitura “all rights reversed” preceduta dal simbolo del copyright “©” rovesciato, che a volte accompagna i lavori rilasciati con licenze *copyleft*.

33 Gnu General Public License (GNU/GPL o GPL) - <https://www.gnu.org/licenses/licenses.it.html#GPL>

figure storiche dell'informatica libera, hanno fondato la Open Source Initiative³⁴, il cui intento è mitigare le posizioni estremiste della FSF e concentrarsi sull'aspetto funzionale dell'apertura del codice sorgente³⁵ e sulla definizione (Open Source Definition³⁶) di licenze meno restrittive (ma comunque ispirate per molti versi alla “pionieristica” GPL).

Le clausole della GPL sono state riprese anche nell'elaborazione di licenze per la libera distribuzione di opere non software; la prima licenza in questo senso è la GNU Free Documentation License (GFDL o FDL), elaborata dalla FSF nel 2000 soprattutto per gestire la documentazione relativa al software. L'apparato di licenze più importanti nell'ambito *open content* è quello fornito dalla Creative Commons Corporation³⁷, fondata da Lawrence Lessig nel 2001. Le licenze Creative Commons risultano dalla combinazione di 4 clausole (versione 4.0):

- **Attribuzione (BY):** permette che altri copino, distribuiscano, mostrino ed eseguano copie dell'opera e dei lavori derivati da questa a patto che vengano mantenute le indicazioni di chi è l'autore dell'opera;
- **Non commerciale (NC):** permette che altri copino, distribuiscano, mostrino ed eseguano copie dell'opera e dei lavori derivati da questa solo per scopi non commerciali;
- **Non opere derivate (ND):** permette che altri copino, distribuiscano, mostrino ed eseguano soltanto copie identiche dell'opera. Non sono ammesse opere derivate;
- **Condividi allo stesso modo (SA):** permette che altri distribuiscano lavori derivati dall'opera solo con una licenza identica o compatibile con quella concessa con l'opera originale.

A partire dalla versione 2.0 la clausola “Attribuzione” è stata resa obbligatoria in tutte le licenze Creative Commons; la clausola “Condividi allo stesso modo” ricalca il meccanismo virale del *copyleft* introdotto dalla GPL. Se una delle condizioni non viene rispettata la validità della licenza decade, e con essa anche le autorizzazioni concesse (copia, distribuzione ecc.); come nel caso del *copyleft*, le licenze Creative Commons hanno esclusivamente funzione autorizzativa, e la tutela dell'opera è affidata all'orientamento esclusivo delle norme vigenti sul diritto d'autore.

34 Open Source Initiative (OSI) - <https://opensource.org/>

35 È in questo contesto che Raymond ha coniato il termine *open source*.

36 Open Source Definition (OSD) - <https://opensource.org/osd>

37 Creative Commons (CC) - <https://creativecommons.org/>

La cattedrale e il *bazaar*

I progetti legati al *free software* e all'*open source* sono stati i precursori di una serie di prospettive innovative che si sono in seguito diffuse in tutta la cultura dell'*open content*. Il principio etico della “libertà” di accesso ai contenuti, l’idea di “apertura” come incremento di efficienza (si consideri il caso della customizzazione di software libero), un nuovo modello di *business* basato sulla condivisione delle tecniche e dei saperi, le politiche innovative di proprietà intellettuale delle licenze d’uso: molti di questi aspetti stanno “contagiando”, con esiti alterni, le diverse declinazioni del mondo *open*.

La fluidità con cui si manifestano tali acquisizioni metodologiche è senza dubbio legata alla digitalizzazione dei contenuti: uno dei criteri generali che definiscono se un contenuto è aperto è la sua codifica in un formato *machine-readable*. La conseguenza spontanea è che tutto il mondo *open* manifesta un legame stretto con la scienza informatica, che assurge sempre più frequentemente al ruolo di paradigma sia tecnico che culturale.

A questo punto è interessante osservare un ulteriore comportamento innovativo introdotto dall'*open source*, la cui prima applicazione di successo, la progettazione del *kernel* Linux³⁸ ad opera di Linus Torvalds nel 1991, è raccontata in uno scritto di Eric Raymond:

Linux stravolse gran parte di quel che credevo di sapere. Per anni avevo predicato il vangelo Unix degli strumenti agili, dei prototipi immediati e della programmazione evolutiva. Ma ero anche convinto che esistesse un punto critico di complessità al di sopra del quale si rendesse necessario un approccio centralizzato e a priori. Credevo che il software più importante (sistemi operativi e strumenti davvero ingombranti come Emacs) andasse realizzato come le cattedrali, attentamente lavorato a mano da singoli geni o piccole bande di maghi che lavoravano in splendido isolamento, senza che alcuna versione beta vedesse la luce prima del momento giusto. Rimasi non poco sorpreso dallo stile di sviluppo proprio di Linus Torvalds – diffondere le release presto e spesso, delegare ad altri tutto il possibile, essere aperti fino alla promiscuità. Nessuna cattedrale da costruire in silenzio e reverenza. Piuttosto, la comunità Linux assomigliava a un grande e confusionario bazaar, pullulante di progetti e approcci tra loro diversi [...]. Un bazaar dal quale soltanto una serie di miracoli avrebbe potuto far emergere un sistema stabile e coerente. (Raymond, 1999, cap. *La cattedrale e il bazaar*)

38 Il *kernel* è il nucleo di un sistema operativo, che fornisce ai processi in esecuzione l’accesso alle componenti hardware. Esso si occupa di assegnare porzioni di tempo-macchina (*scheduling*) e di gestire l’accesso simultaneo all’hardware da parte di molteplici processi (*multitasking*). Il progetto GNU non ha ancora sviluppato un proprio *kernel* sufficientemente stabile, per cui deve appoggiarsi ad un *kernel* esterno; prima dell’arrivo di Linux, GNU si serviva di un *kernel* proprietario di tipo UNIX. In seguito all’adozione di Linux, che è distribuito liberamente, GNU ha finalmente ottenuto l’effettiva indipendenza da qualsiasi software proprietario. Il sistema che deriva dall’uso congiunto di Linux e GNU prende il nome di GNU/Linux. Si veda a questo proposito l’opinione di Stallman <https://www.gnu.org/gnu/linux-and-gnu.it.html>.

Nella sua analisi Raymond contrappone un metodo classico di programmazione, simile alla progettazione architettonica fatta a tavolino, al metodo innovativo applicato da Torvalds, che ricorre alla cooperazione di migliaia di co-sviluppatori sparsi per il mondo e collegati tramite la rete.

L'approccio a *cattedrale* tende ad accentrare il lavoro, a non ammettere intrusioni, e a distribuire le versioni del software solo dopo essersi accertati di avere risolto la maggior parte dei problemi affrontati. Il modello a *bazaar* sovverte questo comportamento: le *release* sono distribuite velocemente e ci si affida al lavoro indipendente svolto dagli stessi utenti, che sono trattati come se fossero dei co-sviluppatori.

«Distribuisi presto. Distribuisi spesso. E presta ascolto agli utenti.» è il motto coniato da Raymond (1999, cap. *Distribuisi presto e spesso*) per descrivere il metodo di lavoro di Torvalds, il cui merito è stato quello di averlo applicato con un'intensità tale³⁹ da riuscire a gestire un livello di complessità strutturale altissimo come è quello della progettazione di un *kernel* stabile. L'efficienza del sistema è condensata da Raymond nella cosiddetta “Legge di Linus”: «Stabilita una base di beta-tester e co-sviluppatori sufficientemente ampia, ogni problema verrà rapidamente definito e qualcuno troverà la soluzione adeguata» (*ibidem*).

La differenza fondamentale tra i due stili di sviluppo sta nel modo in cui affrontano l'attività di *debugging*⁴⁰. Se per lo stile a cattedrale i *bug* rappresentano una grossa insidia, che richiede tempi lunghi di risoluzione e comporta lunghe attese tra la realizzazione di una versione e l'altra, per lo stile a *bazaar* essi sono fenomeni marginali che, esposti all'attenzione di migliaia di co-sviluppatori, sono risolti in tempi brevi e con soluzioni originali⁴¹.

Lo stile a *bazaar* non può ovviamente essere applicato a partire da zero: non si può dar vita a un progetto in questo modo, è necessario partire almeno da una base di *design* di codice che sia «forte e attraente» (*ivi*, cap. *Le precondizioni necessarie per lo stile bazaar*). Ma il successo ottenuto dal progetto Linux ha il merito supplementare di rivalutare l'importanza del lavoro di *debugging* in confronto al concepimento del *design*.

Il problema con il fatto di essere bravi e originali nel design del software è che tende a divenire un'abitudine – prendi a fare cose carine e complicate quando invece dovresti tenerle semplici e robuste. Proprio per questa ragione mi sono crollati addosso vari progetti. [...] Consideriamo Linux. Supponiamo che Linus Torvalds avesse

39 Raymond riferisce che nei momenti di maggior fermento Torvalds pubblicava le *release* più di una volta al giorno.

40 Il *debugging* è l'individuazione e la risoluzione degli errori (*bug*) presenti nel software, in seguito a test sulle versioni *beta* (o comunque non definitive) distribuite a questo scopo.

41 «Linus fece notare come la persona che si rende conto e risolve il problema non necessariamente né di norma è la stessa persona che per prima lo mette a fuoco. “Qualcuno scopre il problema,” dice Linus, “e qualcun altro lo comprende. E secondo me il compito più difficile è proprio trovarlo”. Ma il punto è che entrambe le cose tendono ad accadere piuttosto rapidamente.» (Raymond, 1999, cap. *Distribuisi presto e spesso*).

cercato di tirar fuori le innovazioni fondamentali dal design del sistema operativo nel corso dello sviluppo; quali probabilità esistono che il kernel risultante fosse ugualmente stabile ed efficace come quello che abbiamo ora?
(*ibidem*)

In altri punti del suo scritto Raymond insiste sull'importanza del riuso dei codici e della capacità di sapere abbandonare ciò che non serve in favore di soluzioni, generalmente più semplici, proposte da altri. Fino a concludere che «la maggior parte della scienza, dell'ingegneria e dello sviluppo del software non viene realizzata da alcun genio originale, il contrario della mitologia dell'hacker» (*ivi*, cap. *Fetchmail diventa adulto*).

Secondo un'opinione diffusa, le moderne comunità di *hacker* condividono una serie di "valori": l'idea di accesso libero alle risorse, la decentralizzazione e condivisione di mezzi e conoscenza, la cooperazione, il significato "positivo" della tecnologia informatica per il bene della comunità, la semplificazione e la creatività. Ma una prerogativa che sembrerebbe scontrarsi con quelle elencate è il senso di onore e credibilità, quell'«intangibile, egoistica reputazione e soddisfazione che si guadagna tra gli altri hackers» (*ivi*, cap. *Il contesto sociale del software open source*) che Raymond, appassionato di fantascienza, assimila al cosiddetto *egoboo*⁴².

A prima vista sembra impossibile che una comunità auto-referenziale che condivide una cultura dell'*ego* possa dar vita a progetti collaborativi su larga scala senza esprimere una tendenza alla frammentarietà, ostilità e segretezza. Secondo Raymond in queste occasioni la comunità di *hacker* raggiunge un equilibrio grazie all'adozione di strategie di adattamento assimilabili a quelle dei sistemi ecologici o del *free market*; in essi l'azione congiunta di un gruppo di attori indipendenti, finalizzata alla massimizzazione di una funzione utilitaristica, produce un ordine spontaneo «più elaborato ed efficiente di quanto avrebbe potuto raggiungere qualsiasi pianificazione centralizzata» (*ibidem*).

I singoli sviluppatori di Linux cercano di incrementare la propria reputazione all'interno della comunità di *hacker*, tramite dimostrazioni di abilità nel trovare la soluzione più semplice, diretta e creativa al problema (la «funzione utilitaristica»); nel caso di un progetto di ampio respiro come Linux, tale soluzione può darsi soltanto nel contesto di un lavoro collettivo: l'equilibrio collettivo è garantito dalla soddisfazione (egoistica) del singolo.

Ci si domanda se il modello di sviluppo a *bazaar* possa essere esportato in altri ambiti della cultura *open*, magari ricorrendo alle comunità che popolano i *social network*. Lo stesso Raymond, che ha utilizzato il metodo di Linux per sviluppare un *client* di posta elettronica, si è detto entusiasta del

42 Il termine *egoboo* proviene dal linguaggio gergale del *science-fiction fandom* ed è impiegato per descrivere il piacere ricevuto dal riconoscimento pubblico del proprio lavoro volontario.

feedback ricevuto: eravamo nel 1996 e i *beta-tester* di cui egli disponeva erano 250. Oggi il *social web* dà spesso origine a comportamenti devianti negli utenti, e sembra difficile pensare che sia semplice sviluppare un lavoro di rifinitura di oggetti complessi in un contesto di totale apertura e accessibilità⁴³. Uno dei motivi di tale diffidenza sta nella constatazione di una diversità di base: nel mondo dell'*open source* gli utenti sono trattati come co-sviluppatori, ai quali è conferito un potere decisionale (possibile in virtù di una politica di condivisione dei saperi); nel mondo del *social web* contemporaneo gli utenti sono consumatori, acquirenti nei quali si inoculano bisogni digitali sempre crescenti. E la condivisione delle tecniche e dei saperi è spesso un'illusione.

I comportamenti virtuosi volti a sostenere generosamente la causa del *free software* e dell'*open source* non sono diffusi solo al livello dello sviluppo del codice: esiste una efficiente comunità di supporto agli *end-user* che produce documentazione aggiornata ed assistenza online gratuita.

Lo spirito della comunità di supporto è inclusivo: si cerca di fornire agli utenti meno esperti la possibilità di acquisire nuove conoscenze ed auto-gestirsi nella soddisfazione delle proprie necessità. Se gli utenti acquistano indipendenza, maggiori risorse possono essere dedicate alla soluzione di problemi più importanti.

43 Si tornerà sull'argomento discutendo della *social peer-review* nel capitolo "Accesso libero alla letteratura scientifica".

Accesso libero alla letteratura scientifica

Il movimento *open access* persegue l'accesso libero alla letteratura scientifica: i risultati della ricerca, soprattutto quella finanziata con fondi pubblici, sono considerati un bene comune di cui tutti devono poter beneficiare. I documenti di disseminazione dei risultati devono pertanto essere resi pubblici ed accessibili gratuitamente, svincolati da restrizioni relative alla proprietà intellettuale; la rapidità dei tempi di pubblicazione e la gratuità sono condizioni ulteriori che favoriscono lo sviluppo della ricerca.

La “Dichiarazione di Berlino” del 2004 (intitolata “Accesso aperto alla letteratura scientifica”) individua due requisiti fondamentali che un contributo ad accesso aperto deve soddisfare⁴⁴. Il primo è legato alle licenze d'uso; gli autori devono garantire al contributo

diritto d'accesso gratuito, irrevocabile ed universale e l'autorizzazione a riprodurlo, utilizzarlo, distribuirlo, trasmetterlo e mostrarlo pubblicamente e a produrre e distribuire lavori da esso derivati in ogni formato digitale per ogni scopo responsabile, soggetto all'attribuzione autentica della paternità intellettuale. (De Robbio, 2007a, p. 324)

Il secondo requisito riguarda l'accessibilità dell'opera, che, codificata in un formato elettronico standard, deve essere:

depositata (e dunque pubblicata) in almeno un archivio in linea che impieghi standard tecnici adeguati e che sia supportato e mantenuto da un'istituzione accademica, una società scientifica, un'agenzia governativa o ogni altra organizzazione riconosciuta che persegua gli obiettivi dell'accesso aperto, della distribuzione illimitata, dell'interoperabilità e dell'archiviazione a lungo termine. (*ivi*, p. 325)

Nel riconoscere che «il passaggio all'accesso aperto modifica la disseminazione della conoscenza nei suoi aspetti legali e finanziari» i firmatari si impegnano a sviluppare «i mezzi e i modi per valutare i contributi ad accesso aperto e le pubblicazioni in linea, così da preservare gli standard qualitativi della validazione e della buona pratica scientifica» e a difendere «il riconoscimento delle pubblicazioni ad accesso aperto ai fini delle valutazioni per le promozioni e l'avanzamento delle carriere» (*ibidem*).

La valutazione della letteratura scientifica, soprattutto quella delle discipline STM (scienza, tecnica e medicina), è strettamente legata alle politiche dei gruppi editoriali che si occupano della sua

44 La dichiarazione, redatta nel 2003 nel corso di una conferenza sull'accesso aperto tenutasi a Berlino presso la società Max Planck, è stata firmata da 569 istituzioni scientifiche (dati aggiornati al 13/9/2016). <https://openaccess.mpg.de/3883/Signatories>

Una traduzione del testo in italiano è pubblicata in De Robbio (2007a, pp. 323-326).

pubblicazione; un gruppo ristretto di editori di prestigio ha elaborato un sistema di valutazione bibliometrica (qualitativa e quantitativa) che rappresenta uno tra i criteri più importanti nell'ambito della più generale valutazione scientometrica della ricerca accademica. I risultati delle valutazioni scientometriche sono molto importanti in quanto contribuiscono a indirizzare scelte istituzionali di carattere politico ed economico (per esempio sono impiegate da finanziatori pubblici o privati come criterio per l'assegnazione di nuovi contributi ai gruppi di ricerca).

I prodotti del movimento *open access* sono tendenzialmente relegati ai margini di questo sistema di valutazione; ciò contribuisce ad incrementare le critiche e l'accusa di auto-referenzialità storicamente mosse verso il sistema stesso. Senza entrare nel merito della questione, che è complessa soprattutto perché si manifesta in modi molto differenti a seconda delle discipline, ciò che a noi interessa è osservare l'emergere nel movimento *open access* di nuovi comportamenti (relativi alle strategie di valutazione dell'affidabilità delle informazioni, di rifinitura di "oggetti" grezzi e di sviluppo di metodologie collaborative) che possono avere una valenza trasversale nel mondo dell'*open content*.

Le prime esperienze di politiche legate all'accesso aperto sono riconducibili alla nascita dell'OAI (Open Archive Initiative) nel 1999⁴⁵. A partire dall'inizio degli anni '90, nell'ambito accademico statunitense era già diffuso l'uso repositories per l'autoarchiviazione di *e-print* (pubblicazioni digitali, *pre-print* e *post-print*), con lo scopo di comunicare i risultati di ricerche accademiche in corso prima della *peer review* e della pubblicazione in riviste⁴⁶; il deposito degli *e-print* avveniva ad opera dei ricercatori stessi in un numero sempre crescente di repositories, afferenti a diverse discipline⁴⁷. Il merito di OAI è aver stabilito un'infrastruttura per l'interoperabilità tra repositories e utenti, ottenuta tramite il disegno di un'architettura funzionale e di un protocollo di *harvesting* di metadati, entrambi utilizzati tuttora nell'ambito del trasferimento e della pubblicazione di ingenti quantità di risorse⁴⁸.

45 <https://www.openarchives.org/>.

46 La storia dell'Open Archive Initiative e un'introduzione tecnica al protocollo OAI-PMH sono consultabili a questo indirizzo: <https://www.oaforum.org/tutorial/>.

47 Il primo esempio è stato quello di <https://arxiv.org/>, piattaforma creata nel 1991 per immagazzinare lavori di ricerca della fisica delle alte energie, tuttora attiva.

48 Con il termine *harvesting* si indica una serie di procedure automatizzate volte alla "raccolta" di metadati, anche in grande quantità.

L'architettura del protocollo OAI si basa sulla distinzione fondamentale tra *data provider* e *service provider*. Il *data provider* è il soggetto che espone i metadati; gestisce un repository (nel quale sono contenuti i metadati ed eventualmente le risorse stesse) in grado di elaborare le richieste inviate dall'*harvester*. Il *service provider* raccoglie i dati e fornisce servizi a valore aggiunto agli utenti che si connettono ad esso.

Il protocollo di *harvesting*, denominato OAI-PMH (Open Archive Initiative Protocol for Metadata Harvesting), è giunto alla versione stabile 2.0 nel 2002; ha una struttura molto semplice, basata sulla trasmissione, tramite protocollo HTTP, di metadati codificati in XML. OAI-PMH consente di gestire quantità di metadati ingenti e di effettuare *harvesting* selettivi in base alla data dell'ultima modifica, al set di appartenenza o al formato dei

Nel corso degli anni '90 sia la comunità scientifica statunitense che quella europea hanno manifestato un crescente malessere legato al rapporto instauratosi tra ricerca ed editoria scientifica.

Il disagio si fonda nella constatazione di un paradosso: l'edizione tecnico-scientifica si nutre del lavoro dell'accademico, delle sue ricerche, dei suoi testi, delle sue revisioni e, in cambio, lo stesso ricercatore che alimenta il sistema, quando vuole consultare queste stesse fonti per poter continuare il suo lavoro, è obbligato a pagare un plusvalore che guadagna l'editore. Se a questo aggiungiamo che una gran parte della ricerca scientifica è sostenuta da fondi pubblici e che i vantaggi vanno alle aziende private, allora il paradosso aumenta.

Guédon descrive questa situazione come un mercato "anelastico", vale a dire, con una domanda totalmente prigioniera, visto che non dispone di offerta alternativa di contenuti. Questo facilita la posizione dei venditori che possono fissare i prezzi liberamente. (Abadal, 2014, p. 28)

Tali motivazioni di carattere economico, documentate da una serie di studi analitici⁴⁹, hanno condotto alla stesura delle Dichiarazioni di Budapest (2002), Bethesda (2003) e Berlino (2004) con le quali si è sancita la nascita del movimento *open access*⁵⁰. La Dichiarazione di Budapest ha individuato due strade alternative per l'implementazione dell'accesso aperto alla letteratura scientifica:

- pubblicare in riviste ad accesso aperto (*golden road*);
- archiviare in depositi ad accesso aperto (*green road*).

La *golden road* cerca di stabilire una continuità con il sistema editoriale tradizionale e pone importanti questioni relative alla propria sostenibilità economica e alla valutazione (in termini qualitativi e di impatto sulla comunità scientifica) delle pubblicazioni.

La sostenibilità economica è una delle sfide principali dell'*open access* (OA); se le riviste tradizionali erano sovvenzionate dagli abbonamenti sottoscritti dagli utenti (soprattutto biblioteche istituzionali ed accademiche), nel caso delle riviste OA questo costo è a carico dell'autore. L'idea di fondo è che parte dei finanziamenti dei progetti di ricerca vadano destinati alla disseminazione dei risultati, tramite pubblicazione in riviste OA; è certo che la sostenibilità della *golden road* può

metadati; tali caratteristiche consentono lo *scheduling* di aggiornamenti con procedura automatizzata. Il formato minimo richiesto per la codifica XML dei metadati è *Dublin Core Unqualified*.

49 Le cifre emblematiche riportate dallo studio di Kyrillidou (2012), citato in Abadal (2015), affermano che «nel periodo 1986-2011, l'aumento del costo degli abbonamenti a riviste accademiche fu del 402%, mentre l'inflazione arrivò solo al 106% (ossia, un incremento quattro volte superiore)» (p. 29).

50 In Italia la maggior parte degli Atenei ha sottoscritto nel 2004 la Dichiarazione di Messina, nella quale si dichiara l'adesione programmatica alla Dichiarazione di Berlino. Le firme dei rettori di 71 Università sono state depositate presso la Società Max Planck; nel 2006 ha aderito alla Dichiarazione di Berlino anche l'AIB (Associazione Italiana Biblioteche), che svolge un ruolo di primo piano nella diffusione del tema *open access*. Il CRUI (Conferenza dei Rettori delle Università Italiane) ha istituito un "Gruppo di lavoro Open Access" che ha elaborato le linee-guida per fornire indicazioni sulle migliori pratiche dell'accesso aperto. <https://www.cruil.it/open-access.html>

essere assicurata solo dalla ristrutturazione dei flussi economici relativi alle pubblicazioni editoriali⁵¹.

La valutazione della letteratura scientifica tradizionale si avvale di due approcci: uno qualitativo, che concerne i meccanismi *ex-ante* di *peer-review*, e uno quantitativo che riguarda l'analisi citazionale ed avviene *ex-post*. Il fatto che le riviste OA non siano spesso conformi a questi criteri ha sollevato critiche sulla qualità degli articoli pubblicati.

L'analisi citazionale è dominata dall'impiego dell'Impact Factor (IF), un indice dell'impatto avuto da una rivista⁵² sulla comunità scientifica calcolato in base al numero di citazioni ottenute. L'IF è molto criticato in ambito OA perché la sua definizione è auto-referenziale: pubblicato annualmente dalla Thomson Reuters, colosso dell'informazione economico-finanziaria, fa riferimento principalmente a criteri interni all'ambiente delle riviste scientifiche commerciali. Gran parte della fenomenologia delle pubblicazioni OA ne è quindi esclusa, e, più in generale, il rischio di conflitti di interesse è molto alto⁵³.

Sono stati quindi proposti nuovi metodi di valutazione bibliometrica *ex-post*; alcuni di essi si rifanno, come l'IF, alle analisi citazionali, che sono però estese a criteri o ambiti meno esclusivi (EigenFactor, H-index, Web Impactor Factor, CiteBase, CiteSeerX ecc.); altri metodi più sperimentali cercano di stabilire in modo attendibile l'impatto tramite l'analisi dell'uso (Usage Factor) che viene fatto della risorsa (De Robbio, 2007b, pp. 277-278).

Dal punto di vista della valutazione *ex-ante*, la qualità delle pubblicazioni della *golden road* è stata sottoposta a critiche negative in quanto si considera che tra le riviste OA la *peer-review* sia scarsamente praticata o inadeguata. La questione è emersa apertamente in seguito alla pubblicazione di un articolo sulla rivista *Science* firmato da John Bohannon, biologo ad Harvard e giornalista per riviste di divulgazione scientifica. Nel 2013 Bohannon ha condotto un'indagine

con l'obiettivo di smascherare quelle riviste "scientifiche" che pubblicano articoli senza una vera procedura di peer review, interessate soltanto a ricevere il contributo economico pagato dall'autore [...]. L'esperimento [...]

51 Un esperimento importante in questa direzione è rappresentato da SCOAP³ (Sponsoring Consortium for Open Access in Particle Physics Publishing), «consorzio di centri di ricerca, enti di finanziamento e biblioteche che propone un modello economico per trasformare in open access tutte le riviste scientifiche del settore della fisica delle particelle» (Abadal, 2015, p. 43-44). Il modello prevede il coinvolgimento di tutti gli attori della filiera editoriale (autori, editori, biblioteche, enti finanziatori ecc.) che, coordinati da biblioteche ed enti di finanziamento alla ricerca con forte potere contrattuale nei confronti degli editori, pubblichino i risultati della ricerca scientifica ad accesso aperto senza costi per gli autori.

52 L'IF può anche essere calcolato in relazione ad un autore, un articolo, un gruppo di lavoro o una nazione.

53 Per una analisi dei limiti dell'IF si veda De Robbio (2007b, pp. 265-268).

consisteva nell'inviare un "articolo civetta" volutamente insignificante e colmo di errori anche metodologici a un campione di 304 riviste – tutte rigorosamente ad accesso aperto - per trarre conclusioni sulla attendibilità del processo di peer review (quando esistente) e comunque sul controllo di qualità effettuato dalle riviste OA.

Di queste, ben 157 hanno accettato l'articolo a fronte del solo pagamento, senza accorgersi che era fasullo, mentre riviste OA "serie" come PLOS non sono cadute nel tranello e hanno bocciato il falso articolo dopo meticolosi controlli [...]. (De Robbio, 2013b)

L'articolo (Bohannon, 2013) è stato duramente criticato dall'ambiente OA, in quanto fazioso e caratterizzato da gravi carenze metodologiche⁵⁴; ciononostante ha il merito di puntare l'attenzione sul fenomeno dell' "editoria predatoria" che sorge intorno alle riviste OA e al modello non-consolidato di sostenibilità economica che esse propongono. Jeffrey Beall, professore associato a Denver ed ex-bibliotecario accademico ad Harvard, pubblica sul suo blog "Scholarly Open Access" una lista di «potential, possible, or probable predatory scholarly open-access publishers»⁵⁵.

Molti di questi editori, strumentalizzando l'accesso aperto come opportunità, mancano di qualsiasi affidabilità e operano soltanto con lo scopo di lucro, chiedendo denaro agli autori e pubblicando senza un minimo di revisione scientifica. In apparenza editoria scientifica, in realtà non sono altro che forme di editoria di vanità, o editoria a pagamento senza qualità. Si tratta di falso Open Access, che opera attraverso forme di spam via mail, contattando autori e agendo ai limiti dello scam (truffa). (De Robbio, 2013a)

Ancor più grave è il fatto che questi meccanismi si stiano insinuando anche nella sottomissione di articoli a riviste tradizionali che pubblicano anche in modalità OA:

Numerosi editori tradizionali internazionali che detengono pacchetti con migliaia di riviste con ottimi IF (Impact Factor) stanno trasformando le loro riviste da chiuse a ibride. Sgradito effetto collaterale, a fronte del pagamento di una robusta quota pagata dall'istituzione entro il modello economico noto come Apc (Article Processing Charge) i comitati preposti alla revisione prestano sempre meno attenzione alla qualità. Più un editore pubblica, più incassa: e l'etica scientifica ne rimane compromessa. (De Robbio, 2013b)

Sembra quindi che il fenomeno dell' "editoria predatoria" non possa essere circoscritto alle riviste della *golden road*; inoltre l'abbassamento della qualità e le critiche al sistema generale della *peer-review* possono essere estese anche alle riviste di carattere tradizionale⁵⁶. Le "falle" emerse nel

54 Bohannon è accusato di aver circoscritto il campione esclusivamente a riviste OA, circostanza che non consente di escludere che le riviste non-OA non si comportino allo stesso modo del campione analizzato. Inoltre esso include un buon numero di riviste OA segnalate nella lista di Beall (cfr. *infra*), che notoriamente svolgono attività di "editoria predatoria".

55 La lista, pubblicata a partire dal 2010 e aggiornata periodicamente, contiene qualche centinaio di titoli di riviste. <https://scholarlyoa.com/>

56 Si veda l'articolo sul blog del biologo Michael Eisen (<http://www.michaeliseisen.org/blog/?p=1439>), e i commenti relativi al sensazionale *paper* (pubblicato su *Science*) che annuncia la scoperta, rivelatasi poi infondata, di un ceppo di batteri con DNA a base di arsenico anziché di fosforo.

modello di sostenibilità economica della *golden road* indicano un problema di portata più vasta, che coinvolge l'intero settore dell'editoria scientifica.

Il sistema dell'auto-archiviazione della letteratura scientifica (la *green road*) ha elaborato nuove strategie di valutazione della qualità dei materiali pubblicati.

In primo luogo bisogna tener conto che i repositories OA contengono tipi eterogenei di materiale scientifico: *pre-print*, *post-print*, collezioni di dati, materiale didattico (*slide*), atti di convegni, tesi di dottorato, monografie ecc. Tutti questi materiali, a prescindere da revisioni *ex-ante*, sono considerati “conoscenza aperta” che può circolare apertamente e rapidamente al fine di incrementare lo sviluppo della ricerca. Essi sono indicizzati in modo che l'utente sappia che cosa sta consultando e scelga il grado di attendibilità da conferire al documento: l'operazione di valutazione della qualità è quindi demandata all'utente stesso ed avviene in un contesto di “apertura” dei flussi di comunicazione e condivisione degli strumenti di revisione, piuttosto che attraverso procedure che hanno un carattere di “segretezza”.

In secondo luogo, il mondo OA (sia nella versione *gold* che in quella *green*) non ripudia il confronto tra pari per il raffinamento dei prodotti: se da un lato le riviste OA più prestigiose praticano la *peer-review ex-ante*, dall'altro la *green road* ha indotto alla sperimentazione di nuovi modelli di revisione tra pari, in cui l'operazione avviene *ex-post*, nell'ambito di tecnologie di *social web*.

A questo proposito si parla di *open peer-review*⁵⁷ o *social peer-review* per definire quel «controllo di qualità [che] viene svolto esclusivamente *ex-post* attraverso gli strumenti “sociali” [...] del Web 2.0: blog, wiki, siti di *social bookmarking* quali Connotea, CiteUlike o Mendeley⁵⁸ o piattaforme professionali quali Nature Precedings⁵⁹[...]» (Cassella, 2010, pp. 122-123).

Secondo Cassella (2010) gli strumenti di *social peer-review* sono particolarmente consoni per la valutazione della letteratura scientifica di ambito umanistico; essi infatti

sembrano adattarsi bene alle caratteristiche della ricerca nelle scienze umane che ha un suo punto di forza nella dialettica, nella comunicazione e nell'apertura del dibattito verso l'esterno, fuori dall'accademia, presentandosi

57 Secondo Cassella (2010, p. 122) il termine *open* comporta l'ambiguità semantica tra il concetto di “aperto ai commenti di tutti” e quello di “pubblico in quanto non anonimo”; quindi è preferibile la dicitura *social peer-review*.

58 In riferimento agli strumenti di *social bookmarking* Cassella (2010, p. 122) parla di *soft peer-review*. [NdR]

59 *Nature Precedings* è «iniziativa del gruppo Nature, quale luogo aperto per i ricercatori di condivisione dei lavori di ricerca *pre-print*, non pubblicati, manoscritti, presentazioni, posters, white papers, lavori tecnici, scoperte supplementari e altri documenti del circuito della comunicazione scientifica» (De Robbio, 2007b, p. 261). Il materiale pubblicato non è sottoposto a *peer-review*, e il momento del referaggio viene spostato dopo la pubblicazione; esso avviene apertamente e su larga scala, con la partecipazione di tutta la comunità, compresi gli editori. (<http://precedings.nature.com/>)

Tra le sperimentazioni di *open peer-review* si segnala anche quello attuato da PloS ONE, una delle sette riviste del prestigioso gruppo OA PLoS. (<http://journals.plos.org/plosone/>) [NdR]

così come una possibile alternativa al sistema di revisione qualitativa *ex-ante* mai realmente consolidatosi nelle discipline umanistiche. (p. 123)

La produzione della letteratura scientifica potrebbe quindi orientarsi verso l'adozione di metodi di carattere collaborativo alla costruzione del testo letterario (e allo stesso tempo valutare una revisione del ruolo della paternità intellettuale, molto radicato nell'ambito accademico umanistico); una simile svolta potrebbe implicare anche un ripensamento *a monte* della struttura dei testi scientifici, finalizzato all'ottimizzazione di metodi collaborativi di "rifinitura".

La stampa ci ha abituato a produrre testi lunghi e in sé conchiusi, come piccole astronavi concepite per sopravvivere negli spazi vuoti di una conversazione che la stampa con i suoi filtri anteriori alla pubblicazione rendeva lenta e costellata di lunghissimi intervalli di silenzio. Il nuovo sistema, di contro, richiederebbe interventi brevi e disposti ad arricchiarsi non più in virtù della loro autosufficienza, ma della loro connessione. (Pievatolo, 2015)⁶⁰

Anche se non è corretto istituire facili parallelismi, il confronto tra questa proposta radicale di *social peer-review* e l'attività di *debugging* tipica di alcuni progetti di sviluppo *open source* è abbastanza immediato. In entrambi i casi si parte da un *design* di base (la bozza dell'articolo) i cui *bug* sono individuati e corretti da proposte di emendamento (la *peer-review*) elaborate da singoli revisori appartenenti a una comunità collegata in rete. Come spiega Raymond, nel caso dell'*open source* l'efficacia di questo sistema si affida ad un equilibrio della comunità di *hacker* che ha qualcosa di "miracoloso"⁶¹; non è detto che questo modello sia esportabile in tutti gli ambienti di sviluppo.

Una cosa è certa: il modello *open access* propone l'abbandono o la revisione di pratiche consolidate nell'editoria e nella ricerca accademica; i "vuoti di potere" generati da tali slittamenti di responsabilità e di potere decisionale generano nella comunità scientifica una sensazione di "perdita di controllo" (dei percorsi, degli strumenti, delle garanzie ecc.) a cui si può far fronte anche aderendo a politiche coraggiose, sperimentali e dall'esito imprevedibile.

60 A questo proposito si segnala la sperimentazione di *open-peer review* condotta dal "Bollettino telematico di filosofia politica" - <http://commentbfp.sp.unipi.it/>. Lo strumento utilizzato per le revisioni è CommentPress, un tema di WordPress realizzato dall'Institute for the Future of the Book che consente una revisione delle singole sezioni che costituiscono i testi (articoli o monografie). <http://futureofthebook.org/commentpress/>

61 È lo stesso Raymond a parlare di "miracolo" all'inizio del suo scritto.

La centralità del formato XML

Diffusione e pregi

XML (EXtensible Markup Language) è un linguaggio di marcatura di testi fondamentale per qualsiasi attività di catalogazione in ambito digitale; esso offre un insieme di notazioni convenzionali (i *tag*) la cui applicazione ad un testo ne definisce la struttura formale. La sua capacità di gestire agilmente e in maniera interoperabile documenti contenenti informazione strutturata ha reso XML lo strumento elettivo per la gestione di metadati catalografici.

Lo sviluppo di XML ha avuto inizio a metà degli anni '90 come sottoinsieme del linguaggio SGML⁶². XML 1.0 è divenuto una raccomandazione del W3C nel 1998; da subito esso ha dimostrato di essere particolarmente stabile: nel corso degli anni la versione iniziale ha subito lievi revisioni, e l'ultimo aggiornamento risale alla quinta edizione della versione 1.0, pubblicata nel 2008⁶³. XML è inoltre supportato da una serie di tecnologie che consentono di estendere le funzioni del linguaggio stesso⁶⁴.

Essendo sviluppato dal W3C secondo le regole degli *open standard*, XML è indipendente da tecnologie e formati proprietari; le sue specifiche tecniche sono liberamente accessibili⁶⁵ e, di conseguenza, esso è utilizzato da un numero molto elevato di tecnologie. La sua struttura di linguaggio di markup consente ai file XML (semplici file di testo marcati da appositi *tag*) di essere sia *machine-readable* che *human-readable* (a condizione di non essere troppo complessi).

Un ulteriore fattore di successo di XML è la sua pronunciata duttilità. Ad un esame più accurato esso non risulta un linguaggio completo come è invece HTML, le cui specifiche descrivono sintassi e semantica dei marcatori di testo. XML non definisce una semantica predefinita dei *tag*, ma solo un insieme di convenzioni standard per la descrizione di linguaggi di *markup*; esso quindi non è un linguaggio predefinito ed *estensibile*, bensì un meta-linguaggio per la costruzione di nuovi linguaggi adatti a particolari domini, la cui semantica è definita in funzione del tipo di informazioni che si desidera rappresentare.

62 SGML (Standard Generalized Markup Language) è un linguaggio di *markup* molto versatile e potente inventato per l'IBM negli anni '70, la cui diffusione è ridotta a causa della complessità delle sue specifiche tecniche.

63 Una versione semplificata di XML (la 1.1) è stata pubblicata nel 2004, ma il suo utilizzo è poco diffuso.

64 Tra le più importanti: DTD e XML Schema sono linguaggi per la creazione di nuovi linguaggi XML; XSL descrive i fogli di stile che definiscono la formattazione dei file XML; XSLT ne struttura la trasformazione in un altro formato; XPath individua insiemi di elementi all'interno dei file XML, ed è a fondamento di XQuery, tecnologia per effettuare interrogazioni su basi di dati XML ecc.

65 <https://www.w3.org/XML/>

L'insieme di queste caratteristiche ha fatto sì che XML sia non solo ampiamente utilizzato per la definizione della struttura dei documenti, ma anche per il trasferimento delle informazioni nel web⁶⁶ e la serializzazione di un elevato numero di altri modelli e tecnologie⁶⁷. A tal proposito è necessario segnalare che la maggior parte dei modelli impiegati nell'ambito del *semantic web* sono serializzati in XML: in questo modo gli schemi concettuali che essi rappresentano sono tradotti in un formato *machine-readable* e diventano perciò processabili nell'ambito di operazioni automatizzate⁶⁸.

Se si osserva l'insieme di tecnologie rappresentate dal *Semantic Web Stack*⁶⁹ risulta evidente che XML costituisce il sub-strato sintattico di tutta l'architettura del *semantic web*. In questo contesto tecnologico, la creazione in sede di catalogazione di record XML comporta un duplice vantaggio: i metadati possono essere sia *mappati* verso altri formati appartenenti allo stesso livello tecnologico (quindi lungo l'asse orizzontale) sia recuperati per l'implementazione di nuove tecnologie semantiche (lungo l'asse verticale).

66 Il protocollo di trasmissione OAI-PMH utilizza il formato XML per la codifica dei metadati.

67 La serializzazione è un «processo di traduzione di strutture di dati o stato di un oggetto in un formato che può essere memorizzato (in un file o trasmesso in rete) e ricostruito più tardi (deserializzazione), nello stesso computer o in un ambiente diverso» (Guerrini e Possemato, 2015, p. 235).

68 Uno dei due formati raccomandati dal W3C per la serializzazione di RDF (la *lingua franca* del *semantic web*) è RDF/XML. Di conseguenza tutti gli schemi di dati, i vocabolari, le ontologie e i linguaggi per vocabolari ed ontologie del *semantic web* che implementano RDF possono essere serializzati in RDF/XML.

69 Cfr. *infra* p. 55.

Documenti XML

Dal punto di vista concettuale un documento XML è costituito da una struttura gerarchica ad albero. Per i nostri scopi, consideriamo che essa consiste in un insieme di *elementi* organizzati come in fig.1. Gli elementi sono indicati da cerchi ai quali è attribuito un nome, mentre gli archi rappresentano le relazioni *padre-figlio* che intercorrono tra gli elementi.

Ogni albero XML ha in cima un elemento “radice” (in questo caso, l’elemento “A”) che svolge la funzione di contenitore di tutte le sue diramazioni; nella versione testuale della struttura ad albero, l’elemento radice contiene le dichiarazioni di *namespace* e altre importanti istruzioni per l’elaborazione.

La fig.1 presenta una serie di relazioni padre-figlio:

- B è un elemento figlio di A, che a sua volta è padre di B;
- B è sia figlio di A che padre di E e F;
- A è padre di B, C e D.

B, C e D sono *fratelli*, nel senso che, condividendo A come padre, sono posti sullo stesso livello gerarchico; la *sequenza* degli elementi B, C e D è *ordinata*: gli alberi XML considerano infatti l’ordinamento degli elementi come un fattore significativo. Gli elementi che non hanno figli sono chiamati *foglie* (C, D, E e F).

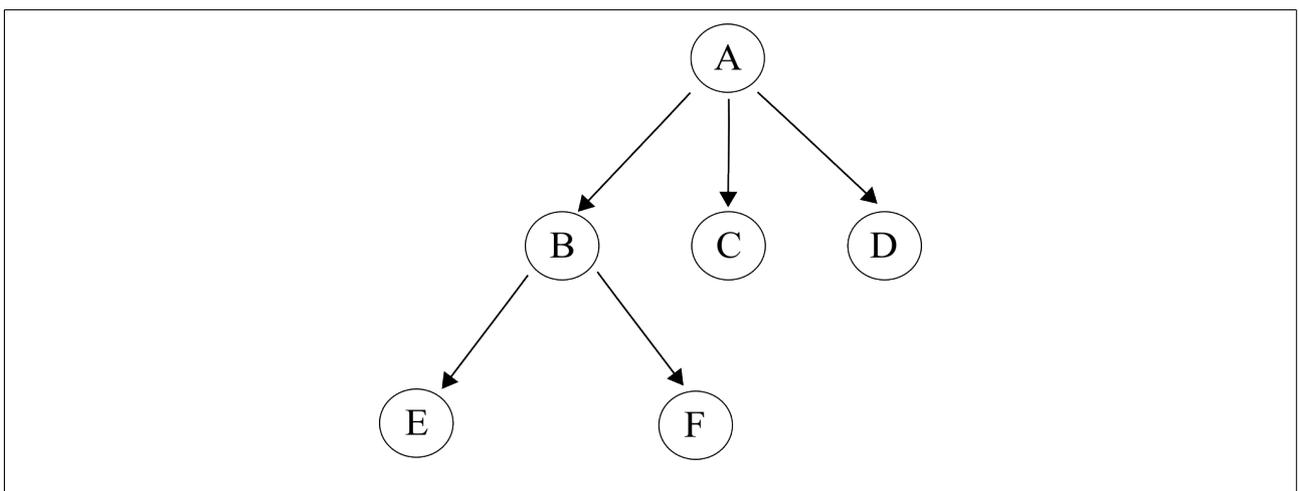


fig.1 Albero XML

Gli elementi si distinguono in base al tipo di contenuto; essi possono:

- contenere altri elementi (come nell'esempio di fig.1);
- essere vuoti;
- contenere stringhe di testo;
- avere un contenuto misto di testo ed elementi.

Nei casi esaminati nel presente lavoro non sono presenti alberi XML con elementi aventi contenuto misto; questo vuol dire che in tutte le strutture XML descritte nei seguenti capitoli, gli elementi che contengono testo sono necessariamente elementi-*foglia*⁷⁰.

Ad ogni elemento possono essere associati uno o più attributi, espressi da una coppia nome/valore. Gli attributi non incrementano la struttura gerarchica di relazioni *padre-figlio* di un albero XML: un attributo non può avere figli, né due attributi posti sullo stesso livello gerarchico possono essere considerati *fratelli*; essi sono solo associati ad un elemento.

⁷⁰ Cfr. *infra* p. 84, dove si parla della differenziazione degli elementi-*foglia* contestualmente alla definizione dello schema XML: siccome l'elemento-*foglia* contiene il dato atomizzato, una sua caratterizzazione conferisce al dato stesso un primo barlume di semantizzazione.

Nell'albero XML di fig.2 gli elementi sono rappresentati da ellissi, i contenuti di testo da rettangoli e gli attributi da parallelepipedi (gli attributi sono riconoscibili anche perché si presentano come coppia nome/valore).

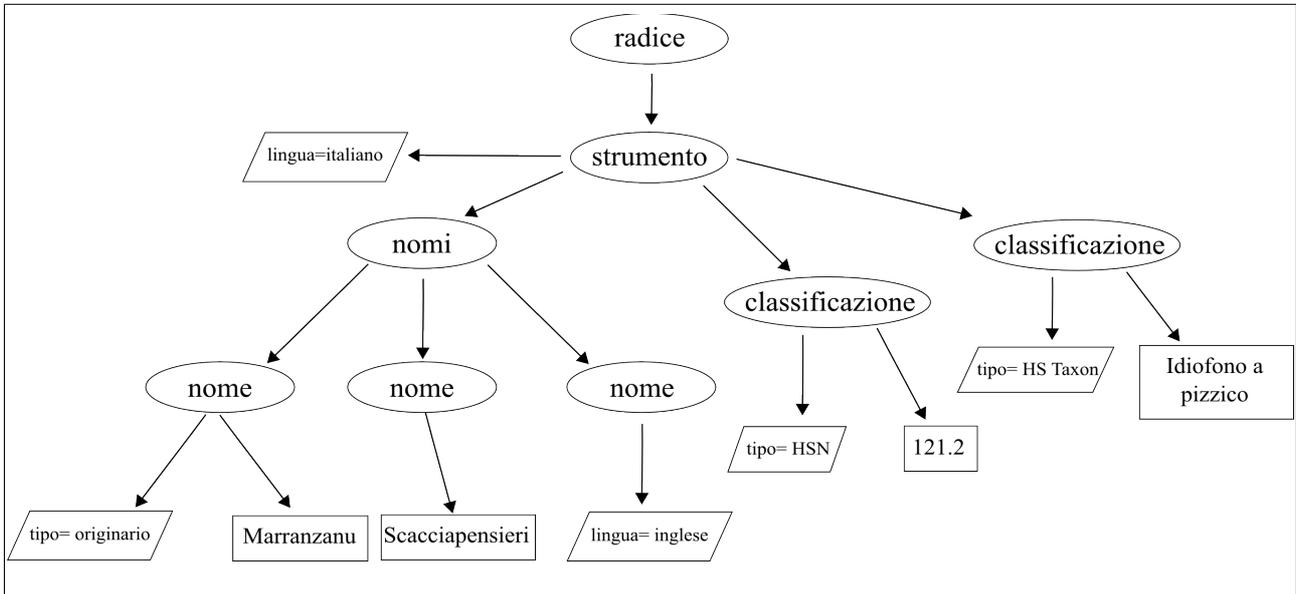


fig.2 Albero XML con elementi, attributi e contenuti testuali.

L'elemento "strumento" ha 3 figli: l'elemento "nomi" più 2 elementi "classificazione"; tutti e 3 gli elementi sono "fratelli" posti sullo stesso livello gerarchico.

Solo gli elementi-*foglia* contengono testo, mentre gli attributi possono essere associati a qualsiasi livello dell'albero: l'elemento "strumento" ha un attributo "lingua=italiano" dal quale non possono partire nuove diramazioni.

Il significato espresso dall'attributo "lingua=italiano" può essere ereditato da tutti i suoi elementi discendenti⁷¹; in questo caso vuol dire che tutti gli elementi dell'albero sono espressi in lingua italiana, tranne il terzo elemento "nome" che specifica la propria lingua tramite un attributo apposito ("lingua=inglese"). Il nome in inglese non viene però specificato: l'elemento "nome" è in questo caso vuoto.

Elementi ed attributi svolgono generalmente una funzione diversa. Gli elementi contengono il dato, l'informazione; gli attributi contengono meta-informazione che specifica il tipo di informazione contenuta dall'elemento (per esempio il "tipo" di classificazione).

71 Ciò dipende da come sono impostati gli algoritmi di lettura dell'applicazione a cui il file XML è diretto.

In un file XML, la struttura ad albero di fig.2 assume una rappresentazione testuale di questo tipo:

```

1  ☐ <radice>
2  ☐ <strumento lingua="italiano">
3  ☐ <nomi>
4  | <nome tipo="originario">Marranzanu</nome>
5  | <nome>Scacciapensieri</nome>
6  | <nome lingua="inglese"/>
7  | </nomi>
8  | <classificazione tipo="HSN">121.2</classificazione>
9  | <classificazione tipo="HS-Taxon">Idiofono a pizzico</classificazione>
10 | </strumento>
11 | </radice>

```

fig.3 Rappresentazione testuale di una struttura ad albero XML

I nomi degli elementi sono racchiusi da specifici delimitatori di *markup*: “<” e “>”⁷².

Ogni elemento ha un contenuto che deve essere racchiuso dai *tag* di apertura e chiusura:

<radice> e </radice>

Gli elementi privi di contenuto possono essere indicati in due modi; con i *tag* di apertura e chiusura privi di contenuto:

<nome lingua="inglese"></nome>

oppure tramite un *tag* vuoto, come nell’esempio riportato in fig.3.

<nome lingua="inglese"/>

Le due scritture sono equivalenti.

Gli attributi sono indicati all’interno dei *tag* di apertura, dopo il nome dell’elemento da cui sono separati tramite uno spazio; essi hanno la forma nome/valore, laddove il nome è separato dal valore da un segno di uguale (=), ed il valore è racchiuso tra due apici (es. lingua="inglese").

A differenza di HTML, il linguaggio XML richiede un certo rigore sintattico e formale. Se un documento XML rispetta determinati criteri sintattici di base, si dice che esso è *ben formato*, e può essere quindi processato da un’applicazione. Uno dei requisiti fondamentali è che i *tag* di apertura e chiusura devono essere *annidati* correttamente: i *tag* di chiusura devono seguire l’ordine inverso dei rispettivi *tag* di apertura. Nel presente lavoro si parla di *annidamenti* proprio per identificare il modo di collocare gli elementi l’uno dentro l’altro, tipico del linguaggio XML.

⁷² I nomi dei *tag* XML sono soggetti ad alcune restrizioni: possono iniziare con una lettera o un underscore e possono contenere lettere, numeri, il punto, l’underscore o il trattino. Non sono ammessi spazi. XML è *case-sensitive*: è sensibile all’uso di maiuscolo e minuscolo.

Schemi XML

Gli esempi visti fin qui hanno riguardato l'analisi della struttura di *documenti* XML, quali possono essere, per esempio, i singoli file XML che contengono i record catalografici. Ciascuno di essi è strutturato secondo precise caratteristiche sintattiche. Ogni file XML:

- presenta determinati nomi di elementi e attributi;
- ha una specifica successione di elementi;
- ammette solo alcuni tipi di contenuto per ciascun elemento;
- struttura elementi ed attributi in base a vincoli di ripetibilità e obbligatorietà.

Uno schema XML è la definizione formale della sintassi di un linguaggio XML, ovvero descrive i *tag* di *markup* ammissibili in un file XML e i vincoli sul loro uso.

Oltre a dover essere intrinsecamente *ben formati*, i documenti XML possono essere sottoposti a un processo di validazione nei confronti di uno schema XML. Il processo di validazione ricorre all'uso

di *processori di schemi*, cioè di strumenti che prendono come input un documento X, chiamato *documento istanza*, e uno schema S [...] e verificano se X è sintatticamente corretto rispetto a S. Se questo è il caso, si dice che X è *valido* rispetto a S. Nel caso il documento non sia valido, la maggior parte dei processori di schemi produce messaggi d'errore di una certa utilità diagnostica. (Møller e Schwartzbach, 2007, p. 95)

Per definire formalmente uno schema XML è necessario fare ricorso a specifici *linguaggi schema*; i linguaggi schema più diffusi sono DTD⁷³ e XML Schema⁷⁴, entrambi sviluppati a partire da direttive del W3C.

Lo studio presentato in questo lavoro è diretto proprio verso l'implementazione nel linguaggio *XML Schema* del modello di scheda catalografica elaborato da Febo Guizzi; l'intenzione quindi è quella di individuare le linee-guida per la definizione di uno specifico schema XML dedicato alla rappresentazione della Scheda Guizzi. I vantaggi della formulazione di uno schema XML sono molteplici:

- una definizione formale dello schema rappresenta una fonte precisa e leggibile da esseri umani; una definizione informale, come potrebbe essere un singolo esempio di scheda catalografica in XML, risulta ambigua e incompleta e rende difficile il compito di dedurre la struttura che

73 DTD (Document Type Definition) - http://www.w3schools.com/xml/xml_dtd_intro.asp

74 <https://www.w3.org/XML/Schema>

soggiace ad essa. La definizione di uno schema quindi assicura che possa avvenire una corretta e comprensibile *disseminazione* del modello di catalogazione;

- uno schema è indispensabile per verificare la validità di ogni nuovo documento XML: il processo di validazione assicura che i record catalografici siano compilati nel modo corretto;
- uno schema è indispensabile per costruire un DBMS (DataBase Management System), un editor/interfaccia che ottimizzi il lavoro di compilazione di nuovi record catalografici.

La scelta del linguaggio schema si è orientata verso *XML Schema* (la cui estensione è .xsd) per una serie di pregi intrinseci al linguaggio e, non secondariamente, in virtù del fatto che anche lo schema XML di LIDO è espresso nel medesimo linguaggio⁷⁵.

Una caratteristica importante di *XML Schema* è che è espresso in XML: non ricorre a un nuovo linguaggio per spiegare sé stesso ed è quindi auto-esplicativo.

Per facilitare il riuso delle descrizioni e migliorare la struttura degli schemi, *XML Schema* ha elaborato il concetto di *tipo*; attraverso le *definizioni di tipo* (che possono essere *semplici* o *complesse*) si stabilisce quale tipo di contenuto può essere associato ad ogni elemento o attributo.

I tipi di contenuto definiti possono essere:

- una famiglia di stringhe di testo Unicode (tipi semplici);
- una collezione di requisiti su attributi, sotto-elementi e dati carattere (tipi complessi).

Le *dichiarazioni di elemento* e le *dichiarazioni di attributo* associano il nome di un elemento o attributo ad un particolare *tipo* di contenuto specificato in una *definizione di tipo*. Affinché il documento XML sia valido è necessario che tutti i requisiti così dichiarati nello schema XML siano soddisfatti.

Senza entrare in una spiegazione dettagliata del sistema dei *tipi*, per i nostri scopi è necessario osservare che un *XML Schema* presenta una serie di definizioni di tipo e di dichiarazioni di elemento ed attributo che indicano i nomi dei *tag* di *markup* ammissibili e li associano a determinati tipi di contenuto.

75 Cfr. *infra* il capitolo “*Unwrapping lido-v1.0.xsd*”.

Le dichiarazioni di elemento sono associate ad *attributi di occorrenza* attraverso i quali si definiscono le *cardinalità*, ovvero il numero minimo e massimo di occorrenze che gli elementi devono avere:

- l'attributo `minOccurs` indica il numero minimo di occorrenze;
- l'attributo `maxOccurs` indica il numero massimo di occorrenze.

Il valore di entrambi gli attributi deve essere un numero intero non negativo.

L'attributo `maxOccurs` può avere il valore speciale "unbounded".

Per entrambi gli attributi il valore di default è "1".

Nell'analisi dell'*XML Schema* di LIDO gli attributi di occorrenza sono esplicitamente associati solo ai valori "0" e "unbounded", dal momento che il valore "1" è implicitamente dichiarato come valore di default. Le configurazioni possibili sono quindi 4:

- nessuna indicazione (default: `minOccurs="1" maxOccurs="1"`): elemento obbligatorio, non ripetibile;
- `minOccurs="0"`: elemento opzionale, non ripetibile;
- `minOccurs="0" maxOccurs="unbounded"`: elemento opzionale, ripetibile;
- `maxOccurs="unbounded"`: elemento obbligatorio, ripetibile.

```

750 <xsd:element name="objectPublishedID" type="lido:identifierComplexType" minOccurs="0" maxOccurs="unbounded">
756 <xsd:element name="category" type="lido:conceptComplexType" minOccurs="0">
762 <xsd:element name="descriptiveMetadata" type="lido:descriptiveMetadataComplexType" maxOccurs="unbounded"/>
763 <xsd:element name="administrativeMetadata" type="lido:administrativeMetadataComplexType" maxOccurs="unbounded"/>

```

fig.4 Esempio di uso degli attributi di occorrenza, tratto dall'XML Schema di LIDO

I LOD (*Linked Open Data*) e il ruolo delle istituzioni culturali nel *semantic web*

Web di documenti e web di dati

Linked Open Data e *semantic web* sono espressioni che stanno entrando nel linguaggio corrente dei professionisti delle “istituzioni della memoria”, ovvero di quel settore che riunisce chi si occupa di conservare, documentare e valorizzare il patrimonio culturale gestito da musei, archivi e biblioteche.

Entrambe le espressioni afferiscono ad uno stesso contesto tecnologico; per *Linked Open Data* si intende un «set of best practices for publishing and connecting structured data on the Web» (Berners-Lee, Bizer e Heath, 2009, p. 1).

It is an expression used to describe a method of exposing, sharing and connecting data via Uniform Resource Identifiers (URIs) on the web. With linked data, in other words, we refer to data published on the web in a format readable, interpretable and, most of all, useable by machine, whose meaning is explicitly defined by a string of words and markers. (Guerrini e Possemato, 2013, p. 67)

I LOD sono un insieme di metodi e tecniche che, tramite l'identificazione dei concetti e la loro interconnessione nell'ambito di rappresentazioni digitali, sono funzionali allo sviluppo del *semantic web*.

Con *semantic web* si intende l'evoluzione del World Wide Web da una rete fatta di documenti (il *web di documenti* che tutti conosciamo) al cosiddetto *web di dati*.

Il modello del web di documenti è una rete di pagine scritte in HTML, il linguaggio di marcatura di testi che ne specifica la visualizzazione e i collegamenti ipertestuali con altri documenti. Questi collegamenti non sono semanticamente strutturati, nel senso che il loro significato non è desumibile dalla macchina; all'utente è affidato il compito di interpretare il significato del documento e aprire i collegamenti ipertestuali di interesse. Il web di documenti si configura quindi come un *file system* globale, «an expression of extreme richness but also particularly monolithic» (*ivi*, p. 71).

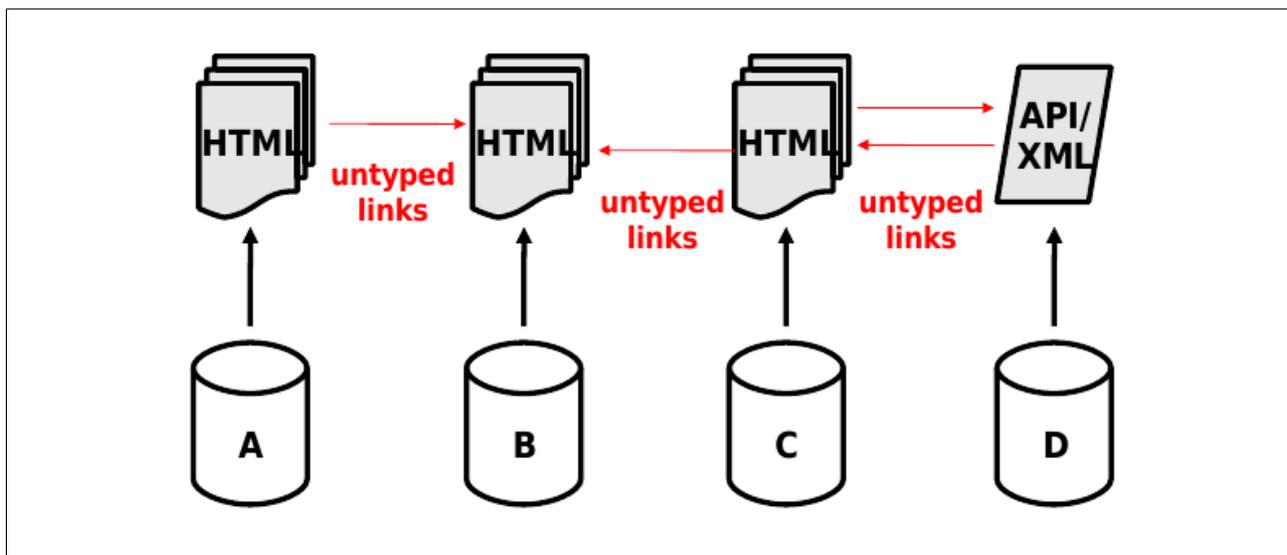


fig.5 Rappresentazione del web di documenti

Bizer, C., Heath, T. e Berners-Lee, T. *Linked data: principles and state of the art*

17th International World Wide Web Conference W3C Track @ WWW2008, Beijing, China 23-24/4/2008.

Il modello del web di dati è invece quello del *database relazionale*; non si tratta più di documenti monolitici ma di *dati* risultanti da un processo di “atomizzazione” dell’informazione:

la pagina HTML evolve a favore di dati singoli, particelle minime autonome, ovvero *atomi*, (da qui il concetto di *atomizzazione*), riaggregabili con modalità e finalità differenti, *on the fly*, secondo il bisogno, le quali producono a loro volta nuove entità. L’attribuzione di significato agli elementi, tramite le ontologie e i vocabolari controllati, spiega il meccanismo di *comprensione*. (Guerrini e Possemato, 2015, pp. 26-27)

I collegamenti tra i dati sono semanticamente connotati: nel *semantic web* chi pubblica i dati rende esplicita la natura della connessione con altri dati. Il significato di tali relazioni è espresso in formati *machine-readable*: se nel *web di documenti* solo l’operatore umano sceglie quale collegamento aprire e come interpretare il documento recuperato, nel *web di dati* si apre un nuovo scenario in cui anche la macchina è in grado di interpretare i dati e fornire risposte che inferiscono nuova conoscenza⁷⁶.

La semantizzazione dei collegamenti conferisce al web di dati la capacità di costituire aggregazioni di informazioni che restituiscono descrizioni di cose del mondo reale: il *semantic web* è definito da Tim Berners-Lee, l’ideatore del World Wide Web, un «web of things in the world described by data on the Web» (Berners-Lee, Bizer e Heath, 2009, p. 2), laddove ciò che interessa non è solo l’aspetto

⁷⁶ «L’aggettivo *semantico* assume la valenza di *comprensibile dalla macchina*: la semantica dei dati consiste nella creazione di informazioni strutturate affinché la macchina possa comprenderle e utilizzarle» (*ivi*, pp. 27-28).

della sua evoluzione tecnica, ma anche l'*estensione* del concetto di “web”. Esso intende ora comprendere la trama di significati che lega le cose del mondo reale («web of things in the world») al fine di restituirne una descrizione fatta di configurazioni di dati «on the Web» (con la “w” maiuscola di World Wide Web). Per costruire un *semantic web* efficace ed affidabile è necessario approfondire la conoscenza del mondo reale e dei modi in cui esso può essere rappresentato (anche, ma non solo, dalle tecnologie digitali).

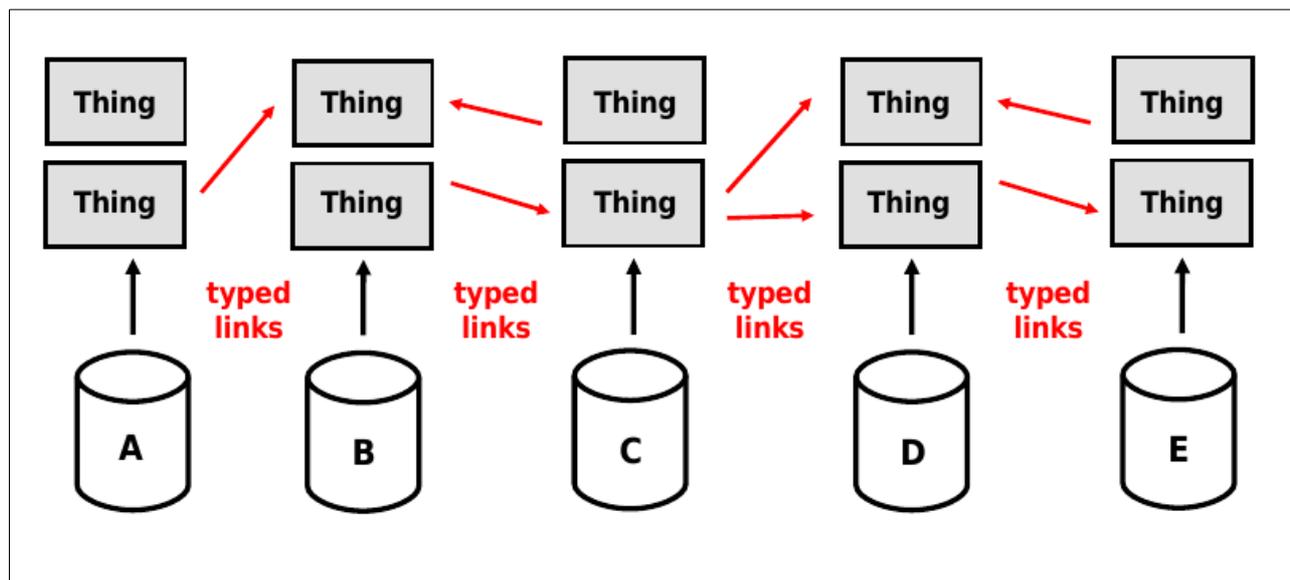


Fig.6 Rappresentazione del web di dati

Bizer, C., Heath, T. e Berners-Lee, T. *Linked data: principles and state of the art*

17th International World Wide Web Conference W3C Track @ WWW2008, Beijing, China 23-24/4/2008.

Sono queste le basi per la realizzazione del “sogno” di Tim Berners-Lee che già nel 1999 prefigurava una rete semantica in cui «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» (Berners-Lee, 2001, pp. 139-140). In questo processo è fondamentale che l’input umano sia scientificamente controllato e organizzato coerentemente; è a questo punto che sorge l’interesse per le istituzioni della memoria: esse hanno alle spalle una lunga tradizione di catalogazione che ha prodotto una mole di dati strutturati e, soprattutto, ha elaborato le teorie sulla strutturazione dei dati e sulla loro valenza “rappresentativa” del mondo reale.

URI e RDF per l'identificazione dei concetti e il collegamento dei dati

Per essere collegati, i dati (o più precisamente i *concetti* rappresentati da dati e gli *oggetti* descritti da aggregazioni di dati) devono essere preventivamente identificati. L'identificazione univoca di oggetti e concetti è un'attività alla base sia del *semantic web* sia della catalogazione tradizionale⁷⁷, e permette di risalire alla loro *identità* indipendentemente dall'espressione testuale e dal contesto in cui sono utilizzati; a questo scopo il Web utilizza gli URI (Uniform Resource Identifiers), stringhe di testo che identificano univocamente qualsiasi tipo di risorsa. Berners-Lee ha individuato 4 regole per la pubblicazione di Linked Data nel Web:

1. Use URIs as names for things.
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL).
4. Include links to other URIs so that they can discover more things⁷⁸.

Il suggerimento è quello di usare la forma HTTP URI, in cui la stringa è formata da un percorso di dominio iniziale e da una parte finale che identifica il singolo oggetto:

- <http://www.mimo-db.eu/InstrumentsKeywords/2861> ;
- http://dbpedia.org/page/Jew's_harp ;
- <http://doi.org/10.1006/jmbi.1998.2354> ⁷⁹.

L'utilizzo di HTTP URI fa sì che la loro gestione sia decentralizzata: ogni amministratore di dominio può creare URI ed è responsabile del loro mantenimento, senza che ci sia la necessità di ricorrere ad un'agenzia centralizzata; ma soprattutto consente la *dereferenziazione* dell'URI, procedimento che è alla base della tecnologia dei LOD e su cui avremo modo di tornare a breve. Un *client* HTTP può dereferenziare l'URI, ovvero può ottenere tramite protocollo HTTP una descrizione della risorsa identificata dall'URI; il formato della risposta ottenuta è negoziabile (*content negotiation*) e può essere, per esempio, in RDF/XML se il contenuto è destinato ad una macchina, o in HTML se è destinato alla lettura da parte dell'uomo.

⁷⁷ In ambito biblioteconomico molto lavoro è stato fatto nell'identificazione e la disambiguazione dei concetti nel corso della redazione di registri di autorità e Soggettari.

⁷⁸ <https://www.w3.org/DesignIssues/LinkedData.html> [NdR].

⁷⁹ È consigliato costruire URI "iconici" che incorporano nella stringa una rappresentazione dell'oggetto/concetto che identificano. Inoltre «è meglio usare chiavi naturali esistenti come parte locale dell'URI, piuttosto che etichette testuali, le quali, per quanto più leggibili dall'occhio umano, offrono, però, minori garanzie di univocità del dato. [...] Nel caso di risorse bibliografiche, per esempio, costruire l'URI utilizzando numeri di controllo quali ISBN, ISSN, DOI, etc.» (Guerrini e Possemato, 2015, p. 43).

RDF (Resource Description Framework) è il modello di descrizione e collegamento di dati raccomandato dal W3C come sostegno all'architettura del *semantic web*⁸⁰; RDF non è un formato di file, ma si propone come una *lingua franca* per l'interoperabilità semantica tra diversi formati e modelli di rappresentazione dei dati nel web. La forma minima di RDF è lo *statement*, un'asserzione unidirezionale di tipo SVO (soggetto-predicato-oggetto) che prende il nome di *tripla RDF*.

Il modello delle triple RDF può essere rappresentato o serializzato in diversi modi a seconda dell'uso e della destinazione. La visualizzazione *human-readable* più accessibile è il grafo, fatto di nodi e archi⁸¹; i formati di serializzazione riconosciuti dal W3C per l'uso di RDF da parte delle macchine sono RDF/XML⁸² (che usa la sintassi XML, difficilmente leggibile dall'uomo) e RDFa⁸³ (Resource Description Framework in Attributes), che incorpora le triple RDF in pagine HTML, coniugando le necessità di leggibilità dell'uomo e della macchina⁸⁴.

I componenti delle triple possono essere di tre tipi:

- URI, stringa che funge da identificativo unico universale, interpretabile dalla macchina;
- *literal*, stringa che esprime un significato, interpretabile dall'uomo;
- *blank node*, stringa che rappresenta un nodo anonimo, cioè non riconosciuto da un identificativo universale⁸⁵.

Vige inoltre il seguente schema di restrizioni:

- il soggetto può essere un URI o un *blank node*;
- il predicato può essere solo un URI;
- l'oggetto può essere un URI, un *blank node* o un *literal*.

Una tripla assume la forma di *tripla literal* se l'oggetto è una stringa che rappresenta un valore (per es. una sequenza di caratteri che esprime un nome di persona, una data o un numero) destinato unicamente all'interpretazione umana. Altrimenti la tripla può essere in forma di *link RDF*, se tutti i componenti sono identificati da URI e sono quindi interpretabili e utilizzabili dalla macchina.

80 <https://www.w3.org/TR/2014/NOTE-rdf11-primer-20140225/>

81 Un grafo è costituito da nodi (che rappresentano il soggetto e l'oggetto di una tripla) e da archi (che rappresentano i predicati). I nodi possono essere ovali se indicano URI, rettangolari se indicano *literal*. I *blank node* sono rappresentati da nodi vuoti, non etichettati.

82 <https://www.w3.org/TR/rdf-syntax-grammar/>

83 <https://www.w3.org/TR/xhtml-rdfa-primer/>

84 Altri formati che coniugano diversamente questi due aspetti sono Turtle, N-Triples, JSON-LD e N3 (Notation3); le specifiche tecniche di tutti i formati di rappresentazione e serializzazione di RDF sono pubblicate tra le raccomandazioni del W3C.

85 L'uso dei *blank node* è sconsigliato perché non consente la pratica dell'*interlinking*.

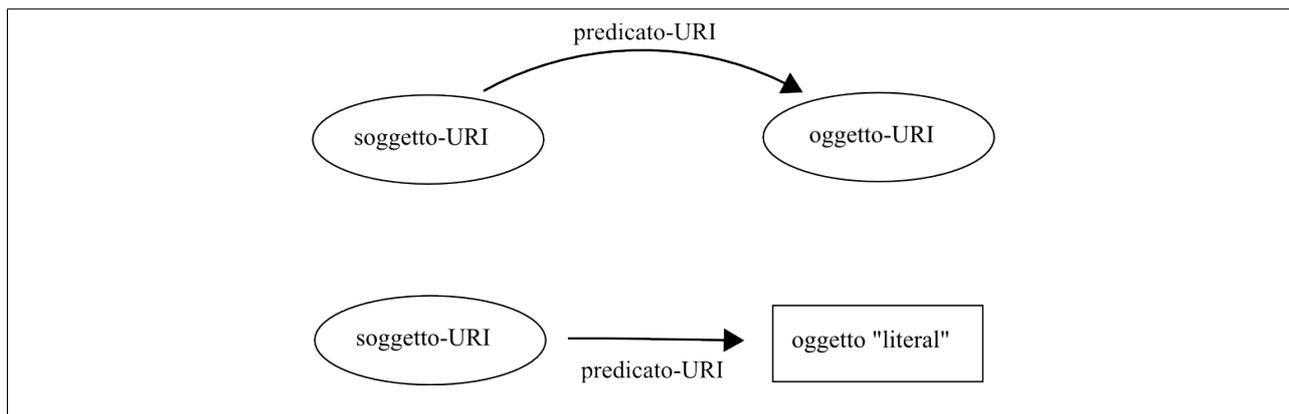


fig.7 Grafi RDF di link RDF e tripla *literal*

Le asserzioni che contengono URI identici possono essere combinate in una successione (*chain*) se l'oggetto della prima tripla è anche soggetto della seconda tripla, o in un grappolo (*cluster*) se lo stesso elemento univocamente identificato è soggetto di entrambe le triple.

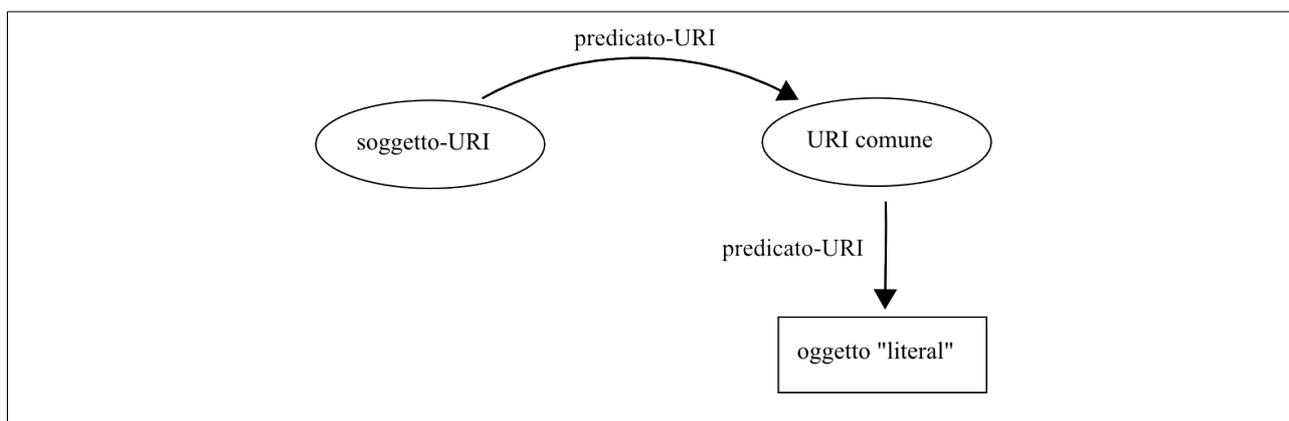


fig.8 Grafo RDF di *chain* di triple RDF

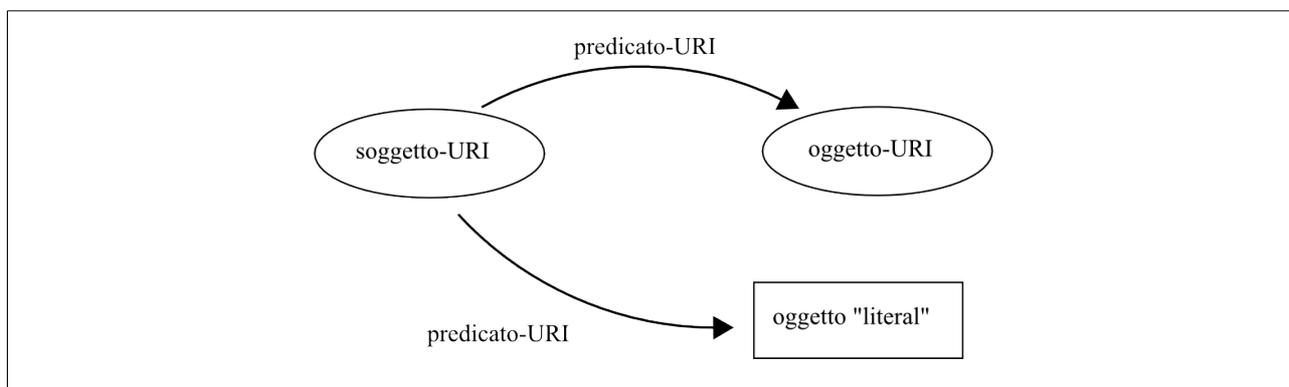


fig.9 Grafo RDF di *cluster* di triple RDF

L'uso di congiunto di tali tecniche di collegamento dà origine a reticoli di asserzioni multiple.

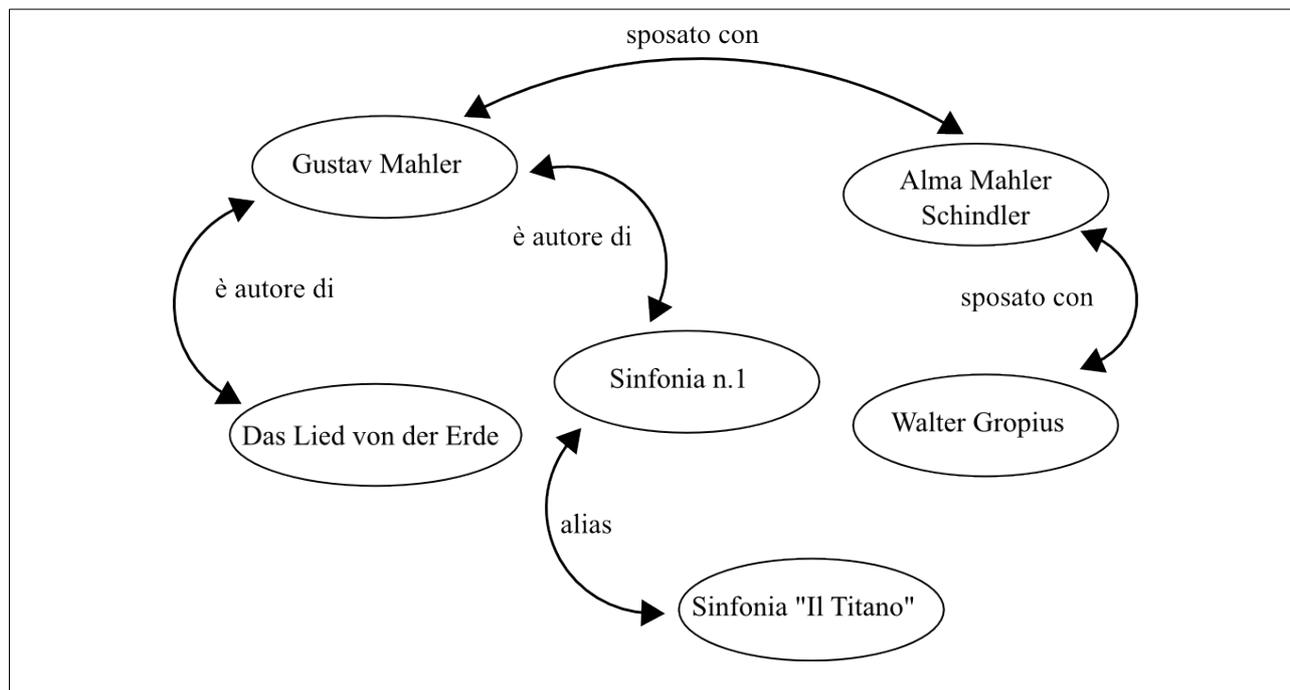


fig.10 Grafo RDF di reticolo di asserzioni multiple

Si parla di *interlinking* quando gli URI del soggetto e dell'oggetto sono referenziati in *namespace* diversi, ovvero appartengono a *dataset* diversi. Questo processo, in cui le informazioni contenute in spazi diversi del Web sono collegate da relazioni qualificate, permette alla macchina di acquisire nuove conoscenze tramite la dereferenziazione degli URI e la conseguente acquisizione di informazioni da altre fonti. Tale pratica risponde a un importante criterio di economicità, la cui valenza non si esaurisce nel "risparmio" derivante dal riuso di informazioni esistenti: infatti «il meccanismo dell'interlinking aumenta la credibilità e l'autorevolezza del dataset e innesca un circolo virtuoso di condivisione e arricchimento dei dati» (Guerrini e Possemato, 2015, p. 67).

Se un *client* HTTP si imbatte in un dato del quale non conosce il significato, può ottenere una descrizione RDF dereferenziano l'URI del dato sconosciuto tramite un'interrogazione HTTP. Nell'esempio riportato in fig.11, il primo grafo asserisce che una risorsa identificata dall'URI:

`http://dig.csail.mit.edu/data#DIG`

ha come membro un'altra risorsa identificata da:

`http://www.w3.org/People/Berners-Lee/`

Quando l'URI del soggetto è dereferenziato tramite protocollo HTTP, il server `dig.csail.mit.edu` risponde inviando una descrizione RDF della risorsa identificata, in questo caso il MIT Decentralized Information Group.

Quando l'URI dell'oggetto è dereferenziato il server del W3C fornisce un grafo che descrive Tim Berners-Lee.

Se si dereferenzia l'URI del predicato `http://xmlns.com/foaf/0.1/member` la macchina ottiene una definizione della proprietà `foaf:member`, descritta in RDF usando RDFS, un linguaggio per la descrizione dei vocabolari. Ogni dato collegato tramite le "buone pratiche" dei LOD risulta quindi auto-descrittivo.



fig.11 Grafo RDF con URI



fig.12 Grafo RDF con URI

I *link RDF* possono essere classificati diversamente, in relazione al tipo di risorsa collegata e alla finalità del collegamento. L'esempio di fig.11 è un *link di relazione*, che determina il rapporto che vige tra oggetti appartenenti a *dataset* diversi. L'esempio di fig.12 è un *link di identità* (di tipo *alias*) che usa la proprietà `owl:sameAs` per affermare che sia il soggetto che l'oggetto si riferiscono alla stessa entità del mondo reale (il film "Pulp fiction"); questo tipo di relazione è fondamentale perché fornisce al *semantic web* lo strumento per la disambiguazione dei concetti. Infine il link ottenuto dalla dereferenziazione del predicato `http://xmlns.com/foaf/0.1/member` è un *link tra vocabolari*, perché collega la proprietà alla definizione che ne dà il vocabolario FOAF⁸⁶.

86 FOAF (Friend Of A Friend) è un'ontologia usata per la descrizione delle persone, delle loro attività e delle loro relazioni con altre persone. Nata per essere utilizzata nell'ambito dei *social network*, è diventata lo strumento più diffuso nel *semantic web* per la descrizione delle persone, al punto da essere impiegata anche in ambito catalografico nonostante presenti alcune evidenti lacune (per esempio non è possibile indicare la data di morte). FOAF è anche un *dataset* di dimensioni importanti che contiene profili di persone pubblicamente accessibili. Ontologie di questo tipo rappresentano il punto di contatto tra il *semantic web* e il *social web*; il lavoro esposto in questa tesi considera il *semantic web* come un potente strumento di consultazione, un enorme archivio digitale, che non implica l'adozione di pratiche del *web 2.0*. Per valutare il significato che il *semantic web* assumerà nei prossimi anni sarà comunque necessario tenere in considerazione l'intreccio che inevitabilmente ci sarà con il mondo dei *social network*. <http://www.foaf-project.org/>

Vocabolari e ontologie

Il *semantic web* si avvale di un complesso di tecnologie che sono usualmente rappresentate come una *pila tecnologica* (*Semantic Web Stack*) in cui ciascun livello della struttura usa le possibilità offerte da quello sottostante. La rappresentazione di fig.13 rende esplicito il fatto che il *semantic web* è, dal punto di vista tecnologico, l'evoluzione del web tradizionale, e che la sua realizzazione può avvenire tramite un "incremento" dei valori strutturali conferiti al web tradizionale e ai dati in esso contenuti. L'immagine evidenzia anche che le modalità LOD di collegamento dei dati fin qui descritte si riferiscono solo alla dimensione prevalentemente *sintattica* della *pila*, necessaria per *identificare* le unità semantiche e i loro percorsi di circolazione.

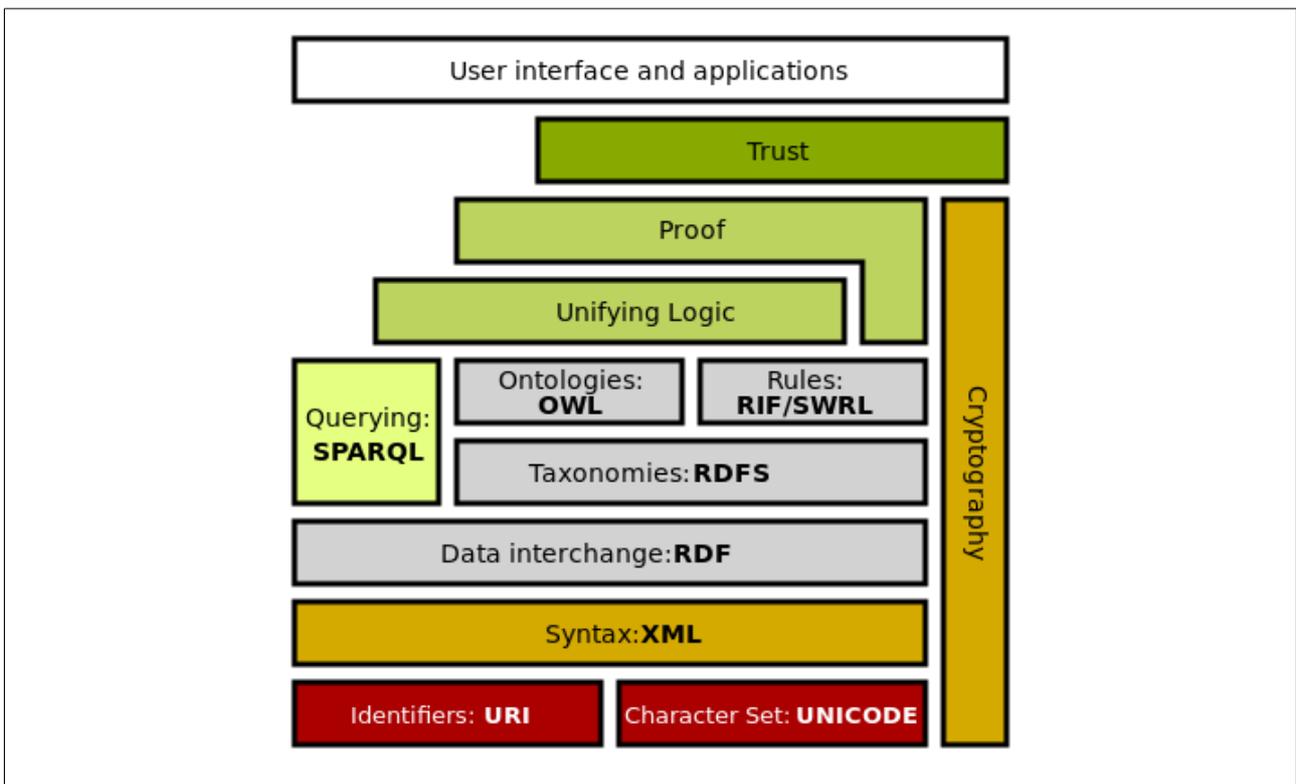


fig.13 Semantic Web Stack – Immagine distribuita da Wikimedia Commons con licenza CC0

Alle istituzioni della memoria è richiesto non solo di condividere i propri dati secondo le "buone pratiche" della modalità LOD, ma anche di fornire la propria competenza per l'elaborazione dei livelli superiori dell'architettura, in particolare quelli relativi alla *semantica* di vocabolari e ontologie dei rispettivi domini di pertinenza.

La terminologia relativa a *vocabolari*, *tassonomie*, *thesauri* ed *ontologie* non è univoca; con il termine generale di *vocabolari* Guerrini e Possemato (2015) si riferiscono a strumenti che forniscono i «termini di specifici domini necessari per descrivere le classi delle risorse e le modalità di relazione esistenti tra di esse» (p. 77). Il W3C non fornisce una definizione univoca dei concetti:

On the Semantic Web, vocabularies define the concepts and relationships (also referred to as “terms”) used to describe and represent an area of concern. Vocabularies are used to classify the terms that can be used in a particular application, characterize possible relationships, and define possible constraints on using those terms. [...] There is no clear division between what is referred to as “vocabularies” and “ontologies”. The trend is to use the word “ontology” for more complex, and possibly quite formal collection of terms, whereas “vocabulary” is used when such strict formalism is not necessarily used or only in a very loose sense.⁸⁷

È possibile comunque riportare alcune considerazioni generalmente valide.

L'unico termine tra quelli elencati che può designare “liste controllate di termini” posti sullo stesso livello gerarchico è *vocabolario*.

I concetti di *tassonomia* e *thesaurus* sono molto vicini nell'ambito del *semantic web*. Una *tassonomia* è una serie di termini organizzati in una struttura gerarchica di natura classificatoria, in cui la relazione padre/figlio esprime significati del tipo tutto/parte, genere/specie o tipo/istanza; esse possono essere anche poligerarchiche (ogni elemento può avere più di un genitore). Il *thesaurus* è un vocabolario di natura prevalentemente onomasiologica che stabilisce tra i termini sia relazioni gerarchiche che relazioni associative (del tipo “related term”, “alternative label”) ed equivalenti (del tipo “same as”, “exact match”).

Il termine *ontologia* si colloca generalmente ad un livello più alto dell'architettura semantica. Essa definisce i termini necessari per configurare una rappresentazione di un particolare dominio della conoscenza; tali termini individuano le entità che popolano il dominio e le relazioni che intercorrono tra tali entità. Una definizione classica, che elabora quella storica data in Gruber (1993), sostiene che «an ontology is a formal specification of a shared conceptualization» (Borst, 1997, p. 28); essa ha quindi a che fare sia con un'operazione di “rappresentazione digitale” (*formal specification*) che con l'individuazione di una concettualizzazione condivisa nel mondo reale. La definizione di *ontologia* acquista maggior contesto se accostata a quella di *semantic web* data da Berners-Lee, della quale coglie la duplice valenza di un *web of things in the world* (shared conceptualization) *described by data on the Web* (formal specification).

87 <https://www.w3.org/standards/semanticweb/ontology>

Esistono ontologie formali per la descrizione di specifici domini (FOAF, Geonames Ontology⁸⁸, Music Ontology⁸⁹) e ontologie *upper-level* per favorire l'interoperabilità semantica tra ontologie diverse (CIDOC-CRM, DOLCE⁹⁰ ecc.).

Per costruire ontologie e vocabolari è necessario utilizzare specifici linguaggi formali; tra i linguaggi per la creazione di ontologie i più diffusi sono OWL⁹¹ e RDFS⁹². Uno dei linguaggi più diffusi per la descrizione di vocabolari, tassonomie e *thesauri* è SKOS⁹³.

88 <http://www.geonames.org/ontology/documentation.html> - GeoNames è un'ontologia per la descrizione di informazioni geospaziali e geopolitiche. GeoNames è anche un *dataset* contenente più di 10 milioni di nomi geografici organizzati in categorie. Il *dataset* è esposto in modalità LOD: ogni nome è identificato da un URI che rimanda ad una descrizione RDF/XML relativa a nome, traduzioni, suddivisioni amministrative, popolazione, coordinate spaziali (usando il Geodetic System 1984 – WGS84), codici postali e link a Dbpedia e ad altri vocabolari LOD.

89 <http://musicontology.com/> - Ontologia usata per descrivere gli aspetti della musica legati al *music business*; non fornisce strumenti per la descrizione della musica stessa.

90 DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) - <http://www.loa.istc.cnr.it/old/DOLCE.html>

91 OWL (Web Ontology Language) - <https://www.w3.org/OWL/>

92 RDFS (Resource Description Framework Schema) - <https://www.w3.org/TR/rdf-schema/>

93 SKOS (Simple Knowledge Organization System) - <https://www.w3.org/2004/02/skos/>

Musei, archivi, biblioteche. La questione dell'affidabilità delle fonti

Le istituzioni della memoria sono interessate al mondo dell'informazione in rete come a uno strumento per garantire la visibilità e l'usabilità dei propri dati. Affinché ciò sia possibile i dati devono essere non solo pubblicati in rete, ma anche realmente accessibili e integrati nelle tecnologie del web; l'adesione alle tecnologie e politiche dei *Linked Open Data* e del *semantic web* sembra essere la strada da percorrere affinché l'informazione e la conoscenza prodotte dalle istituzioni della memoria siano «of the web and not only on the web» (Guerrini e Possemato, 2013, p. 77).

D'altra parte anche il mondo dell'informazione in rete manifesta il proprio interesse nei confronti del patrimonio di musei, archivi e biblioteche: essi hanno alle spalle una lunga tradizione di produzione di informazione strutturata di qualità; essa è stata prodotta sulla base della ricerca e della conseguente miglior comprensione di specifici domini, delle relazioni che intercorrono tra le entità che li popolano, e del modo in cui tali mondi possono essere “rappresentati” in forma discreta.

Il ruolo che le istituzioni della memoria sono chiamate a svolgere nell'ambito del *semantic web* può coinvolgere diversi livelli del *Semantic Web Stack*. Ad un livello più basso della *pila* esse possono identificare i concetti tipici del dominio di propria pertinenza tramite l'atomizzazione del flusso di informazione, e condividere quantità di informazioni di qualità in modalità LOD (quindi subito disponibili per nuovi collegamenti semantici); queste attività sono vicine a quelle tipicamente svolte dalla catalogazione tradizionale. Ad un livello più alto possono fornire le proprie competenze per redigere i vocabolari e le ontologie di dominio.

Bianchini (2014) presenta una lucida riflessione sul mutamento in atto nel mondo delle biblioteche in seguito alla pubblicazione delle RDA⁹⁴, le norme dell'IFLA⁹⁵ per la descrizione delle risorse bibliografiche orientate al *semantic web*:

Nel passaggio di etichetta tra AACR e RDA il termine *catalogazione* è scomparso. [...] È necessario cambiare prospettiva e non parlare più di *catalogazione* [...] quanto di registrazione di dati, da collegare, condividere e distribuire. La registrazione bibliografica tradizionale diventerà la risultante dinamica di un processo di aggregazione di dati, la cui creazione, per una parte sempre più consistente, sarà estranea al lavoro di una

94 Le RDA (Resource Description and Access) nascono per sostituire le AACR2 (Anglo-American Cataloguing Rules) e aggiornare il modello catalogafico ai principi ICP (International Cataloguing Principles), al modello FRBR (Functional Requirements for Bibliographic Records) e alle tecnologie del *semantic web*. Nonostante la provenienza di ambito anglo-americano e biblioteconomico, esse si propongono come standard di descrizione e accesso a qualsiasi tipo di risorsa presente sul web. <http://www.rda-jsc.org/archivedsite/rda.html>
È presente un Gruppo Europeo di interesse per RDA (EU-RIG, <http://www.slainte.org.uk/eurig/>).

95 IFLA (International Federation of Library Associations and Institutions) - <http://www.ifla.org/>

biblioteca o di un'agenzia bibliografica, e potrà essere generata da soggetti diversi, come editori, librai, enti di ricerca e organi amministrativi. All'agenzia bibliografica spetta il compito di individuare, selezionare e collegare dati coerenti e di qualità. [...]

Il ruolo delle agenzie che controllano la qualità dei dati non è affatto secondario. Il concetto di riuso dei dati e, quindi, di *interoperabilità* tra sistemi diversi, che proprio in base all'adozione di standard e vocabolari condivisi, riescono a dialogare tra loro implica necessariamente il concetto di *responsabilità culturale*: le istituzioni culturali che si troveranno a gestire tecnicamente, semanticamente e linguisticamente vocabolari e ontologie, svolgeranno un ruolo determinante nella definizione dei lemmi e delle relazioni tra di essi. Le entità, con relativi attributi e definizioni, saranno infatti utilizzate automaticamente e, dunque, acriticamente dai processi inferenziali compiuti dalle macchine.

Tale approccio non sminuisce l'importanza delle istituzioni culturali; al contrario, il loro ruolo si sposta dal controllo dei dati al controllo della semantica dei dati, cioè dalla fase di produzione dei dati allo sviluppo e al controllo degli strumenti (vocabolari, ontologie, ecc.) che servono per collegarli. (pp. 204-205)

Quest'ultimo aspetto è cruciale e va considerato alla luce di tutte le politiche del *semantic web* relative a distribuzione e collegamento dei dati, costruzione dei significati e verifica dell'affidabilità delle fonti e della conoscenza inferita. Solo grazie alla comprensione dei meccanismi che agiscono a livello globale (e, in parte, alla previsione degli scenari futuri) le istituzioni della memoria possono definire con certezza il proprio ruolo.

Nel 2010 Berners-Lee⁹⁶ ha elaborato un sistema di *rating* dei dati, pubblicato sul sito del W3C per incentivare le persone, ma soprattutto i governi che possiedono dati, alla pubblicazione di *Linked Open Data*. Il sistema si basa su 5 requisiti che si concentrano sulle caratteristiche di *apertura* e *collegamento* dei dati sul web.

★	Available on the web (whatever format) but with an open license, to be Open Data
★★	Available as machine-readable structured data (e.g. excel instead of image scan of a table)
★★★	as (2) plus non-proprietary format (e.g. CSV instead of excel)
★★★★	All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
★★★★★	All the above, plus: Link your data to other people's data to provide context

L'attenzione rivolta nei punti 1. 3. e 4. all'adozione di licenze, formati e standard *open* intende favorire il riuso dei dati e consegnare quindi al Web degli strumenti che sono tanto più potenti

96 <https://www.w3.org/DesignIssues/LinkedData.html>

quanto più sono facili da reperire⁹⁷. Essendo aperti e liberamente riusabili, i dati pubblicati in modalità LOD sono l'oggetto potenziale di qualsiasi tipo di collegamento:

The adoption of the Linked Data best practices has lead to the extension of the Web with a global data space connecting data from diverse domains such as people, companies, books, scientific publications, films, music, television and radio programmes, genes, proteins, drugs and clinical trials, online communities, statistical and scientific data, and reviews. (Berners-Lee, Bizer e Heath, 2009, p. 2)

Un esempio interessante di come questi collegamenti possano avere finalità completamente diverse da quelle per cui i dati sono stati prodotti è rappresentato da un progetto realizzato dal Comune di Firenze.

I 230 data set creati per lo stradario sono stati incrociati con il tesoro del *Nuovo Soggettario*⁹⁸ e con il VIAF⁹⁹, dati bibliografici resi pubblici come Linked Open Data, per consentire di individuare vie e piazze dedicate a personaggi (e più in generale a entità) i cui nomi sono utilizzati nei rispettivi archivi. (Bianchini, 2014, p. 199)

Bisogna considerare che l'impulso per la realizzazione di un progetto di questo tipo non è dato esclusivamente dal cambiamento tecnologico in atto, ma si configura anche come una risposta ai nuovi bisogni manifestati dagli utenti, le cui abitudini di ricerca sono influenzate dalla pratica del *web browsing*.

Oggi il pubblico, il lettore, non ragiona più [...] in termini di dati bibliografici, o archivistici o museali o di qualsiasi altra natura. L'utente si aspetta, legittimamente, di potere condurre una ricerca esplorativa a tutto campo nell'universo documentario, senza pensare che possano esistere limiti imposti da una visione che pone al centro l'istituzione che conserva i supporti e non il servizio che si eroga o il cittadino che lo richiede. Il lettore ha bisogno di informazioni su qualsiasi aspetto riguardi il suo oggetto d'interesse. (*Ibidem*)

97 Inoltre utilizzando formati e standard aperti ci si avvale della collaborazione della comunità di sviluppatori che garantisce il mantenimento a lungo termine delle tecnologie adottate.

98 «Il "Nuovo soggettario" è lo strumento impiegabile nell'indicizzazione per soggetto di risorse di varia natura, realizzato a cura della Biblioteca nazionale centrale di Firenze. È aderente ai principi stabiliti dall'International Federation of Library Associations and Institutions (IFLA) e alle indicazioni degli standard internazionali» <http://thes.bncf.firenze.sbn.it/>. A partire dal 2010 è stato pubblicato in formato RDF/SKOS e i suoi metadati possono essere usati secondo la Licenza Creative Commons Attribuzione 2.5. [NdR]

99 VIAF (Virtual International Authority File) «è un progetto internazionale che ha dato vita ad un'unica base dati di voci di autorità controllate, (circa 20 milioni di autori provenienti da 24 cataloghi collettivi nazionali) a disposizione quale servizio gratuito del web semantico; in essa l'utente può cercare un autore utilizzando la forma del nome utilizzata nella propria lingua, scrittura e alfabeto preferiti e accedere immediatamente ai cataloghi nazionali o internazionali corrispondenti per visualizzare la lista delle loro opere». https://it.wikipedia.org/wiki/Virtual_International_Authority_File. [NdR]

Una rappresentazione emblematica del fenomeno *Linked Open Data*, sia in termini di portata e rapidità di sviluppo che di politiche di implementazione, è offerta dal *Linking Open Data Project*¹⁰⁰; l’iniziativa, nata in seno al W3C, intende individuare i *dataset* esistenti rilasciati con licenza *open*, convertirli in RDF secondo i principi LOD e pubblicarli sul Web. Il progetto mantiene l’aggiornamento del *Linked Open Data Cloud*¹⁰¹, un grafo che rappresenta i *dataset* collegati secondo i requisiti LOD; la crescita esponenziale che ha coinvolto la “nuvola” dal 2007, l’anno della sua creazione, al 2014 è indice della rapidità straordinaria che caratterizza il fenomeno LOD; le cifre aggiornate al 2016 parlano di 9960 *dataset* e di 190 miliardi di triple RDF collegate da 50 milioni di link RDF¹⁰².

100 <https://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>

101 <http://lod-cloud.net/> . La “nuvola”, pubblicata nel formato .svg, è corredata di collegamenti ipertestuali e di statistiche aggiornate al 2014.

102 <http://stats.lod2.eu/>

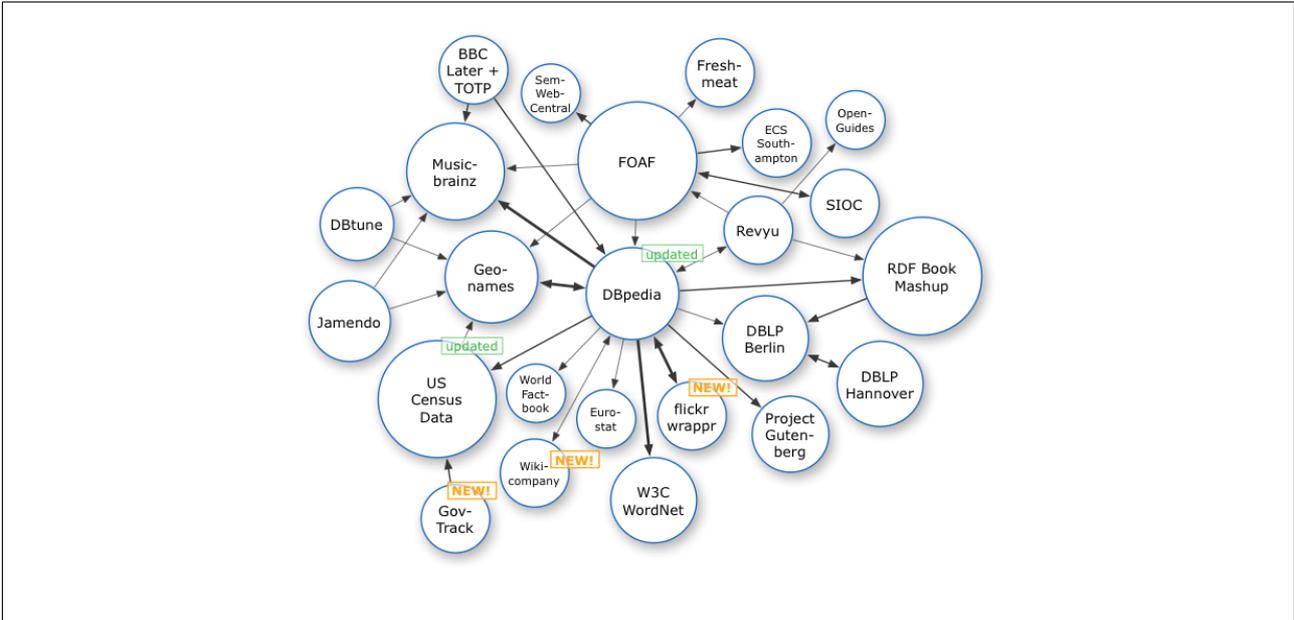
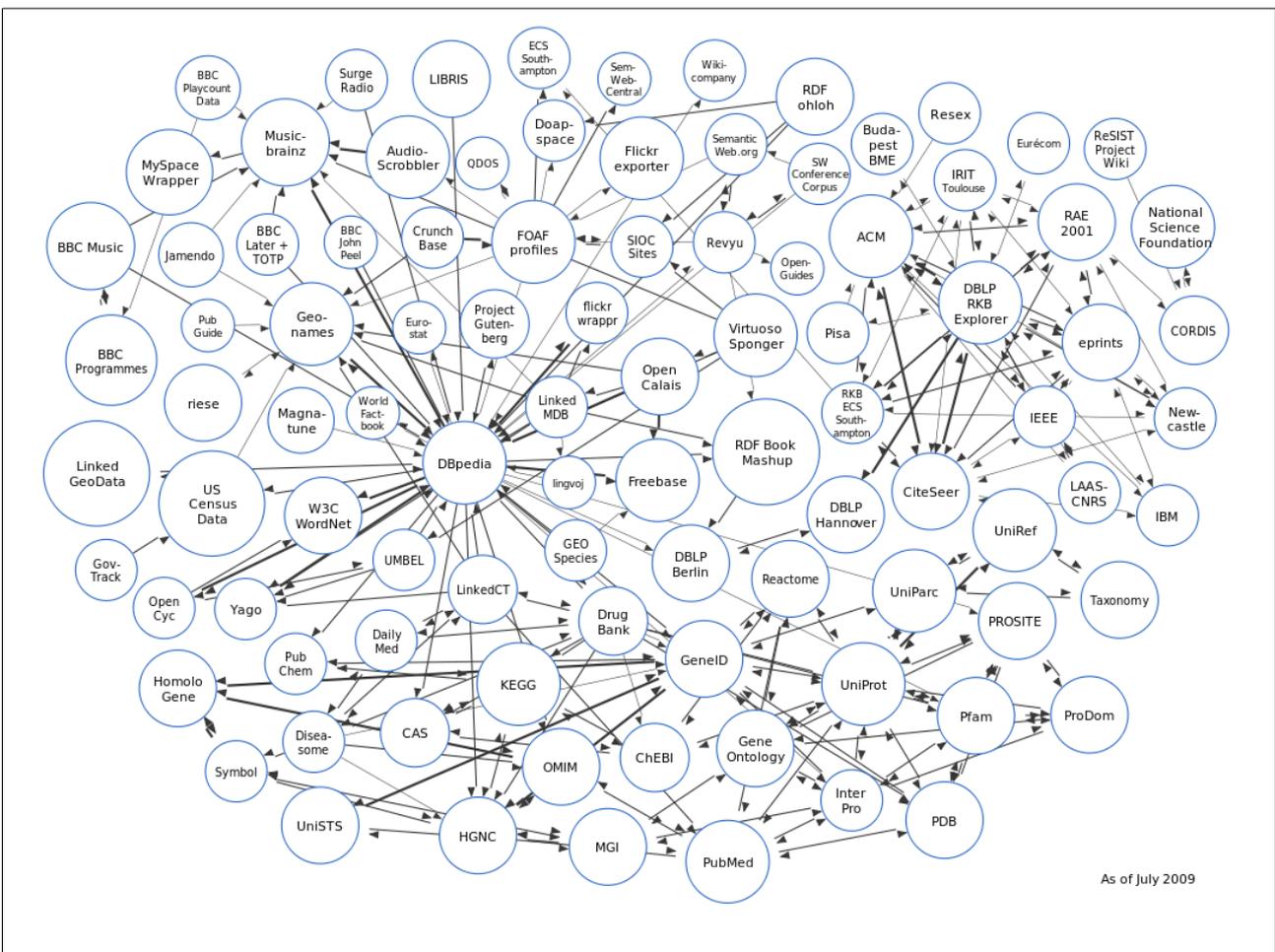


fig.14 Linking Open Data cloud diagram 2007, by Richard Cyganiak

[CC BY-SA 3.0 (<http://creativecommons.org/licenses/by-sa/3.0>)], attraverso Wikimedia Commons



As of July 2009

fig.15 Linking Open Data cloud diagram 2009, by Anja Jentzsch and Richard Cyganiak

[CC BY-SA 3.0 (<http://creativecommons.org/licenses/by-sa/3.0>)], attraverso Wikimedia Commons

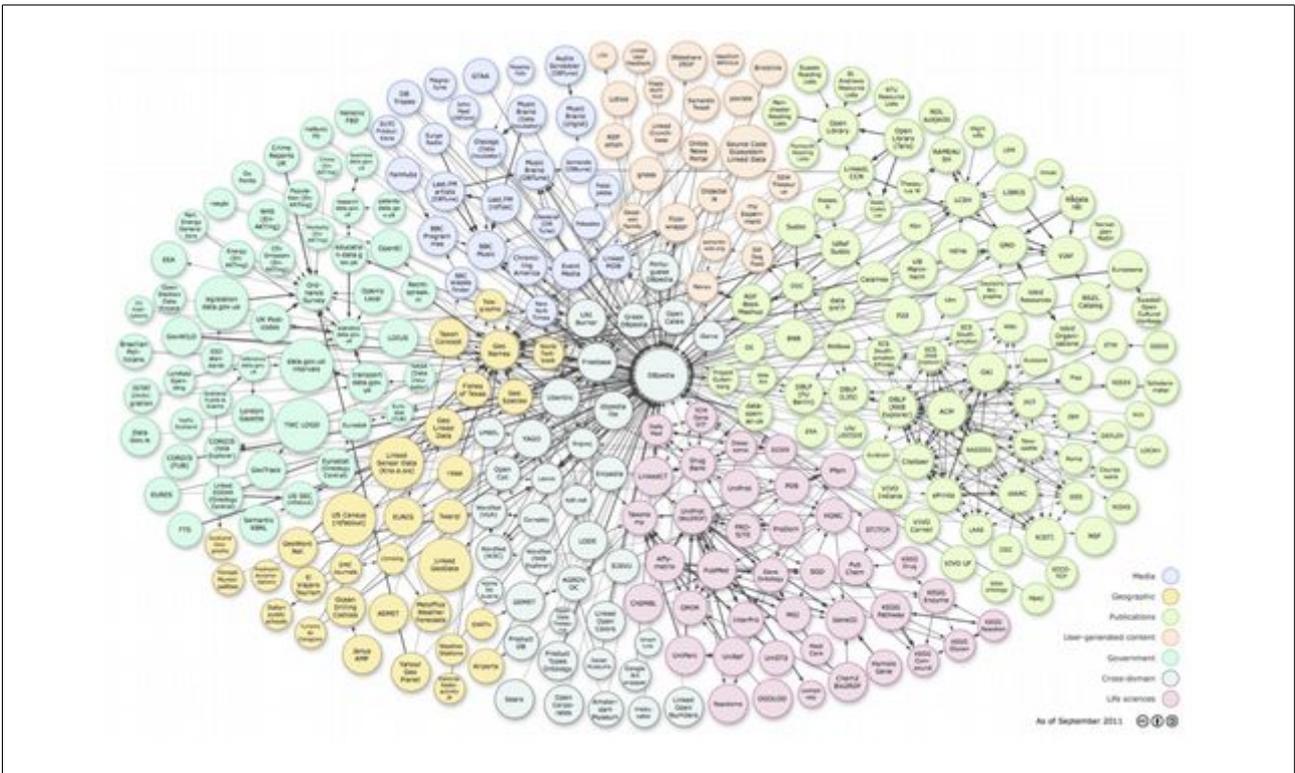


fig.16 Linking Open Data cloud diagram 2011, by Anja Jentzsch

[CC BY-SA 3.0 (<http://creativecommons.org/licenses/by-sa/3.0>)], attraverso Wikimedia Commons

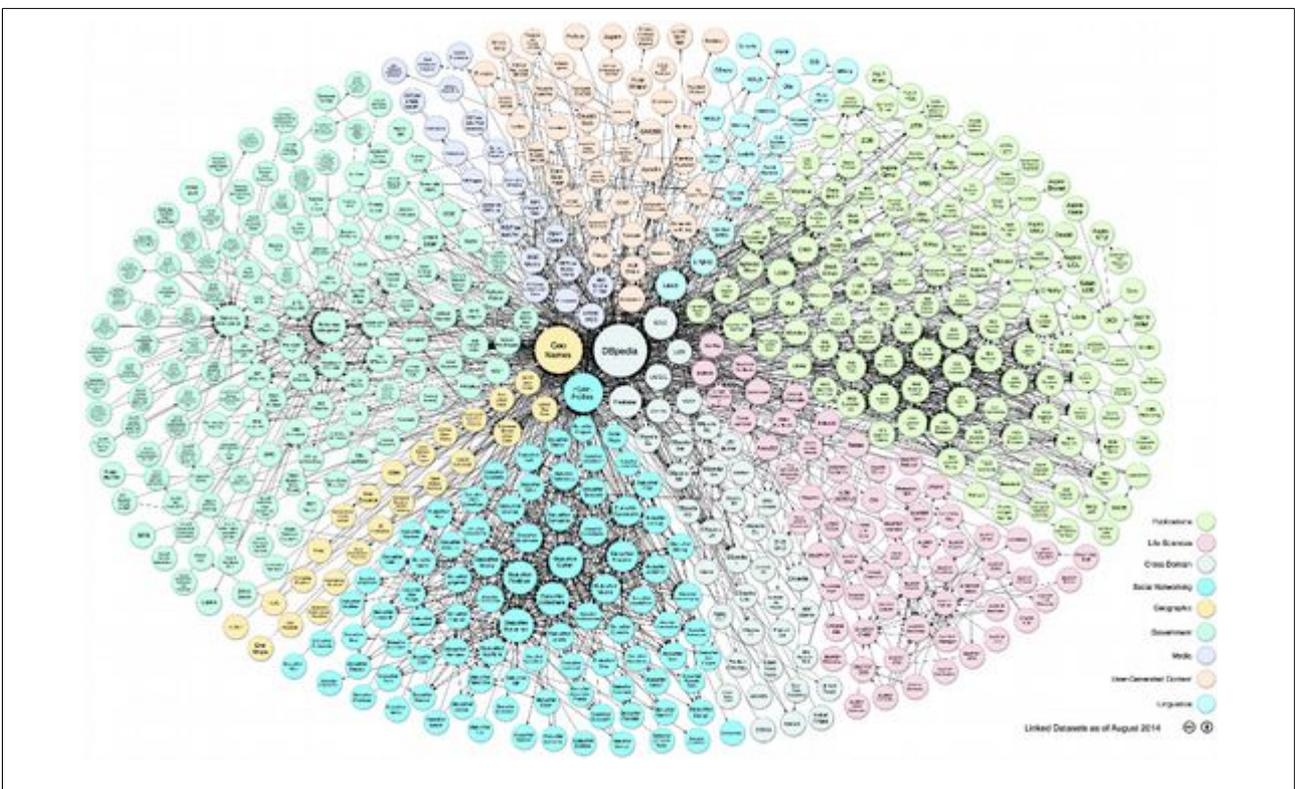


fig.17 Linking Open Data cloud diagram 2014, by Max Schmachtenberg, Christian Bizer, Anja Jentzsch and Richard Cyganiak. <http://lod-cloud.net/>

sottoposti a revisione, *debugging* o sviluppo nell'ambito di una attività collaborativa. Essendo infatti aperti, essi possono, senza restrizioni, divenire oggetto di collegamenti tra *dataset* diversi, oppure essere semantizzati tramite collegamenti a vocabolari ed ontologie. Il carattere collaborativo in questo caso è garantito dal fatto che *chiunque* può istituire tali collegamenti.

In un contesto simile urge affrontare la questione dell'attendibilità delle informazioni pubblicate. Come abbiamo già osservato, la cultura *open* propone un diverso approccio nei confronti dei contenuti; nel caso dell'*open access* per esempio, abbiamo visto che, nei casi in cui viene meno il controllo di qualità *ex-ante*, la politica di apertura consegna all'utente gli strumenti per valutare da sé l'attendibilità dei documenti consultati (indicizzazione dei contenuti, nuovi sistemi di analisi citazionale, *social peer-review*). D'altra parte, e anche in seguito alla constatazione che tali meccanismi stentano ad affermarsi, le istituzioni della memoria devono assumere un ruolo di primo piano nella gestione delle fonti e della loro affidabilità.

The quantity and quality of the information that populates the net are two aspects which are often inversely proportional: much information is of poor quality. The increase in networked information (through publication methods that are increasingly widely-known and used, such as for example, self-publishing, social networks) is not, in fact, always synonymous with quality. The exponential growth and use of information available on the net does not coincide with increasing trustworthiness of the records either: their degree of reliability is low. [...] On which criterion to base the selection? The authoritativeness of the source becomes the key factor, the selection takes place at the outset, preferring to select a resource on the basis of the authoritativeness of its creator, instead of later on, choosing uncritically on the basis of the ranking of the records that appear on the page. The quality of the source, the certainty of the provenance become, therefore, crucial elements in the searcher's exploratory process. (Guerrini e Possemato, 2013, pp. 75-76)

Anche se consentono una prima valutazione *coarse-grained*, gli algoritmi di *ranking* non possono essere impostati come unico criterio di verifica di attendibilità; a questo proposito si sta lavorando nel disegnare interfacce che forniscano direttamente agli utenti gli strumenti per accertare la provenienza dei dati e l'affidabilità dei *data provider*¹⁰⁴.

Nel caso del *semantic web* la questione dell'affidabilità è relativa non solo alle fonti, ma anche ai risultati di processi inferenziali automatici. I *layers* superiori del *Semantic Web Stack* riguardano proprio la definizione di regole logiche unificate (“Unifying Logic”) per programmare processi

104 Si veda la sezione “Trust, Quality and Relevance” in Berners-Lee, Bizer e Heath (2009, p. 19).

inferenziali complessi, e di tecniche per sottoporre i risultati a processi di validazione. I termini “Proof” e “Trust” si riferiscono a

livelli di controllo e validazione applicati dal sistema per verificare che i dati provengano da fonti affidabili (trusted), i quali, uniti all’applicazione dei principi della logica formale per elaborare i dati derivati, consentono di fornire risultati univoci ed affidabili, garantendo un elevato livello di attendibilità agli utilizzatori (Guerrini e Possemato, 2015, p. 232).

Il lavoro in questi ambiti è ancora a uno stadio iniziale e sarà interessante seguirne gli sviluppi nei prossimi anni.

Aderire alle tecnologie dei LOD e del *semantic web* implica anche relazionarsi con le tematiche e le politiche dell’*open content*. La libertà dei contenuti, la disponibilità dei dati e la fiducia in metodi di sviluppo collaborativi vanno di pari passo con la sensazione di perdita del controllo dei meccanismi di circolazione del sapere. È necessario che le istituzioni della memoria si appropriino delle nuove tecniche per non trovarsi impreparate di fronte all’effettivo “vuoto di potere” generato dal mutamento dei mezzi tecnologici.

Seconda parte

MIMO. Nuove tecnologie per la catalogazione di strumenti musicali

Il progetto MIMO

Obiettivi e strategie

Nel campo della catalogazione di strumenti musicali il primo progetto che ha intrapreso la strada della pubblicazione dei dati in modalità LOD è MIMO (Musical Instrument Museums Online) che dal 2009 al 2011 ha digitalizzato, raccolto e pubblicato i dati e le risorse di circa 50000 strumenti provenienti dalle collezioni di diversi musei europei.

MIMO ha creato una piattaforma di *harvesting* dei metadati in formato XML tramite il protocollo OAI-PMH e ha pubblicato i dati e le risorse raccolte su due siti web¹; fornisce inoltre i suoi dati in modalità LOD a Europeana, la biblioteca digitale europea². MIMO ha redatto tre vocabolari funzionali al *metadata enrichment*: una lista controllata di *Instrument Makers*, una tassonomia per la sistematica Hornbostel-Sachs (HS) e un *thesaurus* multilingue di *keywords* relative agli strumenti.

MIMO è un progetto che ha lo scopo di creare, tramite Europeana, «a single access point to information and digital content on the collections of musical instruments held in European museums»³. Esso risponde ad una serie di esigenze individuate in fase progettuale: MIMO intende colmare le lacune create dall'assenza di un punto di accesso comune alle collezioni degli strumenti musicali online, e di uno standard condiviso per la loro pubblicazione. Prima dell'inizio del progetto il patrimonio organologico pubblicato in rete, considerato nel suo complesso, era accessibile principalmente a ricercatori ed esperti del settore. I contenuti erano generalmente pubblicati solo nella lingua d'origine del Museo e le eventuali traduzioni presenti non coprivano tutto il contenuto pubblicato (MIMO, 2009, p. 3).

Un obiettivo chiave di MIMO è stato promuovere un accesso multilingue ai contenuti, che garantisca consistenza di nomenclatura nelle 6 lingue dei partner fondatori del progetto: francese, tedesco, italiano, olandese, svedese e inglese⁴. Contestualmente all'elaborazione dei vocabolari, è stata creata una piattaforma capace di gestire un'implementazione multilingue. L'obiettivo esplicito di coinvolgere il massimo numero di visitatori online ha influenzato i criteri per la compilazione dei

1 <http://www.mimo-international.com/> è il sito ufficiale; <http://www.mimo-db.eu/> è l'interfaccia di *back-office* del database a cui si ha liberamente accesso, con *tools* limitati.

2 Europeana - <http://www.europeana.eu/>

3 Rodger (2011, p. 4). La maggior parte delle informazioni contenute nella presente illustrazione del progetto sono tratte da Rodger (2011), il *Final Report* che ne descrive premesse, svolgimento e *outcomes*.

4 Le lingue di MIMO sono diventate otto in seguito all'adesione di due nuove istituzioni museali la cui lingua (il catalano e il polacco) non era ancora rappresentata.

vocabolari che includono, a fianco di nomenclature specialistiche (la sistematica HS), anche nomi di “uso comune”; si è ritenuto che in questo modo anche l'utente medio potesse avere a disposizione strumenti per compiere efficaci ricerche sia incrementali che esplorative⁵.

Il progetto si è impegnato nell'implementazione delle più moderne tecnologie relative ai *Linked Open Data* applicando, soprattutto nell'elaborazione dei vocabolari, i metodi di lavoro di carattere collaborativo tipici di queste tecnologie. È molto importante segnalare che la piattaforma elaborata da MIMO per la raccolta e la pubblicazione dei contenuti è tuttora attiva: essa è stata concepita in modo da permettere ad altri musei di aderire all'iniziativa una volta che fosse scaduto il termine del suo sviluppo.

Il progetto MIMO, finanziato in parte dal programma *eContentplus* dell'Unione Europea, si è svolto dal 1 settembre 2009 al 31 agosto 2011, e ha previsto un livello di sostenibilità a lungo termine, con scadenza il 31 agosto 2016. Entro tale data è stato possibile aderire all'iniziativa avvalendosi dei contributi stanziati dai partner fondatori. Ora per associarsi al progetto è necessario versare un contributo diretto al mantenimento della piattaforma informatica.

Il consorzio promotore è formato da 11 istituzioni (museali ed accademiche) europee:

- University of Edinburgh;
- Germanisches Nationalmuseum, Nürnberg;
- University of Leipzig, Museum für Musikinstrumente;
- Africamuseum, Tervuren;
- Associazione "Amici del Museo degli Strumenti Musicali", Firenze;
- Cité de la Musique, Paris;
- Musical Instrument Museum, Brussels;
- Università degli Studi di Firenze;
- Horniman Museum, London;
- Stiftung Preußischer Kulturbesitz, Ethnologisches Museum;
- The Stockholm Music Museum⁶.

5 In Bianchini (2012) si distinguono due tipi di ricerche in base ai bisogni dei lettori: «1) la ricerca esplorativa (*exploratory search*), nella quale non si hanno preferenze relative al tema, né si ha familiarità con le parole chiave. Per iniziare questa ricerca si ha quasi sempre bisogno di aiuto e lo scopo è ottenere informazioni generali; 2) la ricerca consapevole o incrementale (*known item search*), nella quale si ha una specifica risorsa in mente (un articolo, una rivista, un libro ecc.), si ha familiarità con le parole chiave o i termini della ricerca e si sa da dove partire (sito web, database ecc.)» (p. 304). La strutturazione dei vocabolari può essere finalizzata a supportare entrambi i tipi di ricerca.

6 È interessante notare che queste istituzioni si occupano di settori organologici non omogenei (liuteria classica, strumenti popolari, elettronici, extra-europei ecc.).

Hanno affiancato il consorzio quattro *Associate Partners*:

- Musikmuseet, Copenhagen;
- Technisches Museum, Wien;
- Pokrajinski Muzej, Ptuj;
- Palais Lascaris-Musée de la musique de Nice.

Il progetto si è definito nell'individuazione di 6 *Work Packages*, ciascuno dei quali guidato da una istituzione leader:

- WP1: Digitalizzazione - Lead Partner: Germanisches Nationalmuseum di Norimberga;
- WP2: OAI *harvesting*, sviluppo database e interoperabilità con Europeana - Lead Partner: Cité de la Musique di Parigi;
- WP3: *Thesaurus* e Classificazione - Lead Partner: Horniman Museum di Londra;
- WP4: Project Management - Lead Partner: Università di Edinburgo;
- WP5: Valutazione - Lead Partner: Africamuseum di Bruxelles;
- WP6: Disseminazione - Lead Partner: Ethnologische Museum di Berlino.

Le due istituzioni italiane hanno avuto il ruolo di *content provider* (Associazione “Amici del Museo degli Strumenti Musicali” di Firenze) e collaboratore al WP3 *Thesaurus* e Classificazione (Università degli Studi di Firenze).

La maggior parte dei partner di MIMO sono membri del CIMCIM⁷, il comitato dedito allo studio e alla conservazione degli strumenti musicali che afferisce all'ICOM⁸, la più importante organizzazione internazionale di musei e professionisti museali. Come previsto e auspicato dai partner del progetto (Rodger, 2011, p. 40) si è assistito a un crescente interesse del CIMCIM verso MIMO, che si è tradotto in un impegno attivo nella fase post-progettuale (in particolare in riferimento allo sviluppo del sito <http://www.mimo-international.com/MIMO>).

Allo scadere del progetto (31 agosto 2011) il database di MIMO ha collezionato materiale da 21 musei (di cui 9 partner) per un totale di 50815 record di strumenti, dei quali 48273 provenienti dalle collezioni dei partner (Bailly, 2011, p. 13).

7 CIMCIM (Comité International des Musées et Collections d'Instruments et de Musique) <http://network.icom.museum/cimcim/>

8 ICOM (International Council of Museums) - <http://icom.museum/>

Il 15 luglio 2016 i *data provider* sono saliti a 30⁹, per un totale di 55984 record¹⁰.

MIMO ha posto come obiettivo a lungo termine la creazione di centri di aggregazione nazionale che favoriscano l'adesione al progetto di istituzioni di dimensioni e risorse ridotte; il primo di questi centri è in fase di sviluppo nel Regno Unito.

9 In aggiunta ai 9 *content provider* iniziali, hanno aderito a MIMO: il Museu de la Musica di Barcellona, il Royal College of Music Museum di Londra, The Norwegian Museum Of Cultural History di Oslo, il Rijksmuseum di Amsterdam e The Nydhal Collection di Stoccolma. È presente anche un nutrito gruppo di musei francesi (Musée Du Palais Lascaris - Nice, Musée De L'hospice Comtesse - Lille, Musée Auguste Grasset - Varzy, Pôle Accordéons - Tulle, Musée Des Instruments À Vent - La Couture-Boussey, Museon Arlaten - Arles, Musée De La Musique Mécanique - Les Gets, Musée Des Instruments De Musique - L'aigle, Musée De La Lutherie Et De L'archèterie - Mirecourt, Musée De La Castre - Cannes, Musée Des Musiques Populaires – Montluçon) e un consorzio di collezioni polacche che hanno aderito recentemente seguendo un percorso comune (Institute Of Music And Dance - Warsaw, Museum Of Folk Musical Instruments - Szydlowiec, National Museum - Poznan, The State Ethnographic Museum - Warsaw, Jadwiga And Marian Sobieski Collection – Szydlowiec).

10 I dati aggiornati sono pubblicati sulla homepage del sito <http://www.mimo-international.com/>.

La disseminazione online

Le informazioni sul progetto e i dati pubblicati da MIMO sono disponibili su diversi siti web:

- <http://www.europeana.eu/>

è l'unico strumento di pubblicazione previsto da MIMO in fase progettuale (MIMO, 2009, p. 3).

Il percorso di caricamento dei dati su Europeana è stato particolarmente travagliato a causa di ritardi nella mappatura da un formato all'altro (Rodger, 2011, p. 16; Bailly, 2011, pp. 8-10). Il risultato è stato infine raggiunto ed ora tutti i dati caricati progressivamente sul database di MIMO sono automaticamente raccolti da Europeana e mappati in EDM (Europeana Data Model) il più recente formato disponibile.

Il fatto che, nel corso del progetto, Europeana non avesse ancora pubblicato i dati già pronti sugli strumenti musicali, ha spinto MIMO a cercare nuove strategie che hanno condotto alla presentazione di uno *showcase* all'interno del sito <http://www.mimo-project.eu/> e di una *virtual exhibition* sul portale Europeana.

Al di là delle iniziali complicazioni, Europeana consente a MIMO di implementare i propri metadati in modalità LOD, di incrementare il numero di vocabolari collegati e di acquisire visibilità all'interno di un portale noto che offre all'utente non specialistico una impressione di “navigabilità” tra i contenuti.

- <http://www.mimo-international.com/>

il sito, pubblicato nel marzo 2014, è una piattaforma autonoma indipendente da Europeana, intesa come il *focus* della disseminazione *post-project*. È destinato al pubblico più ampio, in funzione del quale sono stati concepiti veste grafica, livello di usabilità e navigabilità. È collegato al database di MIMO ed è progettato per fornire una «full search capability» (Rodger, 2011, p. 25). Consente di effettuare ricerche esplorative e incrementali a partire dal *thesaurus* delle *keywords* (disponibile nelle otto lingue di MIMO più il cinese) e dall'elenco dei musei *data provider*; in alternativa si può iniziare la ricerca utilizzando una comune barra di *quick search*. I contenuti indicizzati per la ricerca sono: il *thesaurus* delle *keywords*, la classificazione HS, la lista controllata di *Instrument Makers*, l'elenco dei musei *data provider*, luogo e data di costruzione. I risultati riportano un elenco di *item* che soddisfano i criteri cercati, e, sulla colonna sinistra, un elenco di tutti gli indici mediante i quali è possibile affinare ulteriormente la ricerca. La sistematica HS è l'unico criterio indicizzato a non

comparire sulla colonna sinistra, né è possibile consultarne la tassonomia: nel caso della HS non sono implementate le ricerche esplorative e incremental; in altre parole è necessario conoscere esattamente *a priori* il *taxon* o il numero Hornbostel-Sachs da cercare.

La pagina “How to join us” è dedicata alle istituzioni interessate ad aderire al progetto, ed offre informazioni, risposte alle FAQ, contatti e documenti (*Deliverable*) contenenti le specifiche tecniche.

Il sito è stato realizzato dalla Philharmonie di Parigi e finanziato da cinque partner MIMO: University of Edinburgh, Cité de la musique, Germanisches Nationalmuseum, Musical Instrument Museum di Bruxelles e Associazione “Amici del Museo degli Strumenti Musicali”; le pagine web fanno riferimento anche a contributi elargiti dal CIMCIM. <http://www.mimo-international.com/> è inteso come il primo passo verso la creazione di un portale MIMO indipendente da Europeana e volto a incoraggiare l'adesione di musei extra-europei.

www.mimo-db.eu

è l'interfaccia di *back-office* del database; nel corso del progetto si è deciso di renderne pubblico l'accesso con *tools* limitati, senza password né login. La piattaforma offre a qualsiasi utente la possibilità di effettuare ricerche avanzate. I criteri di ricerca a disposizione sono più numerosi rispetto a quelli del sito “ufficiale”. Per esempio si può impostare la ricerca in base al tipo e alla data di un determinato “Evento”, al tipo di “Media” allegato al record e, soprattutto, in base alla tassonomia HS la cui struttura è consultabile preventivamente. I risultati della ricerca sono ricchi di informazioni ed è possibile visualizzare il record XML originale.

Il sito permette inoltre di esplorare i vocabolari delle *keywords* e della sistematica HS sia nella pagina di ricerca avanzata che attraverso “Vizkos”, uno specifico *tool* di visualizzazione di SKOS, il linguaggio in cui essi sono rappresentati.

<http://exhibitions.europeana.eu/exhibits/show/musical-instruments-en>

“Explore the World of musical instruments” è una *user-friendly virtual exhibition* presentata da MIMO su Europeana. È stata pubblicata nel corso del progetto (2009-2011) per mantenere vivo l'interesse del pubblico, nonostante Europeana non avesse ancora caricato i dati di MIMO; è tuttora accessibile.

MIMO ha considerato questa *exhibition* come l'occasione per «reviewing Europeana Music themes, and those that provided social contexts and events which could be linked to object records in order to enrich them» (Rodger, 2011, p. 20).

La *virtual exhibition* introduce i *Music themes* con queste parole:

Musical instruments play an important part not only in art, but also in many other social contexts, such as religion, the military, celebrations and ceremony. [...] This exhibition, drawn from the collections of nine of Europe's major musical instrument museums, offers a selection of instruments that illustrates a range of styles and cultural uses.

Il risultato è molto accattivante in quanto a originalità degli strumenti, qualità delle risorse e design delle pagine ma il tema del contesto sociale, la cui revisione aprirebbe questioni importanti dal punto di vista della ricerca, potrebbe essere approfondito con maggior rigore scientifico.

<http://www.mimo-international.com/mimo-db.html>

“MIMO Toolkit” è una presentazione del progetto che contiene tutte le informazioni e i documenti utili alle istituzioni interessate all'adesione.

www.mimo-project.eu

era il sito ufficiale attraverso cui MIMO pubblicava aggiornamenti sull'avanzamento del progetto. Ora non è più disponibile.

La piattaforma tecnica e i protocolli di trasmissione

Il *Work Package 2*, guidato dalla Cité de la Musique, ha predisposto la piattaforma tecnica per l'invio e la raccolta dei dati.

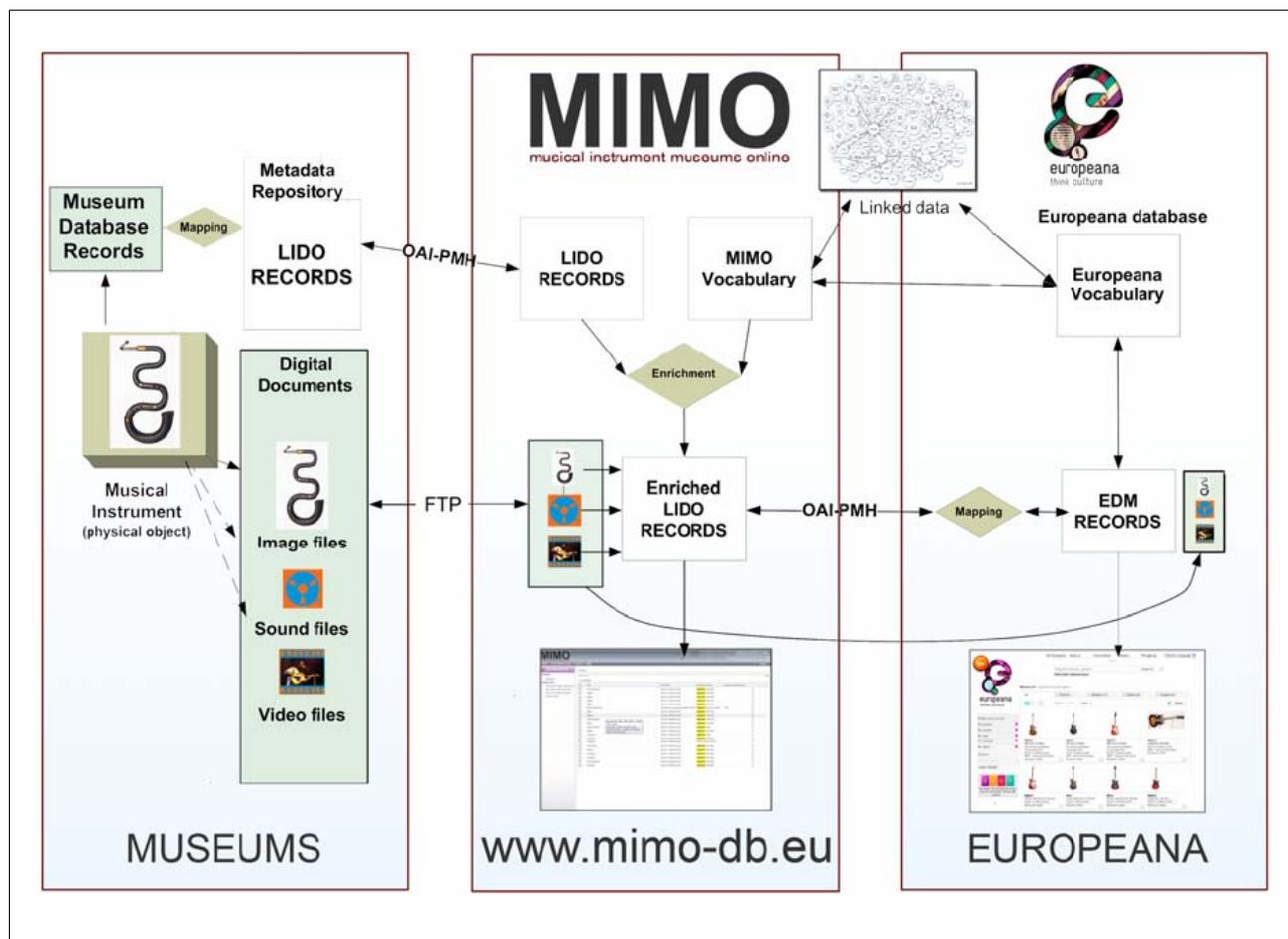


fig.1 The MIMO harvesting process

Norman Rodger (2011, p. 15)

Per ogni strumento della collezione il museo dispone di un record contenente i metadati stoccati in un database, e di riproduzioni digitali (fotografie e file audiovisivi); il museo deve occuparsi di effettuare un *mapping* del proprio database nel formato LIDO (il modello di descrizione dei dati adottato da MIMO), ed assicurarsi che la documentazione digitale sia conforme allo standard elaborato dal *Work Package 1*, guidato dal Germanisches Nationalmuseum di Norimberga (Bär e Pfefferkorn, 2011).

MIMO ha creato un database per l'aggregazione dei dati ("mimo-db") su un server della Cité de la Musique. Qui sono caricati i documenti digitali via FTP, e i *record* in formato LIDO via OAI-PMH¹¹. Durante la raccolta avviene un processo di *metadata enrichment*: il sistema cerca di stabilire collegamenti fra specifici campi dei record LIDO, i vocabolari MIMO e i termini di GeoNames; ne risulta che i record sono *arricchiti* in quanto collegati alla rete dei *Linked Open Data*. I dati arricchiti sono ricondotti alle risorse digitali, stoccati nel database e pubblicati su <http://www.mimo-db.eu/>; da qui sono successivamente caricati su <http://www.mimo-international.com/>¹².

Europeana raccoglie i record arricchiti in formato LIDO via OAI-PMH, effettua una mappatura da LIDO a EDM¹³, e li collega ai propri vocabolari LOD e alle risorse digitali. Al termine di questo processo i record relativi agli strumenti musicali sono pubblicati sulla biblioteca digitale europea.

Secondo le linee-guida elaborate dal *Work Package 2* (Bailly e Le Meur, 2010) i musei possono scegliere a propria discrezione una delle tre tecniche di implementazione di OAI-PMH indicate dall'Open Archive Initiative; esse sono, in ordine di decrescente complessità, le tecniche del repository dinamico, fisico e statico.

L'implementazione di un repository dinamico o fisico è un procedimento tecnicamente impegnativo, che potrebbe rappresentare un ostacolo all'adesione al progetto, soprattutto per i musei di dimensioni e risorse ridotte. Per questo motivo MIMO offre la possibilità di usare la tecnica più semplice del repository statico, per la quale mette a disposizione uno *static repository gateway*.

11 Una specifica interfaccia grafica permette agli amministratori di "mimo-db" di settare nuove fonti di raccolta (nuovi musei), il tipo di raccolta (completa o incrementale) e la frequenza di aggiornamento. Il meccanismo di *harvesting* è completamente automatizzato, e i repository di ogni partner sono periodicamente *harvested* in modo da recuperare ogni modifica ai contenuti.

12 Quest'ultimo passaggio non è presente nella fig.1: il sito è stato creato dopo il termine del progetto, di cui il *Final Report* da cui è tratta l'immagine (Rodger, 2011) è un resoconto.

13 EDM (Europeana Data Model). Quando MIMO era in fase di chiusura del progetto, Europeana stava ancora sviluppando EDM, il suo modello di descrizione dei dati orientato al *semantic web*. È stato necessario un periodo ulteriore di lavoro per completare la mappatura tra i due formati, che avviene ora attraverso un procedimento automatico descritto in un file XSLT (Bailly, *EDM Case Study*).

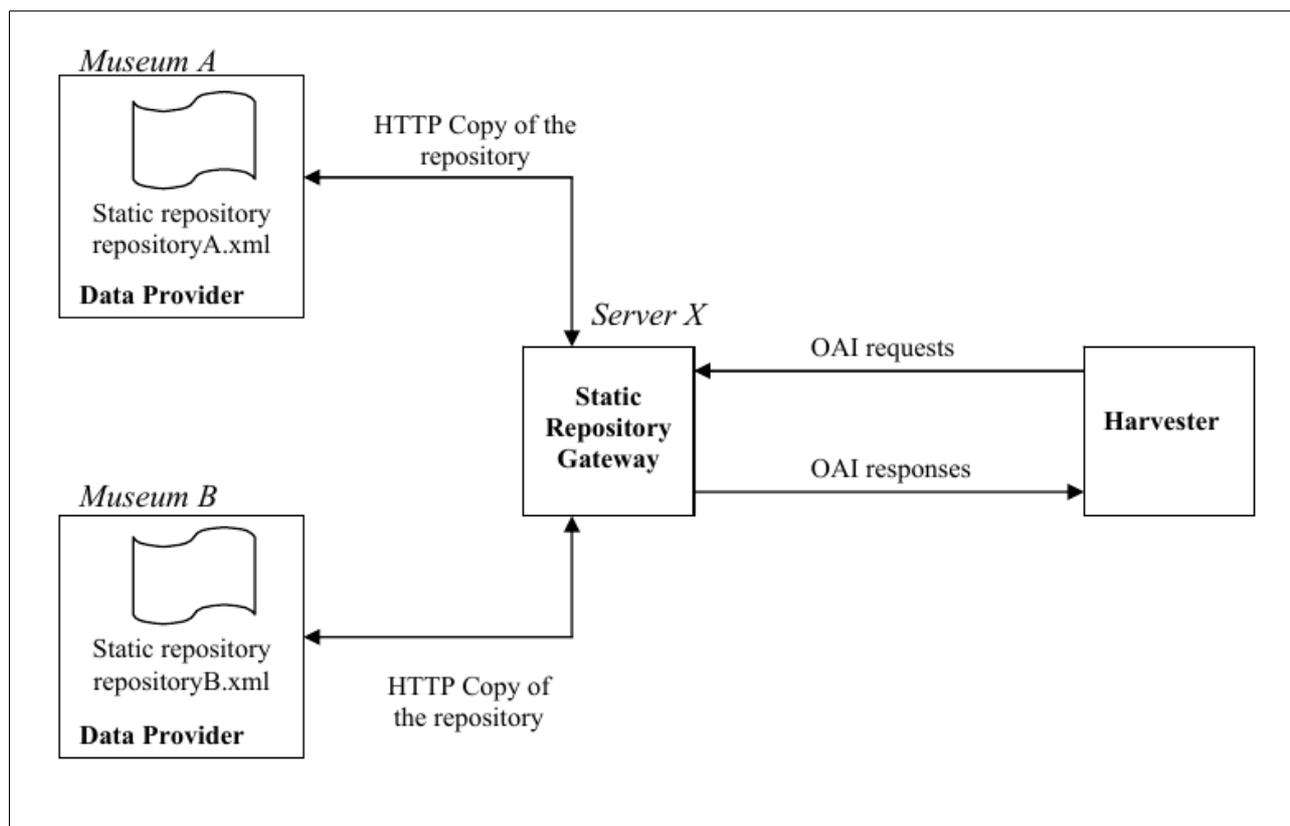


fig.2 OAI-PMH Static repository
(Bailly e Le Meur, 2010, p. 45)

Questa tecnica è stata proposta dall'Open Archive Initiative proprio per le collezioni di piccole dimensioni (indicativamente, fino a 5000 record) e relativamente statiche¹⁴; in questo caso il *data provider* è tenuto esclusivamente a compilare un file XML (contenente tutti i record), garantire che esso sia presente a un URL HTTP persistente e fornire alcune semplici informazioni necessarie all'implementazione. Il repository statico quindi non è un vero repository OAI-PMH: è solo un file XML, non un server che sappia processare le sei *request* del protocollo. Questo lavoro di intermediazione è svolto dallo *static repository gateway* predisposto da MIMO; attraverso di esso ogni repository statico è esposto *come se* fosse un repository OAI-PMH individuale¹⁵.

14 Le linee-guida per l'implementazione del repository statico sono pubblicate sul sito dell'Open Archive Initiative: <http://www.openarchives.org/OAI/2.0/guidelines-static-repository.htm>.

15 Il *gateway* assegna ad ogni *static repository* un unico "base URL" preceduto dal prefisso URL del *gateway*.

In riferimento alla fig.2:

«Museum A exposes a static repository which URL is

<http://www.MuseumA.org/repositoryA.xml>

Museum B exposes a static repository which URL is

<http://www.MuseumB.org/repositoryB.xml>

The static repository gateway URL is

<http://gateway.WebsiteX.org/oai>

MIMO assume quindi il ruolo di *service provider* all'interno di un'architettura OAI-PMH: un raccogliitore di metadati che offre agli utenti un servizio aggiunto. Nelle linee-guida MIMO individua nel protocollo OAI-PMH l'occasione per fare un primo passo verso le pratiche *open access*, in quanto i musei possono costituirsi come *data provider* anche verso altri aggregatori:

Implementation of OAI-PMH on a Museum Collection Management System can be the starting point of multiple automated data exchange with other information systems located in other Institution, or even in the same institution as the Museum, such as the institutional web site (Bailly e Le Meur, 2010, p. 42).

In una nota seguente, si incoraggiano i *data provider* ad esporre i propri dati sia in LIDO (che è il formato obbligatorio per MIMO) che in Dublin Core, con la consapevolezza che quest'ultimo è il formato richiesto per essere compatibili con la crescente comunità di *open data provider*. L'adozione della tecnica del repository statico comporta però alcune limitazioni: per evitare ridondanze, ogni *data provider* può esporre i propri dati ad un unico *gateway*¹⁶. Affinché sia davvero implementata la rete dell'*open access*, ovvero che i musei siano visibili via OAI-PMH ad altri *service provider* oltre che a MIMO, è necessario che essi si assicurino che non ci siano ostacoli all'esposizione pubblica dello *static repository gateway*.

The harvester will act as if it addresses 2 different repositories of which URLs are

<http://gateway.WebsiteX.org/www.MuseumA.org/repositoryA.xml> and

<http://gateway.WebsiteX.org/www.MuseumB.org/repositoryB.xml>» (Bailly e Le Meur, 2010, p.45)

16 Inoltre non si possono implementare set, compressioni e *resumption token*. Di conseguenza un *service provider* non può fare *harvesting* selettivo sui set (la categoria di appartenenza dei record), né viene promossa questa pratica.

LIDO: lo schema per la strutturazione dei metadati

Il modello e l'applicazione all'ambito degli strumenti musicali

Il *Work Package 2* si è occupato di individuare il modello di descrizione dei dati più funzionale al contesto tecnologico di riferimento; come descritto in MIMO (2009), il documento che ne espone le linee progettuali, MIMO ha sin dall'inizio previsto l'adozione di OAI-PMH per il trasferimento dei dati tra le varie piattaforme online (*ivi*, p. 23); di conseguenza la scelta di un modello di descrizione dei dati nel formato XML era obbligata¹⁷.

Anche se il proposito iniziale prevedeva l'uso a questo scopo di *Qualified Dublin Core*¹⁸, MIMO ha preso in considerazione anche altri formati ed ontologie (*ibidem*); come descritto in Le Meur (2010) infatti «analysis of common data models and standards used in various Museums database software, such as MUSEUM DAT, CDWA Lite, Spectrum, CIDOC CRM led us to LIDO, a new model currently designed by eContentplus project ATHENA» (p. 7).

LIDO¹⁹ è uno schema XML concepito appositamente per il trasferimento dei dati delle collezioni museali su diverse piattaforme online; è stato sviluppato dal CIDOC²⁰ dell'ICOM²¹ nell'ambito di diversi progetti che hanno coinvolto una comunità internazionale di professionisti del settore museale. Lo scopo di LIDO è fornire uno strumento di interoperabilità per la trasmissione online dei metadati relativi agli oggetti museali: in questo senso esso definisce un modello standard di descrizione studiato specificamente per l'*harvesting*.

LIDO non è stato quindi strutturato per fungere da base di un *Collection Management System*; per esempio, non supporta le attività di prestito ed acquisizione. Ne deriva che ogni istituzione museale

17 OAI-PMH trasferisce esclusivamente metadati in formato XML.

18 *Dublin Core* è un set di 15 termini che descrive qualsiasi risorsa digitale accessibile in via informatica. È stato sviluppato dal DCMI (Dublin Core Metadata Initiative) nell'ambito dell'OCLC (Online Computer Library Center), il consorzio di biblioteche americano che ha sede a Dublin, nell'Ohio. I termini del *Dublin Core* e del *Qualified Dublin Core* (che ne è un'estensione) sono incorporati nel set *DCMI Metadata Terms*. L'implementazione dei *DCMI Metadata Terms* in diversi formati del web semantico (tra cui RDF/XML) ha reso *Dublin Core* uno dei vocabolari più diffusi nell'ambito dei *Linked Open Data*.

19 LIDO (Lightweight Information Describing Objects) - <http://www.lido-schema.org/>.

20 Il CIDOC (Comité International pour la Documentation) è un comitato dell'ICOM: <http://network.icom.museum/cidoc/>. All'interno del CIDOC c'è un *Working Group* che si occupa specificamente dello sviluppo di LIDO: <http://network.icom.museum/cidoc/working-groups/lido/>.

21 L'ICOM è la più importante organizzazione internazionale di musei e professionisti museali; è importante tenere in considerazione il fatto che i modelli, gli standard e le *policies* promossi da un'organizzazione così influente sono naturalmente intrecciati: la modellazione di LIDO per esempio è stata influenzata dalla compatibilità con l'ontologia CIDOC-CRM, sviluppata anch'essa da un comitato dell'ICOM.

è invitata a conservare il proprio metodo di catalogazione locale, a partire dal quale è possibile effettuare un'operazione di *data mapping* verso LIDO²².

Il modello è stato elaborato in funzione della descrizione di oggetti museali generici, con l'obiettivo di includere «all kinds of object, e.g. art, architecture, cultural history, history of technology, and natural history» (Coburn, Light, McKenna, Stein e Vitzthum, 2010, p. 3).

MIMO ha quindi organizzato workshop con tutti i *content provider* al fine di testare l'uso di LIDO nel campo specifico degli strumenti musicali.

Outcomes of these workshops helped us to delimit a framework for the use of LIDO in MIMO's platform. This framework defines:

- Applicable fields;
- Specific lists of terms (ie: list of Event Types);
- Missing fields²³.

In seguito la documentazione non fa più riferimento alla soluzione di problemi relativi a *missing fields*. MIMO usa una parte (molto consistente) degli elementi ed attributi che costituiscono l'XML Schema di LIDO. In questo senso la struttura delimitata da Le Meur (2010) è un sottoinsieme di LIDO:

The MIMO data model is based on LIDO (Lightweight Information Describing Objects) version 0.9 draft, the XML schema for contributing content to cultural heritage repositories.

In this version of the document, all specified elements are extracted from the LIDO data model. [...]

In the description of an element, we have chosen to list only the attributes and sub-elements belonging to this element, which are necessary for MIMO data model and not all the attributes and sub-elements as they are specified in LIDO model (p. 8).

Le Meur (2010) indica come *MIMO specificities* quegli elementi o attributi della struttura MIMO che si differenziano dalle specifiche LIDO. Si tratta comunque sempre di *restrizioni* all'interno delle possibilità offerte da LIDO; ad esempio:

- elementi o attributi MIMO *obbligatori* entro le possibilità di elementi o attributi LIDO *opzionali*;
- elementi MIMO *non ripetibili* entro le possibilità di elementi LIDO *ripetibili*;
- valori obbligati o scelti all'interno di una lista raccomandata da MIMO, entro le possibilità di *qualsiasi* termine di LIDO.

22 Se possibile, è consigliabile programmare un *automated data mapping*, un processo automatico di conversione dei metadati da un formato all'altro. Per attuare questo procedimento, che accorcia drasticamente i tempi e riduce le possibilità di errore, è necessario avere a disposizione un *database* di partenza già strutturato.

23 Le Meur, T. (2010, p. 7). Il documento è l'esito della prima fase di lavoro del WP2 e descrive dettagliatamente le specifiche del modello XML di dati richiesti da MIMO per la descrizione degli strumenti musicali. È pubblicato tra i documenti da consultare per intraprendere un'adesione al progetto.

In fase di acquisizione e arricchimento dei record, MIMO controlla che le *MIMO specificities* siano soddisfatte e restituisce dei messaggi d'errore se, per esempio, specifici campi LIDO non corrispondono a un termine appartenente a una lista controllata. I file XML compilati secondo tali restrizioni sono comunque validi a fronte dell'XML Schema LIDO.

Se invece si estendesse il file XML a comprendere elementi non previsti da LIDO, il risultato non sarebbe più un record LIDO valido²⁴. In conclusione, avendo optato per l'adozione di LIDO, MIMO non ha tecnicamente avuto la possibilità di aggiungere nuovi campi allo schema XML e risolvere i *missing fields*.

Tra i campi mancanti segnaliamo l'assenza di riferimenti ad aspetti specifici degli strumenti musicali, come le caratteristiche acustiche e le tecniche esecutive.

Citiamo a questo proposito, dal sito web del CIMCIM, le considerazioni di Arnold Myers sui caratteri originali che la catalogazione degli strumenti musicali può avere:

A catalogue of museum instruments should relate the knowledge of the craftsmanship of instrument making to the knowledge of historical practice in music making. The description should not only include the physical description of objects, but also an indication of the resources provided by instruments to performers during their working timespan. It is the synthesis of these elements that makes the documentation requirements of musical instruments different from those of non-functioning museum objects.²⁵

Myers introduce, tra le informazioni utili alla catalogazione, alcuni aspetti relativi alla fattura degli strumenti e alla sua connessione con le tecniche costruttive ed esecutive; la rappresentazione di queste informazioni richiede la messa a punto di strumenti nuovi e più efficaci. La documentazione multimediale potrebbe rivelarsi uno strumento utile: andrebbero quindi elaborati nuovi modelli di descrizione degli strumenti musicali che, comprensivi di un uso documentale e descrittivo delle risorse multimediali, includano anche aspetti acustici e performativi.

Uno dei punti di forza di MIMO è aver puntato sull'aggregazione di istituzioni e competenze diverse, anche per l'istituzione di gruppi di lavoro dedicati alla revisione di temi disciplinari e all'individuazione di adeguate soluzioni tecnologiche. Una riflessione comune su un modello specifico per la descrizione degli strumenti musicali non è però stata preventivata: sin dall'inizio infatti il progetto prospettava l'adozione di un modello non specifico come *Dublin Core*.

24 Questo accade in sede di validazione di un file XML a fronte di un XML Schema (XSD): se il *tool* di validazione individua nel record XML dei *tag* non permessi dallo schema XSD, il file XML risulta *not valid*.

25 Myers e Karp (2005). L'argomento è sviluppato con maggiore dettaglio in Myers (1989).

L'adozione di LIDO da parte di MIMO è stata probabilmente dettata dalla convergenza di più fattori che hanno visto MIMO operare scelte di avanguardia nel campo dell'interoperabilità e dell'apertura (anche tramite l'adozione di tecnologie ancora in fase di sperimentazione).

La prima versione ufficiale LIDO 1.0 è stata pubblicata nel novembre 2010: MIMO ha adottato la versione LIDO *draft* 0.9 già nel giugno 2010 (Le Meur, 2010, p. 8)²⁶. LIDO si è proposto come il nuovo formato XML disegnato specificamente per la descrizione degli oggetti museali, compatibile con il CIDOC-CRM, promosso dall'ICOM come standard per l'*harvesting* tramite OAI-PMH; Europea ha subito tenuto conto di LIDO (e del suo modello evento-centrico) nello sviluppo di EDM, il nuovo modello di descrizione dei dati orientato al *semantic web*²⁷.

26 Successivamente alcuni *MIMO data provider* hanno aggiornato il modello alla versione 1.0.

27 MIMO è stato anche il primo aggregatore di dati ad implementare EDM (Bailly, *EDM Case Study*). Un elenco dei progetti che hanno adottato LIDO (molti dei quali espongono i propri dati in Europea) è consultabile a questo indirizzo: <http://network.icom.museum/cidoc/working-groups/lido/lido-community/use-of-lido/>

Unwrapping lido-v1.0.xsd

L'XML Schema di LIDO è consultabile a questo indirizzo:

<http://www.lido-schema.org/schema/v1.0/lido-v1.0.xsd>²⁸

Se si vuole comprendere la complessità della sua struttura, è necessario aprirlo con un editor XML, quindi collassare tutti i *tag*²⁹. Lo schema presenta una sequenza di definizioni di tipo e dichiarazioni di elemento e attributo (descrizioni a livello globale). La sequenza è in ordine alfabetico: lo schema non è comprensibile dall'essere umano nella sua globalità, ma solo aprendo dei percorsi a partire da un singolo punto.

Se apriamo, per esempio, l'elemento `resourceWrap` notiamo che:

- è seguito da una `annotation`, che fornisce indicazioni su come utilizzare l'elemento; il testo delle annotazioni, è il medesimo riportato anche nelle linee-guida pubblicate da MIMO (Le Meur, 2010);
- contiene una sequenza di descrizioni a livello locale (altre dichiarazioni di elemento e definizioni di tipo); esse possono risolversi nell'annidamento stesso oppure, come in questo caso, far riferimento a `resourceSetComplexType`, un tipo complesso definito a livello globale. Anche `resourceSetComplexType` è costituito da una sequenza di elementi, in una struttura ramificata complessa;
- le definizioni di tipo complesso sono spesso formate da una sequenza di elementi a cui sono associati gli attributi `minOccurs` e `maxOccurs`.

La struttura di LIDO è impostata in modo da consentire la massima riusabilità degli elementi più piccoli (le “foglie” dell’albero), che sono numericamente limitati rispetto alle dimensioni dell’intero schema. Gli elementi “foglia” sono descritti da tipi complessi con contenuto semplice e sono caratterizzati da:

- definizione del modello di contenuto (es. testo Unicode, numeri interi, formato di data ecc.)³⁰;
- lista degli attributi che l’elemento può o deve avere.

28 Il documento è distribuito con licenza Creative Commons “Attribution – Share Alike 3.0”. Le dichiarazioni di copyright sono incluse in un commento interno al file `.xsd`.

29 Sul web sono disponibili numerosi editor XML *open source* e gratuiti, sviluppati dalla comunità informatica.

30 Un elemento “foglia” potrà contenere esclusivamente il modello di contenuto descritto dallo schema XML.

È importante notare che, nell'ottica del *semantic web*, l'elemento "foglia" contiene il "dato atomizzato" a cui conferisce un primo barlume di semantizzazione. Nel caso di LIDO le "foglie" sono differenziate e definite in modo da essere compatibili con il contesto tecnologico ed ontologico di riferimento. Molte delle "foglie" più importanti sono predisposte per una mappatura in modelli ontologici RDF (*in primis* il CIDOC-CRM):

- `lido:appellationValue` contiene un nome; è alla base della concezioni dei "nomi" del CIDOC-CRM;
- `lido:conceptID` contiene un valore identificativo;
- `lido:earliestDate` e `lido:latestDate` indicano un *time-span* e si combinano in un elemento `lido:date`;
- `lido:term` contiene un termine appartenente una lista controllata;
- altri elementi (soprattutto "note descrittive") sono definiti da un `textComplexType` che contiene "testo libero".

Le "foglie" si possono combinare in tipi complessi più grandi, anch'essi concepiti in vista di un *mapping* semantico:

- `conceptComplexType` è l'identificatore di un concetto, che può essere espresso da un `lido:conceptID` e/o da un `lido:term`;
- `measurementSetComplexType` è un set che descrive una misurazione tramite la sequenza di elementi obbligatori `lido:measurementType`, `lido:measurementUnit` e `lido:measurementValue`.

LIDO definisce anche gli attributi (e le relative occorrenze) che le "foglie" possono avere; tra questi i più importanti in vista di un *mapping* semantico sono:

- `lido:pref` definisce il livello di preferenza attribuito all'elemento (es. "preferred" o "alternate");
- `lido:type` indica il "tipo" di elemento. È fondamentale per ricondurre i concetti al "tipo" di vocabolari a cui appartengono, per esempio nel caso delle classificazioni; è alla base della concezione dei "tipi" del CIDOC-CRM;
- `xml:lang`³¹ indica la lingua in cui è espresso l'elemento.

31 L'attributo `xml:lang` (che appartiene a un *namespace* diverso da `lido:`) può essere utilizzato in un record LIDO anche senza una specifica dichiarazione di *namespace* nella *root* `lido:lidoWrap`. `xml:lang` è infatti già dichiarato implicitamente nell'XML Schema `lido-v1.0.xsd` tramite l'elemento:

```
<xsd:import namespace="http://www.w3.org/XML/1998/namespace"
  schemaLocation="http://www.w3.org/2001/03/xml.xsd"/>
```

MIMO data model organization

Il modello di organizzazione dei dati usato da MIMO può essere rappresentato da una struttura ad albero di elementi annidati (cfr. fig.3)³². Ogni elemento è caratterizzato da specifici attributi di occorrenza: può essere obbligatorio o opzionale, ripetibile o non ripetibile.

Per formare un record MIMO valido è obbligatorio compilare almeno i seguenti elementi ed attributi (Bailly e Le Meur, 2010):

- `lido:lidoRecId` l'identificatore del record nella piattaforma MIMO;
- `xml:lang` (attributo di `lido:descriptiveMetadata`) il linguaggio usato per i `descriptive metadata`;
- `lido:objectWorkType` indica se si tratta di “strumento musicale” o “parti di uno strumento musicale”;
- `lido:classification` la *keyword* che designa in modo univoco lo strumento nella lista MIMO;
- `lido:titleSet` il nome dello strumento;
- `lido:workID` il numero di inventario dato allo strumento dall'istituzione di custodia;
- `lido:repositoryName` il nome dell'istituzione di custodia;
- `xml:lang` (attributo di `lido:administrativeMetadata`) il linguaggio usato per gli `administrative metadata`;
- `lido:recordID` l'identificatore del record nel sistema locale del *data provider*;
- `lido:recordType` il livello di collezione (obbligatorio per MIMO: “item”);
- `lido:recordSource` l'identificativo dell'istituzione che gestisce i metadati.

32 Quella che segue è una descrizione ragionata del modello di descrizione MIMO. Nella stesura si sono seguite alcune note redazionali: i termini in carattere “Courier” non preceduti da namespace (es. `descriptive metadata`) fanno riferimento ai generici campi MIMO presenti in fig.3; i termini in carattere “Courier” preceduti da namespace (es. `lido:descriptiveMetadata`) fanno riferimento a specifici elementi del file XML. Per una analisi dettagliata ed esaustiva delle specifiche tecniche si rimanda a Le Meur (2010).

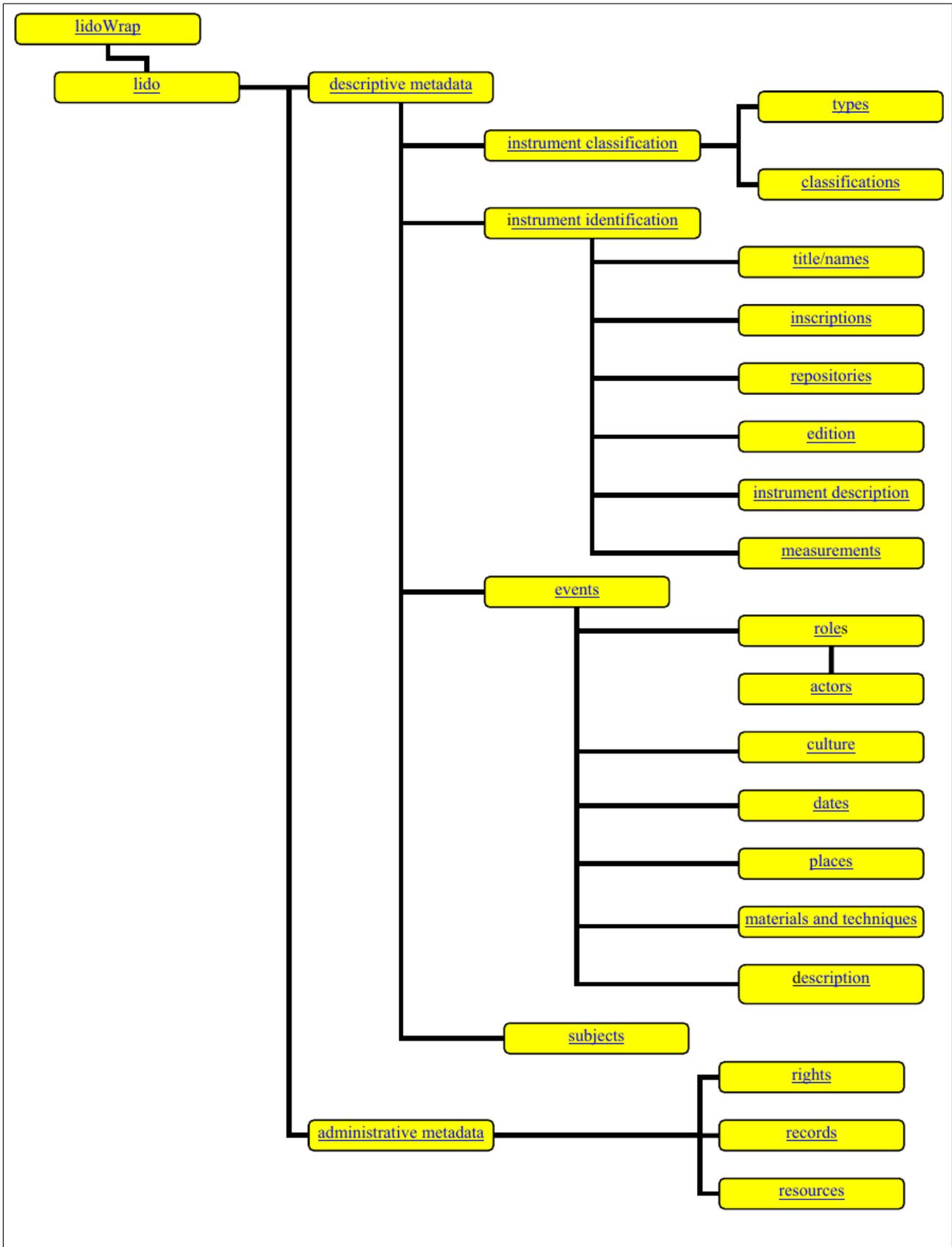


fig.3 MIMO data model organization – Le Meur (2010, p. 9)

Tutti gli altri elementi sono opzionali. Ogni elemento può essere formato da sotto-elementi che hanno a loro volta specifici attributi di occorrenza; nel caso di un sotto-elemento obbligatorio entro un elemento opzionale, il sotto-elemento *deve* essere compilato solo *se* è compilato l'elemento opzionale³³.

lidoWrap

Come tutti i documenti XML, LIDO ha un elemento *root* (`lido:lidoWrap`) che contiene le “dichiarazioni di namespace”.

La *root* contiene tutti i metadati di un'intera collezione, distribuiti in una sequenza di sotto-elementi ripetibili `lido:lido`. Questo vuol dire che `lido:lidoWrap` è il “pacchetto” che contiene infiniti elementi `lido:lido`.

lido

È l'elemento che include tutti i metadati di un singolo strumento, distribuiti nei seguenti sotto-elementi:

- `lidoRecID`;
- `descriptive metadata`;
- `administrative metadata`.

lidoRecID³⁴

È un elemento obbligatorio che identifica univocamente il record all'interno del database MIMO. L'identificativo si costruisce usando il seguente *pattern*:

```
<contributor identifier>":"<record identifier>
```

33 La tendenza a definire un numero ristretto di campi obbligatori è caratteristica dei formati e delle ontologie (come LIDO e CIDOC-CRM) che puntano all'interoperabilità e all'armonizzazione di dati provenienti da fonti diverse. Generalmente in questi casi gli elementi obbligatori servono a identificare il record, il *data provider*, il nome e il tipo di oggetto e il nome dell'istituzione di custodia; non sono definiti elementi obbligatori circa il modello di descrizione dell'oggetto. Questa non-prescrittività fa sì che differenti varietà di completezza di informazione siano consentite, e che i *data provider* possano decidere quali dati hanno intenzione di fornire a un portale pubblico online. Come abbiamo visto LIDO è concepito per la trasmissione di dati online; siccome si suppone che questi provengano da fonti eterogenee, è necessario che siano armonizzati in un modello descrittivo ampiamente inclusivo. Le istituzioni museali sono comunque incoraggiate a mantenere la specificità delle proprie descrizioni locali e possibilmente a tenere in conto una futura mappatura verso lo schema LIDO.

34 `lidoRecID` non compare nella struttura ad albero di fig.3; è un elemento “fratello” di `descriptive metadata` e `administrative metadata`.

Il <record identifier> è l'identificativo unico del record all'interno del sistema locale del *data provider*, e il <contributor identifier> è un prefisso che identifica univocamente il *data provider* all'interno di MIMO³⁵.

Esempio:

```
<lido:lidoRecID lido:type="local">
  CM:0162260
</lido:lidoRecID>
```

Il prefisso CM e la stringa numerica "0162260" identificano rispettivamente la "Cité de la Musique" e il record all'interno del database locale.

descriptive metadata

Contiene tutti i metadati che descrivono lo strumento, e si divide in:

- `instrument classification` include gli elementi obbligatori `lido:objectWorkType` e `lido:classification`;
- `instrument identification` include gli elementi obbligatori `lido:titleSet` `lido:workID` e `lido:repositoryName`;
- `events` opzionale;
- `subjects` opzionale.

administrative metadata

Contiene le informazioni "amministrative" relative ai metadati, al record e alle risorse digitali correlate. Si divide in:

- `rights` opzionale;
- `record` include gli elementi obbligatori `lido:recordID` `lido:recordType` e `lido:recordSource`;
- `resources` opzionale.

³⁵ Entrambi gli identificativi sono contenuti in due elementi obbligatori di MIMO (`lido:recordID` e `lido:recordSource`), all'interno degli `administrative metadata`. MIMO ha usato questi identificativi anche per costruire gli HTTP URI dei record: es. <http://www.mimo-db.eu/CM/0162260> .

xml:lang

Sia i *descriptive metadata* che gli *administrative metadata* hanno l'attributo `xml:lang` obbligatorio. Esso fa riferimento a una procedura standardizzata che permette alle macchine di identificare immediatamente il linguaggio in cui un set di metadati è scritto. Lo standard è quello definito dallo IETF³⁶, organizzazione che si occupa di sviluppare e promuovere standard Internet, in stretta cooperazione con il W3Ce ISO.

Uno dei propositi tecnologici più interessanti di MIMO, congiuntamente ad Europeana, è quello di fornire dati che siano recuperabili da motori di ricerca multilingue; affinché ciò sia possibile è necessario che tutti i dati siano riconducibili al linguaggio in cui sono espressi. Il valore attribuito a `xml:lang` è ereditato da tutti gli elementi discendenti di *descriptive metadata* e *administrative metadata*; ma `xml:lang` può essere impiegato a qualsiasi livello della struttura: MIMO incoraggia i *data provider* ad utilizzarlo soprattutto per quelle informazioni espresse in un linguaggio diverso da quello indicato nei livelli superiori. Siccome per molti campi LIDO consente di fornire diverse descrizioni parallele, l'attributo `xml:lang` è fondamentale per strutturare record bilingue³⁷ e multilingue.

instrument classification

Contiene alcuni elementi obbligatori relativi alla classificazione dello strumento.

Si divide nei campi `types` e `classifications`.

types

Contiene l'elemento obbligatorio e non ripetibile `lido:objectWorkType`³⁸.

Gli unici due valori ammessi da MIMO sono “musical instruments” e “parts of musical instruments”.

36 IETF (Internet Engineering Task Force) - <https://www.ietf.org/>

37 Il Museum of Musical Instruments di Bruxelles fornisce record bilingue (olandese e francese).

38 Questa è una *MIMO specificity*, perché in LIDO l'elemento è opzionale e ripetibile.

classifications

Contiene un elemento ripetibile `lido:classification` attraverso il quale si possono esprimere multiple classificazioni dello strumento.

Tra queste MIMO richiede che lo strumento sia obbligatoriamente classificato secondo il *thesaurus* di *keywords*. La *keyword* deve essere indicata entro un elemento `lido:term`; se sono presenti altre classificazioni espresse da `lido:term` (per esempio un *taxon* HS) la *keyword* deve essere contrassegnata da un attributo `lido:pref="preferred"`. Il termine inserito deve corrispondere esattamente a quello riportato nel vocabolario MIMO; nel caso in cui nessuna *keyword* MIMO sia funzionale alla classificazione dello strumento è possibile proporre una nuova.

La classificazione HS non è obbligatoria. Il *taxon* si indica con un `lido:term` con attributo `lido:pref="alternate"`; il numero Hornbostel-Sachs con `lido:conceptID` e attributo `lido:type="SH_Class"`. Anche in questo caso il codice deve corrispondere esattamente a quello del vocabolario HS³⁹. Esempio di classificazioni multiple tratto da Le Meur (2010, p. 18):

```
<lido:classification>
  <lido:term lido:pref="preferred">
    Accordion
  </lido:term>
</lido:classification>
<lido:classification>
  <lido:conceptID lido:type="SH_Class">
    412.13
  </lido:conceptID>
  <lido:term lido:pref="alternate">
    Free reeds aerophones
  </lido:term>
</lido:classification>
<lido:classification>
  <lido:term lido:pref="alternate">
    Wind
  </lido:term>
</lido:classification>
```

39 Deve esserci consistenza di nomenclatura per poter dare avvio al processo di *metadata enrichment*.

Instrument Identification

Contiene tutte le descrizioni che contribuiscono a identificare lo strumento. Con “identificazione” si intende sia l’ “individuazione” dello strumento in base a specifiche “coordinate” (titolo, iscrizioni, numero di inventario, numero di serie, istituzione di custodia) sia il “riconoscimento” in base alla descrizione dell'oggetto. Il campo si divide nei seguenti sotto-elementi:

- `titles/names` include l’elemento obbligatorio `lido:titleSet`;
- `inscriptions` opzionale;
- `repositories` include gli elementi obbligatori `lido:workID` e `lido:repositoryName`⁴⁰;
- `edition` opzionale;
- `instrument description` opzionale;
- `measurements` opzionale.

Il meccanismo Wrap/Set

Tutti i sotto-elementi di `Instrument Identification` sono a loro volta articolati in un unico elemento *Wrap* che contiene elementi *Set* ripetibili. Questa articolazione sfrutta le potenzialità di XML Schema nella definizione di strutture ad albero tramite annidamenti e attributi di occorrenza.

```
<lido:repositoryWrap>
  <lido:repositorySet lido:repositoryType="current">
    <lido:repositoryName>
      <...>
        Musée de la musique
      </...>
    </lido:repositoryName>
  </lido:repositorySet>
  <lido:repositorySet lido:repositoryType="former">
    <lido:repositoryName>
      <...>
        Collection Geneviève Thibaud de Chambure
      </...>
    </lido:repositoryName>
  </lido:repositorySet>
</lido:repositoryWrap>
```

40 Nell’XML Schema LIDO questi elementi non sono obbligatori.

In questo esempio⁴¹ un unico “pacchetto” `lido:repositoryWrap` contiene due `lido:repositorySet` che riportano descrizioni di repository diversi (uno “current”, l’altro “former”).

Lo stesso meccanismo viene applicato da LIDO in altre circostanze (anche se non sempre è usata la terminologia *Wrap* e *Set*), ad esempio:

- 1 `lido:titleWrap` può contenere 2 `lido:titleSet` con attributo `xml:lang` diverso (“it” e “en”) per descrivere i nomi italiano e inglese dello strumento;
- 1 `lido:objectDescriptionWrap` può contenere 2 `lido:objectDescriptionSet` con attributo `lido:type` diverso (“general description” e “decorative features”) per descrizioni di aspetti diversi dello strumento;
- 1 `lido:objectMeasurementsWrap` può contenere vari elementi `lido:objectMeasurementsSet` che misurano differenti parti dello strumento.

title/names

Contiene un elemento obbligatorio e ripetibile `lido:titleSet`; almeno un nome deve essere indicato obbligatoriamente. Nel caso in cui ci siano più nomi, almeno uno deve essere contrassegnato dall’attributo `lido:pref="preferred"`. Questo sarà il nome principale con cui lo strumento è rappresentato nel database MIMO. Le varianti linguistiche possono essere specificate dall’attributo `xml:lang`.

inscriptions

Contiene un elemento ripetibile e opzionale `lido:inscriptionSet` in cui inserire “trascrizione” e “descrizione” delle iscrizioni presenti sullo strumento.

repositories

Contiene un elemento ripetibile e obbligatorio `lido:repositorySet` che indica, tramite due sotto-elementi obbligatori, il nome dell’istituzione di custodia (`lido:repositoryName`) e il numero di inventario da essa attribuito allo strumento (`lido:workID`).

Altre informazioni opzionali relative a collocazioni precedenti dello strumento possono essere inserite specificando l’attributo `lido:type="former"`.

41 L’esempio è tratto da Le Meur (2010, p. 24) ed è stato volutamente abbreviato.

edition

Contiene un elemento ripetibile e opzionale `lido:displayEdition` che riporta informazioni che identificano un oggetto seriale (per esempio il numero di serie).

instrument description

Contiene un elemento opzionale e ripetibile `lido:objectDescriptionSet`.

Si possono diversificare le descrizioni, intese come brevi testi *essay-like*, settando l'attributo `lido:type` su un termine scelto all'interno di questa lista (proposta da MIMO):

- general description;
- context;
- decorative features;
- literature references;
- performance characteristics;
- inscription locations.

Se l'attributo non è specificato, "general description" è il `lido:type` di default.

measurements

Contiene un elemento opzionale e ripetibile `lido:objectMeasurementsSet`.

Esso si divide in un `lido:displayObjectMeasurements` e un `lido:objectMeasurements`.

`lido:objectMeasurements` è composto da una triade di sotto-elementi che indicano tipo, unità di misura e valore della misurazione, in una descrizione granulare⁴².

`lido:displayObjectMeasurements` fa riferimento alla stessa misurazione, ma la esprime sotto forma di testo libero⁴³.

Gli elementi *display*

Presentano le informazioni in una sintassi adatta e leggibile dall'utente; nella struttura XML sono "fratelli" di elementi che riportano le stesse informazioni mediante descrizioni più granulari.

LIDO può quindi fornire due tipi di informazione relativi allo stesso ambito descrittivo:

⁴² Per il significato del concetto di "granularità" cfr. *infra* il capitolo "Verbosità e granularità".

⁴³ Non c'è necessariamente una corrispondenza tra livello di precisione e di granularità. L'informazione relativa alle misurazioni di oggetti complessi non è sempre riconducibile con facilità a dati "atomici" gestibili digitalmente in maniera significativa. Talvolta un elemento *display*, fatto di testo libero, è uno strumento che rende conto in maniera più agile di misurazioni complesse.

- un'informazione atomizzata, ottimizzata per l'indicizzazione da parte di motori di ricerca e la mappatura in formati RDF;
- un'informazione *display*, ottimizzata per essere mostrata direttamente all'utente finale.

Nell'esempio seguente, tratto dal record RMAH-110249-NL del Musical Instrument Museum di Bruxelles, l'elemento `lido:displayObjectMeasurement` è costituito dalla concatenazione delle informazioni granulari contenute nella sequenza di elementi `lido:measurementsSet`.

```
<lido:objectMeasurementsWrap>
  <lido:objectMeasurementsSet>
    <lido:displayObjectMeasurements>
      Afmetingen H x B x D: 35,2 x 20,2 x 2,2 cm
    </lido:displayObjectMeasurements>
    <lido:objectMeasurements>
      <lido:measurementsSet>
        <lido:measurementType>
          Hoogte
        </lido:measurementType>
        <lido:measurementUnit>
          cm
        </lido:measurementUnit>
        <lido:measurementValue>
          35,2
        </lido:measurementValue>
      </lido:measurementsSet>
      <lido:measurementsSet>
        <lido:measurementType>
          Breedte
        </lido:measurementType>
        <lido:measurementUnit>
          cm
        </lido:measurementUnit>
        <lido:measurementValue>
          20,2
        </lido:measurementValue>
      </lido:measurementsSet>
      <lido:measurementsSet>
        <lido:measurementType>
          Diepte
        </lido:measurementType>
```

```

        <lido:measurementUnit>
            cm
        </lido:measurementUnit>
        <lido:measurementValue>
            2,2
        </lido:measurementValue>
    </lido:measurementsSet>
</lido:objectMeasurements>
</lido:objectMeasurementsSet>
</lido:objectMeasurementsWrap>

```

events

Tutte le informazioni non descritte nei campi precedenti devono essere espresse sotto forma di “eventi”. LIDO deriva l’impostazione evento-centrica dall’ontologia CIDOC-CRM (Coburn et al., 2010, p. 3).

Gli eventi sono strutturati secondo il meccanismo di relazione fra l’elemento `lido:eventWrap` e il sotto-elemento ripetibile `lido:eventSet`. Quest’ultimo consente di esporre le informazioni sia sotto forma di testo libero (usando il sotto-elemento `lido:displayEvent`) che come dati atomizzati, utilizzando il sotto-elemento `lido:event`.

`lido:event`, a sua volta, si articola in una serie di sotto-campi:

- `eventID`;
- `eventType` (obbligatorio);
- `eventName`;
- `eventActor`;
- `culture`;
- `eventDate`;
- `periodName`;
- `eventPlace`;
- `eventMaterialsTech`.

Ne risulta che tutti i metadati relativi a cultura di appartenenza, persone, date, luoghi e materiali connessi allo strumento devono essere descritti sotto forma di evento.

eventType

È l'unico sotto-elemento obbligatorio di `event` e specifica il tipo di evento descritto. LIDO raccomanda di indicare uno dei tipi di evento previsti dal CIDOC-CRM⁴⁴.

Tra questi MIMO elenca i più utili per la descrizione di strumenti musicali:

- acquisition;
- collecting;
- provenance;
- production;
- creation;
- use;
- restoration;
- exhibition.

eventName

Elemento opzionale e ripetibile che descrive il nome dell'evento.

eventActor

LIDO predispone una struttura molto articolata per i metadati relativi alle persone. Anche in questo caso c'è la possibilità di compilare un semplice campo *display*⁴⁵ oppure strutturare l'informazione in una forma più granulare usando l'elemento `lido:actorInRole`. Esso si divide in un elemento `lido:roleActor` che descrive il ruolo svolto dalla persona, e un elemento `lido:Actor` che "identifica" la persona mediante le informazioni fornite dai seguenti sotto-elementi:

- actorID;
- nameActor;
- nationalityActor;
- vitalDates;
- genderActor.

⁴⁴ Si intende che esso deve corrispondere a un termine appartenente alla lista di sotto-classi dell'entità CRM E5 Event.

⁴⁵ Testo libero in cui inserire in forma leggibile dall'utente tutte le informazioni relative alla persona.

Esempio di utilizzo di `lido:actorInRole` tratto dal record CM:0818487 della Cité de la Musique:

```
<lido:actorInRole>
  <lido:actor lido:type="person">
    <lido:nameActorSet>
      <lido:appellationValue xml:lang="fr">Sax, Adolphe</lido:appellationValue>
    </lido:nameActorSet>
    <lido:vitalDatesActor>
      <lido:earliestDate lido:type="exactDate">1814</lido:earliestDate>
      <lido:latestDate lido:type="exactDate">1894</lido:latestDate>
    </lido:vitalDatesActor>
  </lido:actor>
  <lido:roleActor>
    <lido:term xml:lang="en">instrument maker</lido:term>
  </lido:roleActor>
</lido:actorInRole>
```

culture

Descrive il contesto culturale relativo all'evento. MIMO ha in progetto la redazione di una lista controllata di etnonimi da raccomandare per la sua compilazione.

eventDate

Le date possono essere espresse con un elemento `lido:displayDate` o da una coppia di elementi relativi alle date di inizio e fine dell'evento. La concezione del tempo come "durata" espressa da *time-span* è funzionale alla compatibilità con il CIDOC-CRM.

period

Designa un periodo storico, così come è descritto nell'ontologia CIDOC-CRM.

eventPlace

I luoghi degli eventi possono essere espressi con elementi `lido:displayPlace` o con `lido:place`, in una descrizione più granulare.

`lido:place` può avere due attributi:

- `lido:politicalEntity` indica se si tratta di “country”, “region” o “town”;
- `lido:geographicalEntity` indica se si tratta di “continent” o “area”.

Esso contiene inoltre i seguenti sotto-elementi:

- `placeID` l'identificatore del luogo;
- `namePlaceSet` il nome del luogo;
- `partOfPlace` l'entità geografica più grande di cui `lido:place` è parte.

Settando accuratamente questi elementi è possibile interrogare il database di GeoNames e georeferenziare i propri luoghi. Per una corretta georeferenziazione MIMO (Bailly e Le Meur, 2010, p. 32) raccomanda di:

- specificare il tipo di luogo utilizzando gli attributi `lido:politicalEntity` o `lido:geographicalEntity`;
- specificare nel modo più preciso possibile la gerarchia di luoghi a cui un luogo specifico appartiene, utilizzando la combinazione di elementi `lido:partOfPlace`;
- specificare il linguaggio del nome del luogo utilizzando l'attributo `xml:lang`⁴⁶.

46 Alcuni record del database MIMO (tra cui quelli del Museu de la Musica di Barcelona) inseriscono direttamente l'URI del luogo pubblicato da Geonames nell'elemento `lido:placeID`:

```
<lido:placeID lido:type="geonames">http://www.geonames.org/2509951</lido:placeID>
```

MIMO non raccomanda questo metodo, che però a prima vista sembra più veloce e intuitivo; ad esempio: Riva presso Chieri: <http://www.geonames.org/3169256>; Goreme: <http://www.geonames.org/313947>; Ceriana: <http://www.geonames.org/6537035>.

Esempio di `lido:place` tratto dal record CM:0818487 della Cité de la Musique:

```
<lido:place lido:politicalEntity="town">
  <lido:namePlaceSet>
    <lido:appellationValue xml:lang="fr">Paris</lido:appellationValue>
  </lido:namePlaceSet>
  <lido:partOfPlace lido:politicalEntity="country">
    <lido:namePlaceSet>
      <lido:appellationValue xml:lang="fr">France</lido:appellationValue>
    </lido:namePlaceSet>
    <lido:partOfPlace lido:geographicalEntity="continent">
      <lido:namePlaceSet>
        <lido:appellationValue xml:lang="fr">Europe</lido:appellationValue>
      </lido:namePlaceSet>
    </lido:partOfPlace>
  </lido:partOfPlace>
</lido:place>
```

eventMaterialsTech

Descrive i materiali o la tecnica usata per costruire l'oggetto. Può essere espresso da un elemento `lido:displayMaterialsTech` o da un `lido:materialsTech`⁴⁷.

Una conseguenza importante dell'impostazione evento-centrica di LIDO è che si possono inserire informazioni sui materiali di cui uno strumento è fatto solo nell'ambito della descrizione di un evento (in questo caso, si tratterebbe di un evento di tipo "production" o "restoration").

subjects

È un elemento opzionale, che fa parte di una più ampia sezione di LIDO dedicata alle relazioni fra l'oggetto descritto e altri oggetti o argomenti. In questo senso le descrizioni dei "soggetti rappresentati" dall'oggetto rientrano nell'ambito delle relazioni. Questo campo, pensato soprattutto per le opere figurative, può essere utilizzato talvolta anche per gli strumenti musicali.

rights⁴⁸

Elemento opzionale che descrive i diritti relativi all'uso dei metadati e si divide nei campi:

⁴⁷ Se si sceglie di fornire una descrizione granulare e indicizzabile si può ricondurre il materiale o la tecnica a un identificativo riconosciuto univocamente.

⁴⁸ `rights`, `record` e `resources` sono sotto-campi di `administrative metadata` (cfr. *supra* p. 88).

- `rightsType` indica il copyright o altre attestazioni di proprietà intellettuale;
- `rightsDate` indica la data di validità dei diritti;
- `rightsHolder` indica l'istituzione o la persona detentrici dei diritti;
- `creditLine`.

records

Contiene una serie di tre elementi obbligatori per compilare un record LIDO valido:

- `lido:recordID` indica l'identificatore unico del record nel sistema (locale) del *data provider*;
- `lido:recordType` indica il livello della collezione: stabilisce se il record si riferisce a un unico strumento, una collezione o un gruppo di strumenti (il valore obbligatorio per MIMO è "item": ogni record deve riferirsi a un unico strumento);
- `lido:recordSource` identifica l'istituzione che fornisce e gestisce i metadati.

resources

Contiene le risorse digitali connesse al record. Il meccanismo di elementi *Wrap* e *Set* permette di descrivere molteplici risorse. MIMO e Europeana richiedono che una e una sola immagine sia contrassegnata come "principale". Gli elementi che costituiscono il campo `resource` sono:

- `resourceID` contiene l'identificatore unico della risorsa nel sistema "locale". Una e una sola immagine deve essere contrassegnata dall'attributo `lido:pref="preferred"`;
- `resourceRepresentation/linkResource` contiene l'URL della risorsa nel suo contesto originario (per esempio il sito del museo);
- `resourceType` indica il tipo di risorsa (MIMO raccomanda di inserire un termine appartenente alla lista di risorse elaborata da Europeana);
- `resourceDescription` contiene un testo libero di descrizione;
- `resourceDateTaken` indica la data di creazione della risorsa;
- `resourceSource` indica la fonte della risorsa;
- `rightResources` apre un elemento `rights` analogo a quello degli `administrative metadata`.

Gli ultimi due campi vanno compilati solo se sono diversi da quelli indicati generalmente per tutti gli `administrative metadata`.

Verbosità e granularità

A un primo approccio il modello di descrizione MIMO risulta molto complesso. Come tutti i modelli espressi in linguaggio XML infatti anche LIDO è “verboso”, cioè richiede l’utilizzo di molti *tag* e caratteri rispetto al contenuto di informazione da veicolare: questa inevitabile verbosità è connaturata al linguaggio XML, e consente di definire strutture tanto rigorose quanto flessibili e ricche di alternative⁴⁹.

La stessa impressione di eccessiva complessità si riscontra nella granularità⁵⁰ di LIDO, cioè nella sua tendenza a scomporre l’informazione in numerosi campi al fine di ottenere una atomizzazione dei dati.

In questo contesto accade che per descrivere la data “11 gennaio 1985” si debba suddividere l’informazione in questo modo:

```
<dateSet>
  <dateDisplay>
    11 gennaio 1985
  </dateDisplay>
  <date>
    <earliestDate>1985/01/11</earliestDate>
    <latestDate>1985/01/11</latestdate>
  </date>
</dateSet>
```

Solo il contenuto dell’elemento `date` può essere indicizzato da un motore di ricerca che gestisce le *date* come *time-span*. Più in generale, solo il tipo di informazione atomizzata è utile alle tecnologie del *semantic web*: i dati discreti possono essere identificati, ricondotti univocamente ai vocabolari, e mappati in modelli RDF come *Linked Open Data*.

49 L’XML Schema di LIDO predispone appositi nodi a sostegno dell’articolazione di elementi tra loro “discendenti” (*Wrap* e *Set*) o “fratelli” (elementi granulari e *display*): se da una parte tali nodi aumentano la verbosità, dall’altra consentono a LIDO di disporre di elementi molto utili.

50 Per una “mappatura” dei significati assunti dal termine nella letteratura scientifica contemporanea, si veda Zani (2006). Nell’attuale contesto di persistente indeterminatezza semantica, si è qui scelto di indicare con “granularità” non solo la “possibilità” che un’informazione sia suddivisibile in entità minute o atomiche (che abbia cioè una “grana”, similmente all’emulsione fotografica), ma anche che il processo di atomizzazione sia effettivamente in atto: una descrizione sarà “granulare” se tende ad essere maggiormente suddivisa in entità più piccole rispetto a edizioni passate o vicine. Allo stesso modo la locuzione “più granulare” intende sottolineare l’intensificazione di questa tendenza.

L'atomizzazione dei dati è uno dei compiti delle istituzioni della memoria nell'ambito del *semantic web*; ad esse infatti compete decidere di quali tipi di informazione è necessario o superfluo fornire una descrizione granulare⁵¹. La compilazione dei record XML è un lavoro impegnativo che deve essere alleggerito da processi automatici e dall'uso di interfacce per l'inserimento e la mappatura dei dati. Per programmare questi *tools* in modo efficace è fondamentale comprendere a fondo la struttura sulla quale essi devono intervenire.

Per rivedere alcune considerazioni fatte in precedenza, riportiamo un estratto dal record CM:0818487 della Cité de la Musique. Esso descrive la fabbricazione ("production") di uno strumento, nel 1848 a Parigi, ad opera di Adolphe Sax.

```
<lido:eventWrap>
  <lido:eventSet>
    <lido:event>
      <lido:eventType>
        <lido:term xml:lang="en">production</lido:term>
      </lido:eventType>
      <lido:eventActor>
        <lido:actorInRole>
          <lido:actor lido:type="person">
            <lido:nameActorSet>
              <lido:appellationValue xml:lang="fr">Sax, Adolphe</lido:appellationValue>
            </lido:nameActorSet>
            <lido:vitalDatesActor>
              <lido:earliestDate lido:type="exactDate">1814</lido:earliestDate>
              <lido:latestDate lido:type="exactDate">1894</lido:latestDate>
            </lido:vitalDatesActor>
          </lido:actor>
          <lido:roleActor>
            <lido:term xml:lang="en">instrument maker</lido:term>
          </lido:roleActor>
        </lido:actorInRole>
      </lido:eventActor>
      <lido:eventDate>
        <lido:displayDate xml:lang="fr">1848</lido:displayDate>
        <lido:date>
          <lido:earliestDate>1848</lido:earliestDate>
```

51 Questa decisione va presa anche tenendo conto che spesso le risorse disponibili per questi lavori sono limitate.

```

    <lido:latestDate>1848</lido:latestDate>
  </lido:date>
</lido:eventDate>
<lido:eventPlace>
  <lido:displayPlace xml:lang="fr">Paris / France / Europe</lido:displayPlace>
  <lido:place lido:politicalEntity="town">
    <lido:namePlaceSet>
      <lido:appellationValue xml:lang="fr">Paris</lido:appellationValue>
    </lido:namePlaceSet>
    <lido:partOfPlace lido:politicalEntity="country">
      <lido:namePlaceSet>
        <lido:appellationValue xml:lang="fr">France</lido:appellationValue>
      </lido:namePlaceSet>
      <lido:partOfPlace lido:geographicalEntity="continent">
        <lido:namePlaceSet>
          <lido:appellationValue xml:lang="fr">Europe</lido:appellationValue>
        </lido:namePlaceSet>
      </lido:partOfPlace>
    </lido:partOfPlace>
  </lido:place>
</lido:eventPlace>
</lido:event>
</lido:eventSet>
</lido:eventWrap>

```

Il nome del costruttore è inserito in un campo isolato, in modo che possa essere confrontato in modo univoco con il vocabolario di MIMO.

Il linguaggio delle informazioni è specificato ogni volta con l'attributo `xml:lang`.

Il record è caratterizzato dalla compresenza di descrizioni *display* e granulari.

La data di costruzione è espressa sia con un elemento *display* che in forma di *time-span*.

Il luogo è descritto parallelamente da un elemento *display* e da una gerarchia di `lido:partOfPlace`. L'informazione gerarchica garantisce che durante il processo di *metadata enrichment* si riesca a interrogare con sicurezza il database di GeoNames.

MIMO, Europeana e i LOD

Thesauri e classificazioni

I tre vocabolari redatti da MIMO in modalità LOD sono consultabili sulla pagina di *back-office* del database di MIMO⁵².

Il vocabolario di *Instrument Makers* è una lista di termini suddivisi in tre gruppi: *corporations*, *persons*, *families*; ogni voce della lista è caratterizzata da un nome principale, una lista di eventuali nomi alternativi e le date di nascita e morte (o inizio e fine di attività).

I vocabolari di sistematica HS e *keywords* sono organizzati gerarchicamente; la struttura ad albero è consultabile accedendo al *form* di *advanced search* del sito di *back-office*.

MIMO ha elaborato una propria versione della sistematica HS⁵³: essa a sua volta si basa sulla revisione pubblicata da Jeremy Montagu nel 2009, che è stata integrata con l'aggiunta della classe degli elettrofoni in base al modello elaborato da Marteen Quanten (Weisser e Quanten, 2011, pp. 133-142). La HS di MIMO è una struttura ad albero che, a partire da cinque classi principali, si suddivide in numerosi *taxa* rappresentati da numeri costruiti sulla base di un modello analogo a quello della Classificazione Decimale Dewey; i *taxa* sono pubblicati esclusivamente in inglese.

Il *thesaurus* di *keywords* è una struttura ad albero su tre livelli gerarchici, impostati su quello che è considerato il linguaggio d'uso più comune. La redazione delle *keywords* avviene tramite un processo collaborativo al quale tutti i *data provider* sono incoraggiati a partecipare. È possibile proporre modifiche alle parole esistenti e l'inserimento di parole nuove; di queste bisogna fornire una descrizione e indicare il livello gerarchico e la corrispondenza con un *taxon* HS. Le proposte vengono inviate a un gruppo di lavoro di MIMO che si occupa dell'aggiornamento del *thesaurus*. Inoltre MIMO richiede che i *data provider* si impegnino nella traduzione nella propria lingua-madre delle *keywords* esistenti.

Tutti e tre i vocabolari sono pubblicati in modalità LOD; questo vuol dire innanzitutto che ogni concetto (ogni “cellula” del vocabolario) è identificato tramite HTTP URI.

52 http://www.mimo-db.eu/mimo/infodoc/page-daccueil-infodoc.aspx?_lg=EN-en

Le scelte e le modalità di rappresentazione operate da MIMO nella redazione dei vocabolari sono discusse, da un punto di vista critico-organologico, nel capitolo “Il problema delle *keywords* e della sistematica Hornbostel-Sachs”. Qui ci preme illustrare tecnicamente il processo di *metadata enrichment*.

53 È pubblicata sul sito del CIMCIM: <http://network.icom.museum/cimcim/resources/classification-of-musical-instruments/>.

Ecco, dal database MIMO, tre esempi di HTTP URI di termini appartenenti ai vocabolari:

- <http://www.mimo-db.eu/InstrumentMaker/Person/2889>
- <http://www.mimo-db.eu/HornbostelAndSachs/321>
- <http://www.mimo-db.eu/InstrumentsKeywords/3783>⁵⁴

Anche ad ogni record è attribuito un HTTP URI, che ha questa forma:

- <http://www.mimo-db.eu/UEDIN/5710>

Una volta identificati tramite URI i concetti possono diventare il nodo di collegamenti semantici, espressi tramite il linguaggio SKOS. Nel caso particolare della sistematica HS e delle *keywords* i collegamenti possono essere all'interno del singolo vocabolario, tra i due vocabolari, o tra un vocabolario e un *dataset* esterno.

MIMO ha predisposto “Vizskos”, un *tool* di visualizzazione dei vocabolari strutturati in SKOS; è pubblicato nella sezione “Vocabulary” del sito di *back-office*⁵⁵. “Vizskos” consente di scegliere il tipo di visualizzazione⁵⁶ e il vocabolario da rappresentare. Selezionando i nodi della struttura si naviga all'interno della gerarchia; se invece si seleziona un termine si apre la scheda relativa al “concetto”.

L'esempio di fig. 4 è relativo alla *keyword* “Jew’s harp” e contiene i seguenti elementi SKOS:

- `skos:Concept` rappresenta il concetto relativo alla *keyword*, tramite l'URI che lo identifica univocamente;
- `skos:prefLabel` definisce il termine principale che identifica la *keyword* per ognuna delle otto lingue di MIMO, più la lingua *pivot*⁵⁷;
- `skos:altLabel` elenca le varianti linguistiche;
- `skos:broader` indica l'elemento “genitore” nella gerarchia;
- `skos:narrower` indica gli elementi “figli” nella gerarchia;
- `skos:exactMatch` definisce la corrispondenza esatta con un concetto del vocabolario HS;
- `skos:closeMatch` riporta il link alla voce “Jew’s harp” del *dataset* Dbpedia.

54 Notiamo in tutti gli esempi che l'ultima parte dell'URI, che differenzia un concetto dall'altro all'interno dello stesso vocabolario, è costituita da un numero primo, e non incorpora una rappresentazione del concetto a cui si riferisce (ad es. il numero HS o la parola “zampogna”).

55 “Vizskos” non è di semplice usabilità: è lento e ha problemi di visualizzazione, soprattutto riguardo alla dimensione dei caratteri; è stato pubblicato in via sperimentale.

56 La visualizzazione ad albero è più *user-friendly*.

57 La lingua *pivot* è «the most commonly used term in Europe to describe a class of instruments. This [...] form was needed because its language is not necessarily one of the MIMO's [...] languages» (Bailly, 2011, p. 7) .

ConceptScheme : <http://www.mimo-db.eu/InstrumentsKeywords>

skos:Concept
Guimbardes
<http://www.mimo-db.eu/InstrumentsKeywords/2861>

skos:prefLabel

- Jew's harps (pivot)
- Drumle (pl)
- Guimbardes (ca)
- Maultrommeln (de)
- Trumps (en)
- Guimbardes (fr)
- Scacciapensieri (it)
- Mondharpen (nl)
- Mungigor (sv)

skos:altLabel

- Trumps (pivot)
- Jew's harps (en)

skos:broader

- [Instruments à percussion](#)

skos:narrower

- [Murcang](#)
- [Guimbarde](#)
- [Kubing](#)
- [Dan moi](#)
- [Genggong](#)
- [Tong](#)
- [Tombagl](#)
- [Komuz](#)
- [Susap](#)
- [Mukkuri](#)
- [Afiw](#)
- [Kinaban](#)
- [Kuding](#)
- [Karinding](#)
- [Bikong](#)
- [Cheng](#)
- [Kovízh](#)
- [Marranzanu](#)
- [Drimba](#)
- [Drymba](#)

skos:exactMatch

- <http://www.mimo-db.eu/HornbostelAndSachs/60>

skos:closeMatch

- http://dbpedia.org/resource/Jew%27s_harp

< >

fig.4 Rappresentazione grafica dello schema concettuale della keyword “Jew’s Harp”

<http://www.mimo-db.eu/InstrumentsKeywords/2861>

La scheda del *taxon* di fig.5 è per certi aspetti simile a quella della *keyword*. Il concetto, descritto da un testo libero contenuto nell'elemento `skos:definition`, è collocato nella tassonomia da `skos:broader` e `skos:narrower`. L'elemento `skos:exactMatch` indica le *keywords* a cui il concetto del *taxon* HS corrisponde.

ConceptScheme : <http://www.mimo-db.eu/HornbostelAndSachs>

skos:Concept
311.121.22 Mono-heterochord musical bows with resonator attached < >
<http://www.mimo-db.eu/HornbostelAndSachs/154>

skos:definition
 A single string is stretched between fixed points. The instrument consists of a string bearer with a resonator that can be detached without destroying the sound-producing apparatus. The string bearer is flexible (and curved). The string is of separate material from the string bearer.

skos:prefLabel

- 311.121.22 Mono-heterochord musical bows with resonator attached (pivot)

skos:broader

- [Simbing](#)
- [311.121.2 Mono-heterochord musical bows with resonator](#)

skos:narrower

- [311.121.221 Mono-heterochord musical bows with resonator attached, without tuning noose](#)
- [311.121.222 Mono-heterochord musical bows with resonator attached with tuning noose](#)

d+ Sets of free reeds, with flexible air reservoir, with keyboard

skos:exactMatch

- <http://www.mimo-db.eu/InstrumentsKeywords/3551>
- <http://www.mimo-db.eu/InstrumentsKeywords/3528>

fig.5 Rappresentazione grafica dello schema concettuale del taxon 311.121.22

<http://www.mimo-db.eu/HornbostelAndSachs/154>

Il metadata enrichment

Quando avviene l'*harvesting* dei metadati da parte di MIMO, si attiva un importante processo di *metadata enrichment*. Il sistema cerca di stabilire la corrispondenza tra specifici campi LIDO e i vocabolari in modalità LOD. Interroga, per esempio, il database di Geonames per recuperare l'URI dei luoghi; verifica l'eventuale corrispondenza tra il nome del costruttore della sezione *events* e la lista di *Instrument Makers*, e tra un termine del campo *classification* e la tassonomia di *keywords*.

Nel caso degli *Instrument Makers* il record è associato in modo univoco all'URI che identifica la persona nella lista controllata; tutti i record che condividono questa associazione hanno una proprietà in comune univocamente riconosciuta. Alcuni esempi dei risultati raggiunti grazie all'applicazione di registri di autorità su aggregazioni di dati sono descritti nel *Final Report*⁵⁸.

Nel caso dei vocabolari di tipo gerarchico e associativo (*keywords* e HS) il record è connesso univocamente non solo all'URI del concetto, ma anche a tutti gli altri URI indicati dagli elementi SKOS, in un reticolo di relazioni semantiche “a cascata”.

Molte di queste relazioni supportano il multilinguismo; altre puntano a *dataset* esterni. Se si consulta la voce http://dbpedia.org/resource/Jew%27s_harp a cui punta la *keyword* di fig.4, si realizza l'imponenza delle dimensioni del fenomeno LOD.

Quando Europeana raccoglie i dati dal database MIMO avviene un ulteriore processo.

MIMO ha fornito anticipatamente a Europeana i propri vocabolari in modalità LOD (RDF/XML).

Europeana mappa i record da LIDO a EDM; in quanto orientato al *semantic web*, EDM è

58 «This section of the report gives some examples of how MIMO has opened up access to information through aggregation and through other features including controlled vocabularies and classification. The examples include searches such as those made by the general public, information gathering in learning and teaching, and by researchers.

Many kinds of information search are available to all through MIMO which could previously be attempted only by specialists by use of printed catalogues (generally out of date) and by visits to, or correspondence with, museums (expensive and time-consuming). [...]

Theobald Boehm (1794-18871) not only developed the modern flute but his ideas were applied to produce the modern clarinet and with less success to the oboe and bassoon. Searching in MIMO for 'Boehm' gives 181 instruments of different kinds in MIMO museums, providing material for a study of his influence on woodwinds of all kinds. This is a typical topic for student projects. [...]

The Denner family, active in Nuremburg at the end of the 17th and the first half of the 18th century, were renowned for the quality of their instruments and for their development of a range of woodwind instruments. They invented the clarinet. Through MIMO the 54 Denner instruments in 6 museums can be compared with instruments by other contemporary makers, and an approach can be made to the difficult question of which Denner family member made which instruments.

The Brussels Musical Instrument Museum is planning a major exhibition on Adolphe Sax (1814-1894) to be held in 2014; the greatest part of Sax's surviving instruments (254 instruments in 6 museums) can be viewed through MIMO and compared, thus greatly facilitating exhibition planning» (Rodger, 2011, pp. 35-36).

un'implementazione di RDF: descrive i dati sotto forma di triple RDF fatte di classi e proprietà semanticamente connotate. Solo a questo punto i dati "atomizzati" dei record LIDO e dei vocabolari MIMO sono implementati in RDF, collegati da relazioni semantiche in un'ontologia evento-centrica, e pubblicati come *Linked Open Data*.

Il grafo di fig.6 rappresenta un prototipo di mappatura LIDO/EDM.

CIDOC-CRM: un'ontologia *event-based*

Il Conceptual Reference Model

Il CIDOC-CRM⁵⁹ è stato sviluppato dal CIDOC Documentation Standards Working Group e dal CIDOC-CRM Special Interest Group, due *Working Group* del CIDOC, il Comitato Internazionale per la Documentazione dell'ICOM. LIDO è anch'esso un'emanazione del CIDOC, e può essere considerato un'applicazione del CIDOC-CRM. L'impostazione evento-centrica e la concezione di terminologie (`lido:appellationValue`) e classificazioni (`lido:type`) che caratterizzano LIDO corrispondono infatti ad alcune delle strutture portanti del Conceptual Reference Model (CRM).

Il CRM è un'ontologia, una forma di rappresentazione della conoscenza relativa a uno specifico dominio, in questo caso l'intero dominio del patrimonio culturale. Lo scopo di un'ontologia di dominio è descrivere un intero ambito della conoscenza e quindi facilitare l'integrazione, la mediazione e l'interscambio di informazioni provenienti da fonti eterogenee; in altre parole un'ontologia fornisce una struttura semantica modellata sulle costanti che caratterizzano il dominio rappresentato.

Il CRM non prescrive standard di documentazione: è la ricerca che si occupa di definire le modalità di inserimento dei dati (che cosa inserire, in quale ordine, quali sono i campi obbligatori e quali sono ripetibili ecc.) e di elaborare le terminologie e le classificazioni, dando origine ad “ambienti” semanticamente ricchi e logicamente formati. Il CRM si occupa di fornire quella interoperabilità semantica necessaria affinché tali fonti possano costituirsi in una risorsa globale coerente.

L'ambito della teoria (*Intended Scope*) del CRM include «all information required for the exchange and integration of heterogeneous scientific documentation of museum collections»⁶⁰.

L'insieme dei documenti che entra nel campo d'interesse del CRM è decisamente ampio e diversificato; per questo motivo il CRM può essere considerato più vicino all'idea di “ontologia fondazionale” che a quella di “ontologia di dominio”. La nozione di «museum collections» definita dall'ICOM comprende «collections, sites and monuments relating to fields such as social history,

59 CIDOC-CRM (Conceptual Reference Model) - <http://www.cidoc-crm.org/index.html>

60 Crofts, Doerr, Gill, Stead e Stiff (2011, p. ii). Il documento presenta l'ultima versione ufficiale del CRM, pubblicata nel novembre 2011, e divenuta standard ISO nel 2014 (ISO 21127:2014). La maggior parte delle informazioni sul CRM presenti in questo capitolo sono tratte dal documento citato.

Per una breve introduzione al CRM rivolta a manager, professionisti e ricercatori, si rimanda a Oldman, C. e CRM Labs (2014).

ethnography, archaeology, fine and applied arts, natural history, history of sciences and technology» (Crofts et al., 2011, p. ii).

Con il termine «scientific documentation» non si intende un sapere di ambito accademico; piuttosto si presuppone che la documentazione prodotta dai musei risponda a requisiti di qualità e profondità descrittiva dell'informazione trasmessa sufficiente affinché il CRM possa produrre aggregazioni di dati utili *ai fini* della ricerca scientifica. In questo senso rientrano nell'ambito del CRM numerosi tipi di documenti: gli accordi e i contratti provenienti dagli archivi, le opere letterarie, la documentazione relativa ad amministrazione e conservazione degli oggetti, i cataloghi delle esposizioni, il materiale divulgativo ed educativo, gli articoli e le classificazioni della ricerca accademica.

Ciascuno di questi tipi di documenti è inoltre espressione di una particolare comunità, per cui assume forme diverse a seconda degli orientamenti culturali; ad esempio, nel caso della ricerca scientifica, possono coesistere diversi tipi di classificazione degli stessi oggetti. L'intenzione del CRM non è solo preservare tale eterogeneità, ma anche contestualizzarla: se in un dominio possono convivere due classificazioni differenti e parallele, è anche possibile ricostruirne il processo storico-culturale di definizione e risalire agli autori. Lo stesso principio è applicato nel caso di fonti discordanti, come può accadere nel caso di documenti d'archivio che restituiscono interpretazioni diverse di un certo evento.

I documenti si differenziano inoltre per il livello di conoscenza che rappresentano: ci possono essere informazioni più o meno dettagliate, descrizioni più o meno granulari, aspetti della conoscenza che assumono un ruolo di rilievo o, viceversa, sono ignorati nelle rappresentazioni. Tale variabilità non dipende solo dal grado di accuratezza ma anche dall'orientamento e dall'ambito della ricerca: un biologo o uno storico dell'arte adotteranno strategie diverse nella descrizione del materiale (per esempio “legno”) che costituisce un oggetto; allo stesso modo, la relazione tra oggetto e tempo assumerà significati differenti a seconda che si tratti di un reperto archeologico o di un documento d'archivio digitale.

Il CRM è predisposto per essere molto flessibile e consentire a molteplici sistemi differenti di essere CRM-compatibili, invece che imporre una particolare soluzione. Non è richiesta una corrispondenza completa tra tutte le strutture del sistema CRM-compatibile e quelle del CRM, né è necessario che tutti i concetti e le associazioni del CRM siano sempre implementati. Al contrario, il modello è estensibile (si può ampliare senza che ne sia compromessa la validità) al fine di coprire

nuovi ambiti non descritti in origine⁶¹, o può essere semplificato (se ne può usare solo un *subset*). L'assenza di prescrizioni di campi e valori, spesso interpretata come carenza di determinismo e standardizzazione, è invece un presupposto per una struttura che fornisce interoperabilità semantica per l'integrazione dei dati.

Il CRM è un'ontologia nel senso usato dalle scienze informatiche. È stato espresso tramite un modello semantico *object-oriented*, nella convinzione che questa formulazione sia comprensibile tanto dagli esperti della documentazione quanto dagli informatici, e che sia allo stesso tempo facilmente traducibile in formati *machine-readable*. Il CRM è indipendente da particolari tecnologie (anche se viene spesso implementato nei formati connessi ai linked data, RDF innanzitutto) e può essere espresso nei linguaggi ontologici e nei formati di serializzazione più comuni (RDF Schema, OWL, XML ecc.).

Lo scopo ultimo dell'ontologia è supportare il *computer-based reasoning*, una delle caratteristiche principali del *semantic web*: il CRM è fondato sul concetto di “classi” *object-oriented*, le cui “relazioni” sono definite seguendo regole logiche. Il rispetto di tali regole, messe a punto empiricamente secondo un processo *bottom up*, dovrebbe assicurare la definizione di una struttura semantica coerente di dati; il computer avrebbe così a disposizione un “ambiente” ricco semanticamente, condizione in cui può esso stesso inferire automaticamente proposizioni logiche corrette. Tali proposizioni, che cercano di sfruttare al meglio le possibilità del calcolo computazionale, non intendono, secondo il CIDOC, sostituire la ricerca accademica, ma piuttosto servire come strumento di individuazione di fatti e connessioni tra fatti, lasciandone ovviamente l'interpretazione agli studiosi.

61 Tra le estensioni del CRM, consultabili a questo URL <http://83.212.168.219/CIDOC-CRM/use-cases> (nuovo website del CIDOC-CRM, ancora in fase di test), un cenno particolare meritano le FRBRoo (http://www.cidoc-crm.org/frbr_inro.html) ontologia che rappresenta il contenuto formale dell'informazione bibliografica secondo il paradigma *object-oriented*. Il risultato è il frutto della collaborazione tra IFLA (International Federation of Library Associations and Institutions) e CIDOC (International Working Group on FRBR/CIDOC CRM Harmonisation) ed è volto a migliorare l'integrazione tra l'informazione bibliografica e quella museale.

E2	-	Temporal Entity
E4	- -	Period
E5	- - -	Event
E7	- - - -	Activity
E11	- - - - -	Modification
E12	- - - - -	Production
E13	- - - - -	Attribute Assignment
E65	- - - - -	Creation
E63	- - - - -	Beginning of Existence
<i>E12</i>	- - - - -	<i>Production</i>
E65	- - - - -	Creation
E64	- - - - -	End of Existence
E77	-	Persistent Item
E70	- -	Thing
E72	- - -	Legal Object
E18	- - - -	Physical Thing
E24	- - - - -	Physical Man-Made Thing
E90	- - - - -	Symbolic Object
E71	- - -	Man-Made Thing
<i>E24</i>	- - - -	<i>Physical Man-Made Thing</i>
E28	- - - -	Conceptual Object
E89	- - - - -	Propositional Object
E30	- - - - -	Right
E73	- - - - -	Information Object
<i>E90</i>	- - - - -	<i>Symbolic Object</i>
E41	- - - - -	Appellation
<i>E73</i>	- - - - -	<i>Information Object</i>
E55	- - - - -	Type
E39	- -	Actor
E74	- - -	Group
E52	-	Time-Span
E53	-	Place
E54	-	Dimension
E59		Primitive Value
E61	-	Time Primitive
E62	-	String

fig.6 Reduced CRM-compatible form: subset minimo di classi

Crofts et al. (2010, pp. iv-v)

La rappresentazione tramite indentazione di una struttura poligerarchica costringe alla ripetizione di alcune classi: ad esempio E12 Production è sia una E7 Activity che un E63 Beginning of Existence, per cui è ripetuta due volte (la seconda volta è in corsivo).

Property id	Property Name	Entity – Domain	Entity - Range
P1	is identified by (identifies)	E1 CRM Entity	E41 Appellation
P2	has type (is type of)	E1 CRM Entity	E55 Type
P3	has note	E1 CRM Entity	E62 String
P4	has time-span (is time-span of)	E2 Temporal Entity	E52 Time-Span
P7	took place at (witnessed)	E4 Period	E53 Place
P10	falls within (contains)	E4 Period	E4 Period
P12	occurred in the presence of (was present at)	E5 Event	E77 Persistent Item
P11	- had participant (participated in)	E5 Event	E39 Actor
P14	- - carried out by (performed)	E7 Activity	E39 Actor
P16	- used specific object (was used for)	E7 Activity	E70 Thing
P31	- has modified (was modified by)	E11 Modification	E24 Physical Man-Made Thing
P108	- - has produced (was produced by)	E12 Production	E24 Physical Man-Made Thing
P92	- brought into existence (was brought into existence by)	E63 Beginning of Existence	E77 Persistent Item
P108	- - <i>has produced (was produced by)</i>	<i>E12 Production</i>	<i>E24 Physical Man-Made Thing</i>
P94	- - has created (was created by)	E65 Creation	E28 Conceptual Object
P93	- took out of existence (was taken out of existence by)	E64 End of Existence	E77 Persistent Item
P15	was influenced by (influenced)	E7 Activity	E1 CRM Entity
P16	- <i>used specific object (was used for)</i>	<i>E7 Activity</i>	<i>E70 Thing</i>
P20	had specific purpose (was purpose of)	E7 Activity	E7 Activity
P43	has dimension (is dimension of)	E70 Thing	E54 Dimension
P46	is composed of (forms part of)	E18 Physical Thing	E18 Physical Thing
P59	has section (is located on or within)	E18 Physical Thing	E53 Place
P67	refers to (is referred to by)	E89 Propositional Object	E1 CRM Entity
P75	possesses (is possessed by)	E39 Actor	E30 Right
P81	ongoing throughout	E52 Time-Span	E61 Time Primitive
P82	at some time within	E52 Time-Span	E61 Time Primitive
P89	falls within (contains)	E53 Place	E53 Place
P104	is subject to (applies to)	E72 Legal Object	E30 Right
P106	is composed of (forms part of)	E90 Symbolic Object	E90 Symbolic Object
P107	has current or former member (is current or former member of)	E74 Group	E39 Actor
P127	has broader term (has narrower term)	E55 Type	E55 Type
P128	carries (is carried by)	E24 Physical Man-Made Thing	E73 Information Object
P130	shows features of (features are also found on)	E70 Thing	E70 Thing
P140	assigned attribute to (was attributed by)	E13 Attribute Assignment	E1 CRM Entity
P141	assigned (was assigned by)	E13 Attribute Assignment	E1 CRM Entity
P148	has component (is component of)	E89 Propositional Object	E89 Propositional Object

fig.7 Reduced CRM-compatible form: subset minimo di proprietà

Crofts et al. (2010, pp. iv-v)

Il modello *object-oriented*

Il CIDOC ha mantenuto fino al 1998 un tradizionale modello entità-relazione, con lo scopo di fornire una struttura di armonizzazione dei dati. Alcuni importanti difetti strutturali hanno però costretto al suo abbandono: il modello relazionale si è rivelato carente in flessibilità e pregnanza semantica, ed è stato sottoposto a continue espansioni fino a raggiungere dimensioni e complessità troppo elevate da gestire. A partire dal 1996 il CIDOC ha iniziato a sviluppare il CRM, che a differenza del modello precedente è un'ontologia evento-centrica e adotta un paradigma di modellazione *object-oriented* (quindi non relazionale).

Introdotta dai linguaggi di programmazione OOP (Object Oriented Programming), il paradigma *object-oriented* è un insieme di concetti che viene sempre più spesso implementato in altri contesti della *Information Technology*⁶². Il CRM è definito in base ad una serie di concetti tipici del modello *object-oriented*⁶³: ne deriva una struttura poligerarchica di 90 classi e 149 proprietà⁶⁴.

Una classe è una categoria di oggetti che condividono uno o più tratti in comune; i criteri per identificare gli elementi appartenenti alla classe sono espressi in un testo (*scope note*), e fanno riferimento a concettualizzazioni diffuse nel campo degli esperti del dominio. La somma dei tratti comuni che identificano una classe è detta intensione. Un elemento della vita reale che appartiene a una classe è chiamato istanza di quella classe, ed una classe è associata ad un set aperto di istanze della vita reale, detto estensione; qui il termine “aperto” indica che la conoscenza di tutte le istanze reali di una classe va generalmente oltre le nostre capacità, e lascia spazio a possibili integrazioni ed espansioni future (*Open World Assumption*).

Le classi sono identificate da numeri preceduti dalla lettera E (che sta per “Entity”, altro termine con cui il CRM indica le classi), e seguiti da gruppi nominali con iniziale maiuscola.

62 Nell'ambito dei Database Management Systems, a fianco dei tradizionali modelli relazionali si stanno sviluppando sistemi *object-oriented* e sistemi intermedi (OODBMS – Object Oriented Database Management System e ORDBMS – Object Relational Database Management System).

63 Ad es: classi e proprietà, oggetti come istanze di una classe, particolari meccanismi di relazione all'interno dei livelli gerarchici, come l'ereditarietà multipla ecc.

64 Il dato si riferisce all'ultima versione ufficiale, la 5.0.4 del 2011, divenuta standard ISO (ISO 21127:2014) nel 2014. La precedente versione, la prima approvata come standard ISO nel 2006 (ISO 21127:2006) era costituita da 86 classi e 137 proprietà: si rileva quindi un incremento della struttura tra le due versioni. L'ultima bozza del 2015 invece (versione 6.2.1) evidenzia un ulteriore aumento delle proprietà (168), mentre le classi restano contenute (94).

E2 Temporal Entity	
Subclass of:	E1 CRM Entity
Superclass of:	E3 Condition State E4 Period
Scope note:	This class comprises all phenomena, such as the instances of E4 Periods, E5 Events and states, which happen over a limited extent in time. In some contexts, these are also called perdurants. This class is disjoint from E77 Persistent Item. This is an abstract class and has no direct instances. E2 Temporal Entity is specialized into E4 Period, which applies to a particular geographic area (defined with a greater or lesser degree of precision), and E3 Condition State, which applies to instances of E18 Physical Thing.
Examples:	<ul style="list-style-type: none"> ▪ Bronze Age (E4) ▪ the earthquake in Lisbon 1755 (E5) ▪ the Peterhof Palace near Saint Petersburg being in ruins from 1944 – 1946 (E3)
Properties:	<p>P4 has time-span (is time-span of): E52 Time-Span</p> <p>P114 is equal in time to: E2 Temporal Entity</p> <p>P115 finishes (is finished by): E2 Temporal Entity</p> <p>P116 starts (is started by): E2 Temporal Entity</p> <p>P117 occurs during (includes): E2 Temporal Entity</p> <p>P118 overlaps in time with (is overlapped in time by): E2 Temporal Entity</p> <p>P119 meets in time with (is met in time by): E2 Temporal Entity</p> <p>P120 occurs before (occurs after): E2 Temporal Entity</p>

fig.8 E2 Temporal Entity Declaration

Crofts et al. (2010, p. 2)

La struttura poligerarchica del CRM si distingue da una tradizionale gerarchia ad albero per il fatto che ogni elemento può avere più di un elemento “genitore” (ogni classe può avere più di una e una sola superclasse). Questa caratteristica rende la strutturazione delle classi più ricca semanticamente e contestualmente per diversi motivi; la definizione dell’intensione di una classe, per esempio, non è limitata ad un percorso unidirezionale ma può attingere a molteplici intensioni poste su diversi livelli; si può usare la funzione dell’ereditarietà multipla delle proprietà (cfr. fig.9).

Una sottoclasse è una classe che è la specializzazione di un'altra classe. Il vincolo di specializzazione stabilisce una *IsA Relationship*, secondo la quale:

- tutte le istanze della sottoclasse sono anche istanze della sua superclasse;
- l'intensione della sottoclasse estende l'intensione della superclasse (cioè i suoi tratti sono più restrittivi di quelli della superclasse);
- oltre ad avere le sue specifiche proprietà, la sottoclasse eredita la definizione di tutte le proprietà dichiarate per la sua superclasse (ereditarietà stretta) e di conseguenza, di tutte le proprietà degli “antenati”.

Una sottoclasse che ha più di una immediata superclasse, eredita le proprietà di tutte le sue superclassi (ereditarietà multipla).

Una superclasse è una classe che è la generalizzazione di una o più classi (le sue sottoclassi), cioè sussume tutte le istanze delle sue sottoclassi (inoltre può avere istanze aggiuntive che non appartengono ad alcuna delle sue sottoclassi).

E21 Person	
Subclass of:	E20 Biological Object E39 Actor
Scope note:	This class comprises real persons who live or are assumed to have lived. Legendary figures that may have existed, such as Ulysses and King Arthur, fall into this class if the documentation refers to them as historical figures. In cases where doubt exists as to whether several persons are in fact identical, multiple instances can be created and linked to indicate their relationship. The CRM does not propose a specific form to support reasoning about possible identity.
Examples:	<ul style="list-style-type: none"> ▪ Tut-Ankh-Amun ▪ Nelson Mandela

fig.10 E21 Person Declaration

Crofts et al. (2010, p. 11)

Nel caso di E21 Person, la classe eredita tutte le proprietà delle sue due superclassi E20 Biological Object e E39 Actor, che a loro volta ereditano le proprietà delle rispettive superclassi e così via fino a E1 CRM Entity, vertice della gerarchia e ultimo antenato di E21 Person.

Una classe infine può essere il dominio o il *range* di nessuna, una o più proprietà.

Una proprietà serve a definire la relazione che intercorre tra due classi. Anche la proprietà, come la classe, è caratterizzata da un'intensione, descritta nella *scope note*. La proprietà svolge un ruolo sintattico analogo a quello del predicato verbale, e va definita in riferimento al proprio dominio e *range*, che svolgono rispettivamente la funzione di soggetto e complemento⁶⁵. Come accade per i verbi transitivi, essa può essere interpretata in entrambe le direzioni (dal dominio al *range* e viceversa), ed è perciò descritta da due distinte definizioni a seconda del verso di interpretazione. Un'istanza di una proprietà è la relazione fattuale che c'è tra l'istanza di un dominio e quella di un *range*.

Anche le proprietà sono articolate in una poligerarchia di superproprietà e sottoproprietà; il rapporto tra intensioni, istanze, domini e *range* di tali proprietà è definito dalla *IsA Relationship*.

Le proprietà sono identificate dalla lettera P seguita da un numero, un verbo (che esprime il senso della proprietà dal dominio al *range*) e un ulteriore verbo tra parentesi (che esprime il senso dal *range* al dominio).

P43 has dimension (is dimension of)	
Domain:	E70 Thing
Range:	E54 Dimension
Quantification:	one to many, dependent (0,n:1.1)
Scope note:	This property records a E54 Dimension of some E70 Thing. It is a shortcut of the more fully developed path from E70 Thing through <i>P39 measured (was measured by)</i> , E16 Measurement <i>P40 observed dimension (was observed in)</i> to E54 Dimension. It offers no information about how and when an E54 Dimension was established, nor by whom.
Examples:	An instance of E54 Dimension is specific to an instance of E70 Thing. <ul style="list-style-type: none"> ▪ silver cup 232 (E22) <i>has dimension</i> height of silver cup 232 (E54) <i>has unit (P91)</i> mm (E58), <i>has value (P90)</i> 224 (E60)

fig.11 P43 Property Declaration

Crofts et al. (2010, p. 49)

⁶⁵ Le classi, invece, possono essere definite autonomamente, come “nomi” grammaticali o sostantivi.

L'impostazione evento-centrica

Il nucleo del CRM è formato dalle E2 Temporal Entities, le uniche classi che possono essere collegate al tempo (o, meglio, agli intervalli di tempo). Stephen Stead afferma:

A really powerful way to understand things that have happened in the past is to understand the relationships between those things and where they've met in time and space, so we can think of all of the things that have happened to our objects - our administration of them, our scholarly research, the things that happened to them in the past - as being <part of> events which have had different things there, and which have therefore expressed part of a story. So we can understand lives from the relationships between things and where they've met.⁶⁶

Gli eventi rappresentano quel livello di conoscenza in cui tutte le informazioni che rientrano nell'ambito dell'ontologia CRM possono essere armonizzate. Essi sono intesi come «volumi di coerenza spazio-temporale»⁶⁷ entro i quali avvengono gli incontri tra oggetti persistenti, sia oggetti fisici che concettuali. L'evento descrive quindi sia “ciò che è accaduto all'oggetto” (la creazione dell'opera d'arte, il rinvenimento del reperto archeologico, la nascita del redattore del trattato), sia tutti i processi mediante i quali l'oggetto è amministrato, studiato e interpretato (il restauro, l'acquisizione, la definizione della classificazione, la storia della terminologia che lo identifica). Le E2 Temporal Entities sono disgiunte dalle E77 Persistent Items: un'istanza può appartenere solo ad una delle due classi, le cui estensioni sono perciò due insiemi che non si intersecano.

La differenza ontologica tra le due classi è legata al loro comportamento nel tempo; le E77 Persistent Items sono *enduranti*: sono interamente presenti (tutte le parti che le compongono sono presenti) in ogni momento in cui sono presenti. Le E2 Temporal Entities invece sono *perduranti*: si estendono nel tempo per accumulazione di differenti parti temporali in modo che, in ogni momento in cui sono presenti, lo sono solo parzialmente, nel senso che alcune delle parti che le costituiscono (per esempio le loro fasi passate o future) possono non essere presenti. Le entità *enduranti sono nel tempo*, tutte le loro parti fluiscono con loro nel tempo (non hanno parti temporali), mentre le *perduranti accadono nel tempo* e le loro parti sono fissate nel tempo (hanno parti temporali).

Il CRM trae questa distinzione da DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering)⁶⁸, ontologia progettata dal Laboratorio di Ontologia Applicata dell'Istituto di Scienze e Tecnologie della Cognizione del CNR nell'ambito del progetto europeo *WonderWeb: Ontology Infrastructure for the Semantic Web* (2001-2004). DOLCE è stata sviluppata con «lo scopo di

66 Stead, S. (2008). *The CIDOC CRM, a Standard for the Integration of Cultural Information*, ICS-FORTH, Crete
http://www.cidoc-crm.org/cidoc_tutorial/index.html

67 *Ibidem*.

68 Gangemi, Guarino, Masolo, Oltramari e Schneider (2002).

catturare le categorie ontologiche che emergono nel linguaggio naturale e nel senso comune» (Borgo, Gaio, Guarino, Masolo e Oltramari, 2010, p. 111), entro l'obiettivo più ampio del progetto di sviluppare una “libreria” di ontologie fondazionali integrabili semanticamente.

Nella mappatura di un set di informazioni nel modello CRM si nota che le entità perduranti sono quelle classi che hanno la natura di occorrenza/happening (eventi o attività) entro un limite temporale, mentre le entità enduranti sono iniziatrici, testimoni o presenti a tali eventi e attività e li attraversano preservando la propria identità.

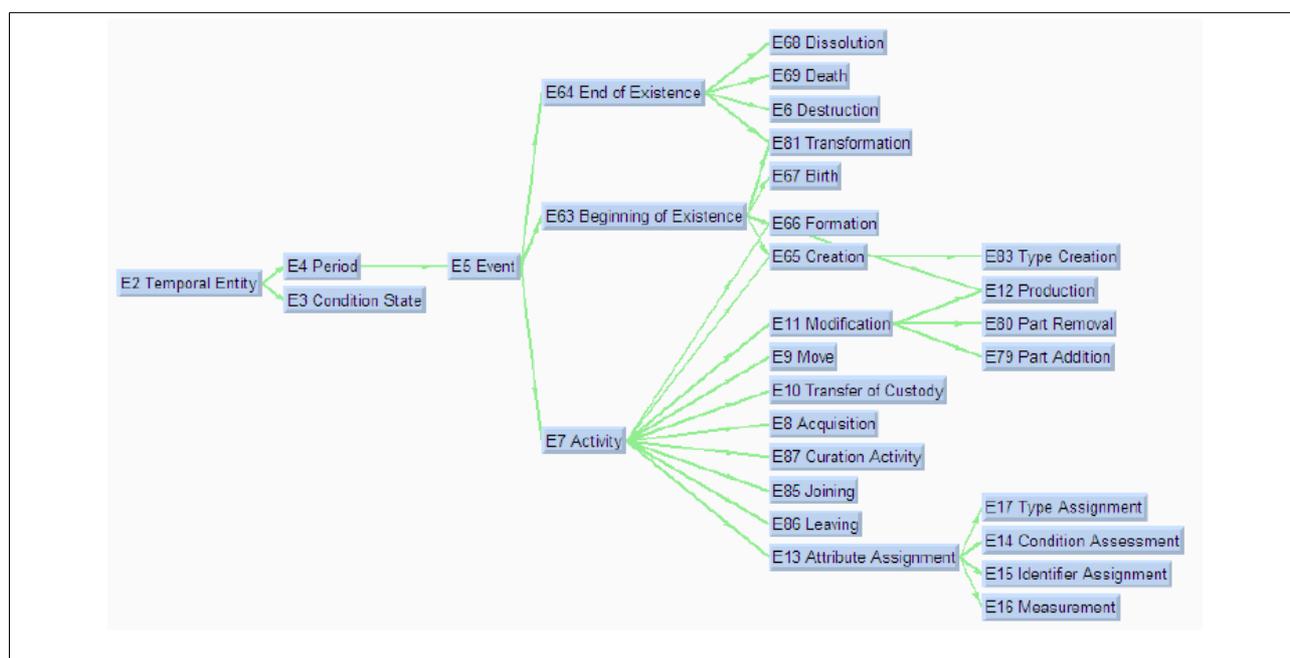


fig.12 Rappresentazione della gerarchia di E2 Temporal Entities (entità perduranti)

Stead, S. (2008). *The CIDOC CRM, a Standard for the Integration of Cultural Information*, ICS-FORTH, Crete

http://www.cidoc-crm.org/cidoc_tutorial/index.html

Il *top-level* del CRM

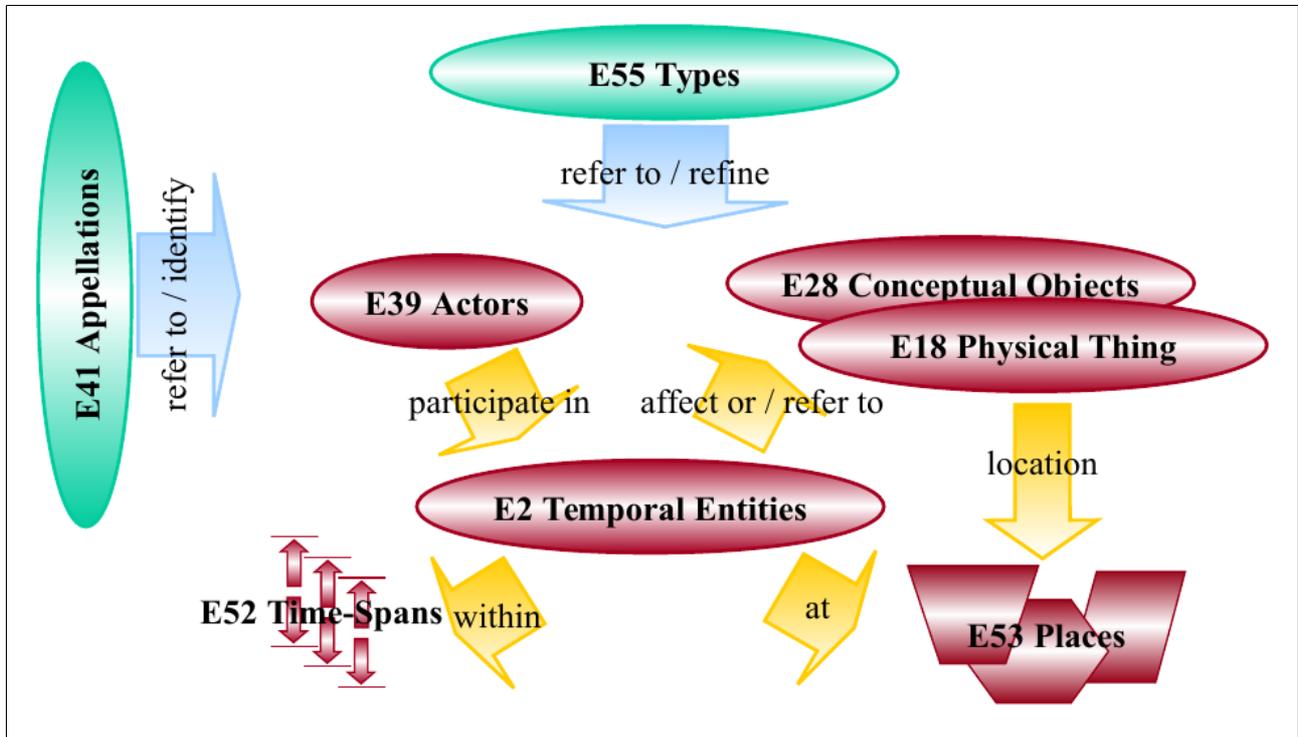


fig.13 The CIDOC-CRM top-level classes useful for integration

Oldman, C. e CRM Labs (2014, p. 8)

Solo le entità temporali possono essere collegate al tempo ed avere E52 Time-Spans (intervalli di tempo). Gli oggetti (E28 Conceptual Objects e E18 Physical Objects), gli attori (E39) e i luoghi (E53 Places) non possono essere direttamente collegati al tempo: per esserlo, devono prendere parte ad un evento.

Il tempo non è mai considerato come un punto ma sempre come durata; il CRM mette a disposizione una serie di proprietà per registrare la durata temporale anche in condizioni di incertezza dell'informazione⁶⁹.

I luoghi sono definiti geometricamente, e non corrispondono necessariamente a una collocazione politico-geografica: un E53 Place può essere sia un luogo geografico, sia la prua di una nave, il retro di una moneta o l'angolo di un quadro. Gli attori (E39 Actors) sono entità con responsabilità

⁶⁹ P81 *ongoing through*, P82 *at some time within*, P83 *had at least duration* e P84 *had at most duration*. È inoltre abilitato l'uso degli operatori di Allen.

legale e possono essere individui o gruppi di persone; interagiscono con gli oggetti persistenti (E77 Persistent Items) che possono essere E18 Physical Things o E28 Conceptual Objects.

E18 e E28 sono classi *disgiunte*: un oggetto può essere istanza di una e una sola delle due classi. La differenza più importante sta nel fatto che le E18 Physical Things possono essere distrutte: un oggetto è distrutto quando cessa di essere funzionale al dominio di documentazione ed esce dall'ambito di interesse del CRM⁷⁰. Un E28 Conceptual Object può essere distrutto solo con l'eliminazione di tutti i suoi "supporti", tra cui la mente umana, il che non è possibile⁷¹.

Una caratteristica importante del *top-level* CRM sono le classi E41 Appellations e E55 Types: il CRM permette di attribuire nomi e classificazioni a ogni oggetto appartenente al suo ambito d'interesse. E1 CRM Entity, il vertice della poligerarchia di classi CRM, è il dominio di "P1 *is identified by*: E41 Appellation" e "P2 *has type*: E55 Types", proprietà che sono automaticamente ereditate da tutte le classi CRM. Applicando infatti la funzione *object-oriented* dell'ereditarietà stretta, ogni classe che discende da E1 CRM Entity (quindi tutto il dominio del CRM) ne eredita le proprietà dichiarate, senza eccezioni⁷².

Quindi gli oggetti sono sempre separati da nomi e classificazioni, che possono essere multipli e cambiare nel tempo. Siccome E41 Appellations e E55 Types sono anche E28 Conceptual Objects, essi possono essere trattati come qualsiasi altra entità collegabile ad un evento, e se ne può registrare la storia. Il CRM conferisce una particolare importanza al processo di nominazione, e mette a disposizione una gerarchia apposita.

70 Se un oggetto invece viene disassemblato, e parti di esso sono utilizzate per creare un nuovo oggetto, quest'ultimo diventa parte del dominio del CRM; l'oggetto originario quindi non va incontro a una E6 Destruction, ma a una E81 Transformation.

71 «If I take a desk and I break off its legs and make it into little pieces of wood, and I'm not going to record what happened to those little pieces of wood, then it's been destroyed and it's no longer part of our domain of documentation. However, if I take something apart into its component pieces and then I keep track of them – I take the legs off and I start tracking the legs - they've not been destroyed. The table has not been destroyed in the sense we talked about, it has been transformed into some other things. (...) Conceptual objects can't be destroyed, because they're a product of the human mind, an idea. Consequently, we can't destroy them – all we can do is destroy the last carrier of them» Stead, S. (2008). *The CIDOC CRM, a Standard for the Integration of Cultural Information*, ICS-FORTH, Crete. http://www.cidoc-crm.org/cidoc_tutorial/index.html

72 Le altre proprietà di E1 CRM Entity ereditate da tutte le sottoclassi sono P3 *has note*: E62 String, P48 *has preferred identifier*: E42 Identifier, P137 *exemplifies*: E55 Type. Con E62 String non si intende solo una stringa di testo, ma qualsiasi BLOB (Binary Large Object): quindi qualsiasi tipo di file contenente una codifica binaria di dati (testi formattati, immagini, file audiovisivi ecc.) può essere associato a ogni E1 CRM Entity.

La classe E59 Primitive Value (composta da 3 sottoclassi: E62 String, E61 Time Primitive e E60 Number) esprime un insieme di valori estranei al dominio del CRM, ma indispensabili per connettere l'ontologia alla tecnologia informatica di riferimento. E59 Primitive Value e E1 CRM Entity sono quindi posti sullo stesso livello. Siccome E62 String non eredita da E1 CRM Entity la proprietà P2 *has type*, esiste una specifica proprietà P3.1 *has type*: E55 Type, che permette di classificare il tipo di contenuto di E62 String. Come risultato, ogni classe dell'ambito E1 CRM Entity può essere collegata a un BLOB (usando P3 *has note*) che può essere classificato come tipo di risorsa (usando P3.1 *has type*).

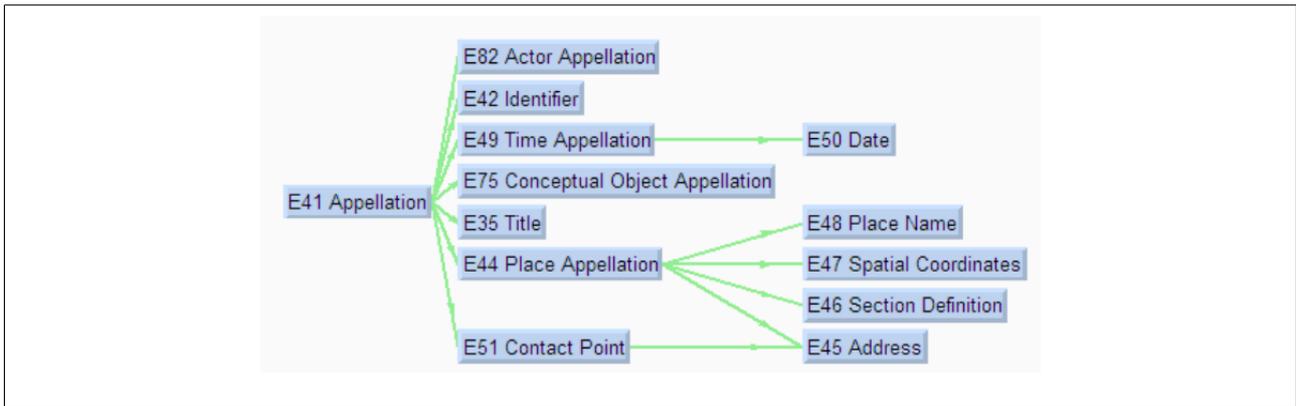


fig. 14 Rappresentazione della gerarchia di E41 Appellations

Stead, S. (2008). *The CIDOC CRM, a Standard for the Integration of Cultural Information*, ICS-FORTH, Crete

http://www.cidoc-crm.org/cidoc_tutorial/index.html

E55 Type rappresenta l'interfaccia tra il CRM e ontologie, vocabolari e *thesauri* specifici di dominio; è possibile, all'interno del CRM, creare una gerarchia di termini collegando istanze di E55 Type tramite la proprietà P127 *has broader term (has narrower term)*⁷³. Inoltre con E83 Type Creation si può registrare il processo accademico e scientifico che ha portato alla definizione di una classe entro una classificazione⁷⁴.

CRM prevede uno specifico evento E17 Type Assignment per registrare la circostanza dell'attribuzione di una classe a uno specifico oggetto. E17 è sottoclasse del più generico E13 Attribute Assignment che ha anche le seguenti sottoclassi: E14 Condition Assessment, E15 Identifier Assignment, E16 Measurement.

Come premesso, *ogni evento* accaduto all'oggetto della descrizione è registrato dal CRM come E5 Event; in questa classe sono comprese anche le attività connesse a ricerca, conservazione e amministrazione. La comune pratica di catalogazione interna alle istituzioni museali, però, non prevede nella maggior parte dei casi la descrizione di processi così dettagliati: non si parla dell'*evento* della misurazione, né si registrano attore, luogo e tempo di tale evento, ma esistono solamente i campi che ne descrivono gli esiti. Il CRM prevede a questo scopo l'uso di *shortcut*, singole proprietà che rappresentano una deduzione di un processo più esteso. P2 *has type*, P43 *has*

⁷³ Sarebbe interessante verificare la possibilità di descrivere la sistematica Hornbostel-Sachs usando la complessa struttura del CIDOC-CRM. Oltre all'uso della proprietà 127 (che permetterebbe di definire solo strutture ad albero) le specifiche CRM della versione 5.0.4 suggeriscono «such hierarchies may be extended with additional properties» (Crofts et al., 2010, p. xvii). Nel caso della HS bisognerebbe disporre di proprietà che rappresentino le funzioni della punteggiatura prevista dagli autori per la descrizione di strumenti poliorganici e per l'apposizione di suffissi.

⁷⁴ Siccome rientrano nell'ambito del CRM anche le *Life Sciences*, particolare attenzione è stata posta alla descrizione del processo di creazione della classificazione, del quale si possono indicare *taxon*, *protologue* e *holotype*.

dimension, *P44 has condition* e *P52 has current owner* sono esempi di *shortcut* mediante i quali rappresentare i campi tipici della documentazione museale, senza sviluppare il percorso completo di *E17 Type Assignment*, *E16 Measurement*, *E14 Condition Assessment* e *E8 Acquisition*. Sta all'esperto del dominio decidere come documentare i dati in origine e come mapparli in CRM (andando eventualmente a cercare gli "eventi nascosti" nel proprio database).

Le strategie messe a punto dal CRM per trattare nomi e tipi (la separazione dall'oggetto, la possibilità di strutturarli in gerarchie e trattarli "storicamente" come qualsiasi altro oggetto concettuale) rispondono ad un più ampio intento di armonizzazione dei dati nell'ottica della loro distribuzione nel Web e preservazione nel tempo. Il CRM intende incorporare le strutture e le discussioni relative a terminologie, *thesauri* e classificazioni nella stessa struttura usata per discutere degli oggetti d'interesse. Quando si trasferiscono i dati da un sistema all'altro è possibile rinvenire nella stessa struttura i dati e le istruzioni usati per classificarli e identificarli. Questo è un obiettivo importante se si cercano soluzioni per preservare a lungo termine dati e conoscenze in formato digitale.

Aspetti critici e modelli alternativi

Il CRM può essere considerato a tutti gli effetti una ontologia *fondazionale* (*upper ontology, top-level ontology*) che descrive concetti molto generali comuni a molteplici sotto-domini; lo scopo di un'ontologia di questo tipo è predisporre un'infrastruttura che supporti l'interoperabilità semantica tra diverse ontologie *domain-specific*.

Se è vera la premessa per cui il CRM rappresenta la conoscenza entro il dominio del patrimonio culturale, è evidente che i confini di tale dominio sono strutturalmente deboli e permeabili e tendono ad ammettere al proprio interno un numero crescente di “porzioni” di conoscenza. Gli autori del CRM ne sono consapevoli, e sostengono che il riferimento a un'unica ontologia *top-level* di riferimento sia la chiave per l'interoperabilità semantica:

«Everybody thinks that ontologies are very domain specific - that when you go for an ontology, what you're actually doing is really hemming yourself into working in a particular small environment. And what we've found with the CRM is almost the opposite - nearly all of it is *not* domain specific. (...) And in fact we've been successfully deploying the CRM into medicine to look at the derivation and processing of cancer samples; <into> biodiversity, processing information about collecting samples and documenting biodiversity sampling. A whole range of different things use the same basic concepts that we use. So it's a notion of scientific method - retrospective analysis, taxonomic discourse, which is common to archaeology and natural history - which show us that this model is useful *across* domains, it's not very specific to museums. We've been able to do it with a very small set of concepts»⁷⁵.

Altrove sono state mosse critiche nei confronti del CRM e delle ontologie fondazionali in genere; ciò che suscita diffidenza è proprio la loro vocazione olistica che, coerentemente supportata dal punto di vista teorico, tende però a generare strutture troppo complesse per poter essere implementate nell'atto pratico.

In seguito a uno studio di implementazione del CRM in ambito archeologico Haslhofer e Nussbaumer (2007) affermano: «this gap between the well-defined conceptual and the undefined technical level could result in additional incompatibility instead of interoperability» (p. 1). Gli autori, nell'ambito del progetto BRICKS⁷⁶, hanno applicato il CRM con lo scopo di fornire un accesso uniforme al patrimonio digitale proveniente da diverse istituzioni archeologiche. Siccome

75 Stead, S. (2008). *The CIDOC CRM, a Standard for the Integration of Cultural Information*, ICS-FORTH, Crete http://www.cidoc-crm.org/cidoc_tutorial/index.html

76 BRICKS – Building Resources for Integrated Cultural Knowledge Services (IST 507457) (2007) - <http://www.brickcommunity.org>

Nella loro ricerca Nussbaumer e Haslhofer hanno trattato dati relativi a reperti archeologici (monete) ritrovati nel Regno Unito e custoditi in diverse istituzioni, ciascuna delle quali adottava uno standard proprietario di catalogazione diverso.

ognuna di queste istituzioni esponeva i propri dati in base a un modello descrittivo diverso, BRICKS ha progettato una mappatura delle fonti originarie in CRM; in questo modo si sarebbe creato un accesso uniforme a dati eterogenei, essendo tutte le fonti raggiungibili interrogando, tramite motore di ricerca, un'unica ontologia di armonizzazione.

Due principali argomenti sono stati individuati come ostacoli all'effettiva messa in pratica di questo progetto. In primo luogo il CRM definisce concetti astratti (es. E49 Time Appellation e E22 Man-made Object) che danno origine a interpretazioni ambigue: accade che diversi esperti di dominio (e di CRM) producano mappature differenti della medesima istanza; cosa che ovviamente allontana dal raggiungimento dell'interoperabilità semantica auspicata.

Un'eccessiva astrazione inoltre, obbliga a disegnare percorsi complessi per definire istanze semplici; una “moneta prodotta in epoca romana” è descritta da questo insieme di classi e proprietà:

E22 Man-Made Object – P108 was produced by – E12 Production Event – P10 falls within – E4 Period – P1 is identified by – E49 Time Appellation – *Roman* ⁷⁷

Tale definizione risulta poco intuitiva e pratica per poter essere implementata in un motore di ricerca indirizzato ad utenti non esperti di CRM; in altre parole, affinché l'*end-user* possa formulare interrogazioni su concetti più concreti e familiari, è necessario programmare un'interfaccia che *nasconda* la complessità della struttura che soggiace all'ontologia.

La seconda critica al CRM è che la documentazione (le “specifiche tecniche”) che lo definisce è carente di istruzioni per l'implementazione delle mappature e la rappresentazione delle istanze; per questo motivo ogni istituzione applica la propria interpretazione e ciò causa ulteriore eterogeneità in uno scenario che originariamente punta all'interoperabilità.

In conclusione, secondo gli autori, il CRM condivide il problema di fondo di tutte le ontologie fondazionali: non possono essere applicate nel caso di scenari di integrazione di schemi descrittivi *già esistenti*; dal punto di vista semantico infatti tali schemi non vanno interpretati come estensioni dell'ontologia globale, bensì come mappature artificiali. Le ontologie fondazionali si rivelano uno strumento utile solo nel caso di ontologie che non sono ancora state implementate, ovvero come concettualizzazione di riferimento sulla quale costruire future ontologie interoperabili. Al *global ontology approach* gli autori contrappongono un *multiple ontology approach*: ogni fonte di informazione è descritta dalla propria ontologia e l'interoperabilità semantica è garantita da mappature bilaterali (Haslhofer e Nussbaumer, 2007, p. 19). Questo metodo garantisce la pluralità

⁷⁷ Haslhofer e Nussbaumer (2007, p. 19)

delle prospettive locali ed ogni singola ontologia può essere trattata in maniera autonoma, senza che variazioni o aggiunte locali debbano essere supportate da un raffinamento della struttura globale. D'altra parte, le mappature bilaterali pongono non pochi problemi di interoperabilità semantica, anche a causa dei possibili diversi livelli di aggregazione e granularità dei concetti.

L'*hybrid ontology approach* descritto da Hubner et al. (2001, pp. 2-3) è un approccio intermedio che supporta l'uso di ontologie multiple; queste però non sono concepite in maniera isolata, ma fanno riferimento ad un vocabolario *top-level* condiviso che fornisce i concetti base e quindi garantisce alle ontologie locali un livello minimo di interoperabilità semantica.

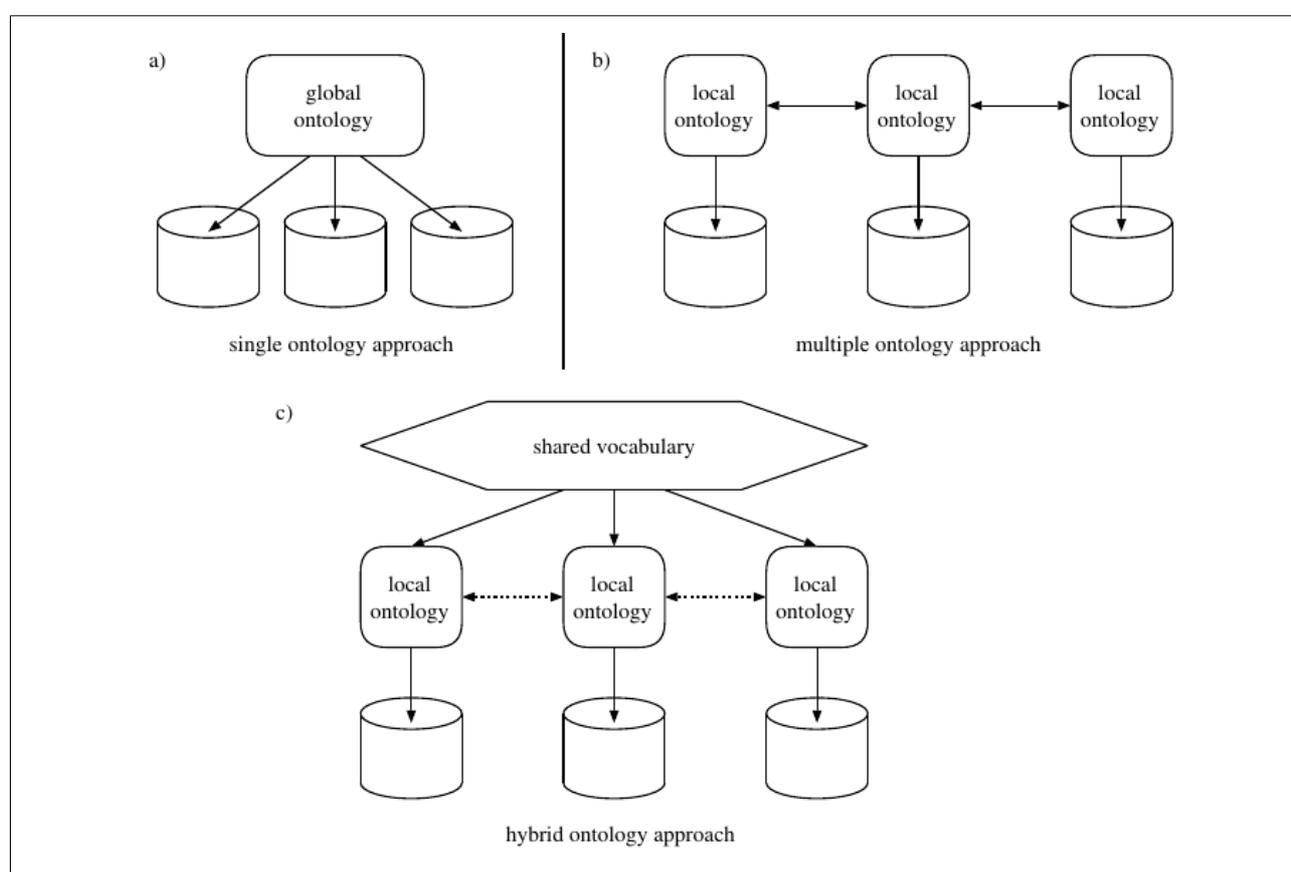


fig.15 The three possible ways for using ontologies for content explication

Hubner et al. (2001, p. 2)

Una soluzione di questo tipo è quella adottata dall'*Open Memory Project*, realizzato dal CDEC⁷⁸ e da *regesta.exe*⁷⁹; il lavoro ha portato alla creazione di un portale web sul quale sono pubblicati, con le tecnologie dei *Linked Open Data*, i dati del CDEC relativi alle vittime della *Shoah* in Italia⁸⁰. Il progetto ha dovuto affrontare il problema dell'integrazione di dati eterogenei provenienti da dipartimenti separati (archivi storico, fotografico e audiovisivo, biblioteche di libri e periodici e un database di circa 60000 identità personali), ciascuno dei quali adottava il proprio standard descrittivo. La soluzione ha condotto all'individuazione dei «nomi di persona» come «elemento univoco di raccordo e quindi di accesso alle diverse fonti informative» (Brazzo e Mazzini, 2015, p. 2), e alla definizione di una libreria di ontologie e vocabolari da utilizzare per la descrizione e interconnessione semantica dei dati. Tale libreria fa riferimento a una serie di vocabolari RDF diffusi e consolidati⁸¹; inoltre è stata adottata *Shoah Ontology*, un'ontologia creata *ad hoc* per descrivere ed esprimere al meglio le caratteristiche specifiche del database locale⁸². Il riferimento a un modello calibrato specificamente sulle relazioni esistenti tra i concetti del dominio “*Shoah*” ha reso più semplice ed efficace recupero, disambiguazione e connessione dei dati; di conseguenza l'applicazione dei LOD in un ambiente già semanticamente ben strutturato, ha permesso di inferire nuova conoscenza con maggiore precisione⁸³.

78 CDEC (Centro di Documentazione Ebraica Contemporanea) - <http://www.cdec.it/>

79 Regesta.exe - <http://www.regesta.com/>

80 *Open Memory Project* è stato presentato come *case history* al seminario “Lavorare con i linked data” svoltosi a Torino il 16 giugno 2016 e promosso da AIB Piemonte, ICOM Piemonte e Valle d'Aosta, ANAI Piemonte e Valle d'Aosta, MAB Piemonte e Dipartimento di Studi storici dell'Università di Torino.

81 FOAF (Friend Of A Friend - <http://xmlns.com/foaf/0.1/>) e Biographic Ontology (<http://purl.org/vocab/bio/0.1/>) per i metadati delle persone, Dublin Core (<http://purl.org/dc/elements/1.1/>) e Dublin Core Terms (<http://purl.org/dc/terms>) per la descrizione di documenti e risorse digitali, OWL (<http://www.w3.org/2002/07/owl#>) per descrivere le relazioni tra ontologie e vocabolari, Geonames (<http://www.geonames.org/ontology#>) base di dati geografici, VIAF (Virtual International Authority File - <http://viaf.org/ontology/1.1/#>) base dati di voci di autorità controllate.

82 «Al momento l'ontologia comprende 9 *classes*, 27 *object properties* e 26 *datatype properties*; sono stati creati indici sui nomi dei campi di concentramento e sterminio nazisti, sui luoghi di raccolta e detenzione, sui convogli e gli eccidi avvenuti in Italia prima della deportazione. Attraverso la creazione di IRI univoche per tutte le risorse del dominio, è stato possibile ricondurre tutte le informazioni a “nodi” univocamente riconosciuti nel dominio senza la proliferazione di informazioni» Brazzo e Mazzini (2015, p. 3). Le informazioni di riferimento di Shoah Ontology sono pubblicate a questo indirizzo: <http://dati.cdec.it/lod/shoah/reference-document.html>.

83 I risultati sono pubblicati sulla Digital Library del CDEC (<http://digital-library.cdec.it/>): nella sezione “Persone” è possibile fare una “Ricerca avanzata sulla Shoah in Italia”. Il progetto Digital Library è stato realizzato con il DBMS Open Source xDams per la catalogazione e inventariazione della documentazione, e la piattaforma *linked data* openDams per la gestione integrata degli archivi e delle basi di dati, entrambi sviluppati dalla società *regesta.exe*. openDams è una piattaforma che guida alla costruzione dei LOD secondo i passaggi del W3C; i dati sono codificati in RDF/XML ed è possibile scegliere o introdurre le ontologie necessarie per descriverli.

Shoah Ontology è un esempio di ontologia costruita attorno al concetto di “Persona”, alternativo a quello di “Evento” caratteristico del CRM; quest’ultimo approccio è da alcuni considerato troppo complesso, soprattutto nel caso di un’ontologia fondazionale che si propone come modello per numerosi domini del patrimonio culturale: sembra che il CRM sia concepito soprattutto per rispondere alle esigenze di rappresentazione del campo archeologico, che è caratterizzato da una particolare complessità nel rapporto fra oggetti e tempo. Ma è probabile che per la resa delle descrizioni di altri domini il peso della logica evento-centrica non sia necessario.

Europeana ha concepito EDM come una struttura aperta, *cross-domain*, in grado di adattarsi alla varietà e ricchezza degli standard nativi dei *data provider*, quali LIDO per i musei, EAD per gli archivi e METS per le biblioteche digitali (Isaac, 2013, p. 5). Anche EDM, come il CRM, può essere considerato un’ontologia *top-level*, nel senso che trascende le prospettive dei singoli domini; ma, a differenza del CRM, non punta a descrivere l’intera complessità del *top-level*, bensì a fornire un piccolo set di proprietà e classi a cui “allegare” descrizioni specifiche più dettagliate.

Nel fare questo EDM permette, per uno stesso oggetto, la coesistenza di un modello descrittivo oggetto-centrico e di uno evento-centrico, con una predilezione per il primo, il cui *core* è sviluppato con maggiore dettaglio. L’approccio ontologico ibrido di EDM consente il riuso di metadati strutturati in *Dublin Core*, il vocabolario più consolidato per le descrizioni oggetto-centriche ed ampiamente diffuso nei database di tutto il mondo: così, mappando i propri dati in EDM, le istituzioni non sono forzate a ripensarne la struttura concettuale. Anche se si impegna a sviluppare maggiormente nel tempo la compatibilità con iniziative come CIDOC-CRM e LIDO, Europeana richiede quindi l’invio di una descrizione «*object-centric core*» a fianco di quella evento-centrica (*ivi*, p.17).

Per analizzare le differenze tra i due modelli è necessario introdurre i tre concetti chiave di EDM:

- **edm:ProvidedCHO** (*provided cultural heritage object*), identifica l’oggetto “reale”;
- **edm:WebResource** identifica una o più rappresentazioni digitali dell’oggetto;
- **ore:Aggregation** identifica il risultato dell’attività del *provider*, ovvero l’aggregazione delle identificazioni dell’oggetto “reale” e delle sue rappresentazioni digitali.

L'aggregazione è definita usando classi e proprietà ORE⁸⁴, ontologia dell'Open Archives Initiative che definisce standard per la descrizione e lo scambio di aggregazioni di risorse web.

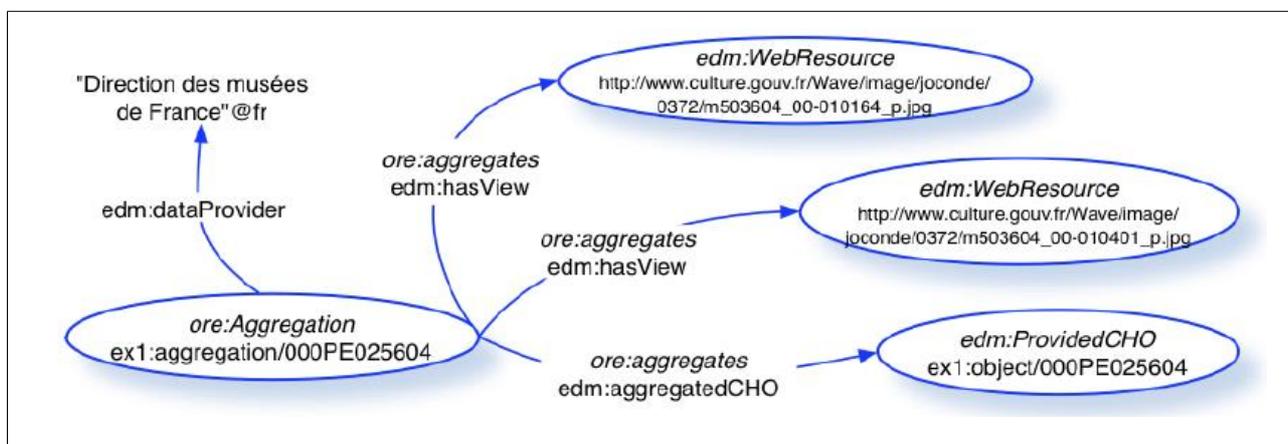


fig.16 Provider's aggregation of web resources and provided CHO

Isaac (2013, p. 11)

[edm:hasView e edm:aggregatedCHO sono sottoproprietà di ore:aggregates]

In una descrizione oggetto-centrica l'informazione si presenta sotto forma di dichiarazioni che forniscono un collegamento diretto tra l'oggetto edm:ProvidedCHO e le sue caratteristiche.

84 ORE (Object Reuse and Exchange) - <https://www.openarchives.org/ore/>

Nell'esempio di fig.17 l'oggetto (la "Gioconda" di Leonardo) è coinvolto in un processo di *metadata enrichment*: l'entità `edm:Agent` infatti rimanda a un URI del database VIAF, che a sua volta collega, tramite i vocabolari `rdaGr2` e `SKOS`, l'ID che rappresenta l'identità del creatore dell'opera (Leonardo) alle sue date vitali e alle varianti linguistiche del nome.

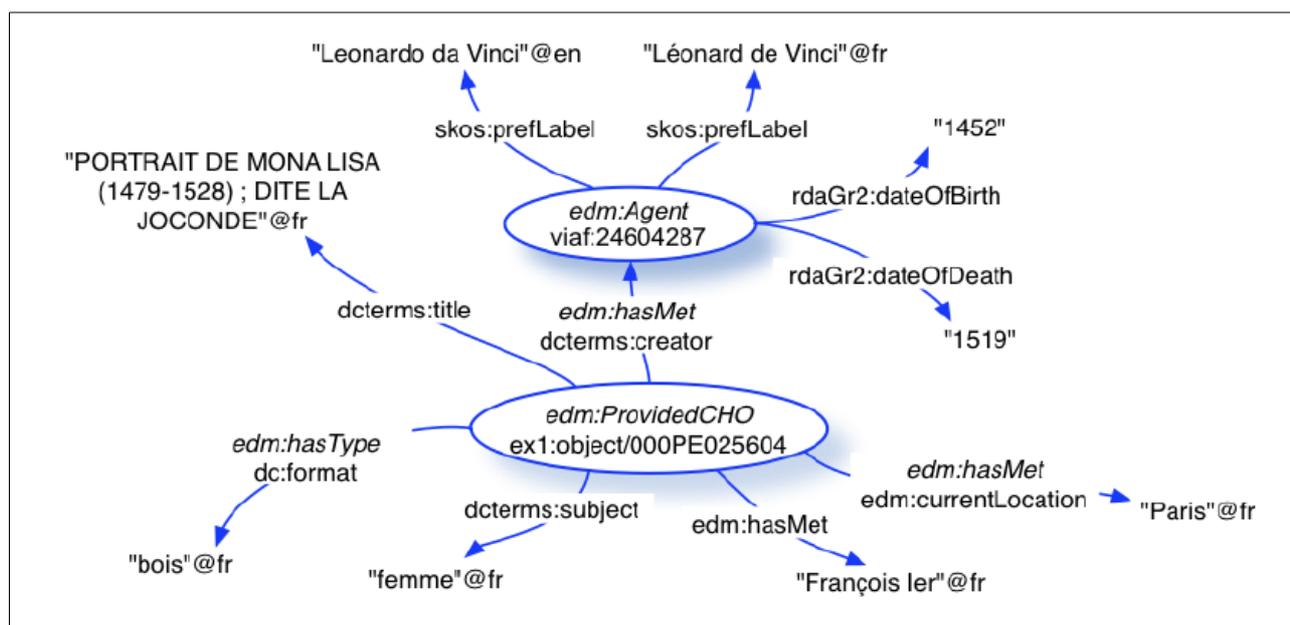


fig.17 An object-centric description enriched with an Agent contextual entity

Isaac (2013, p. 14)

Lo stesso esempio tradotto in un modello evento-centrico richiede l'inserimento di due eventi: 000PE025604-c e 000PE025604-a che denotano rispettivamente la creazione e l'acquisizione del dipinto, in un sistema semanticamente più ricco, ma concettualmente complesso.

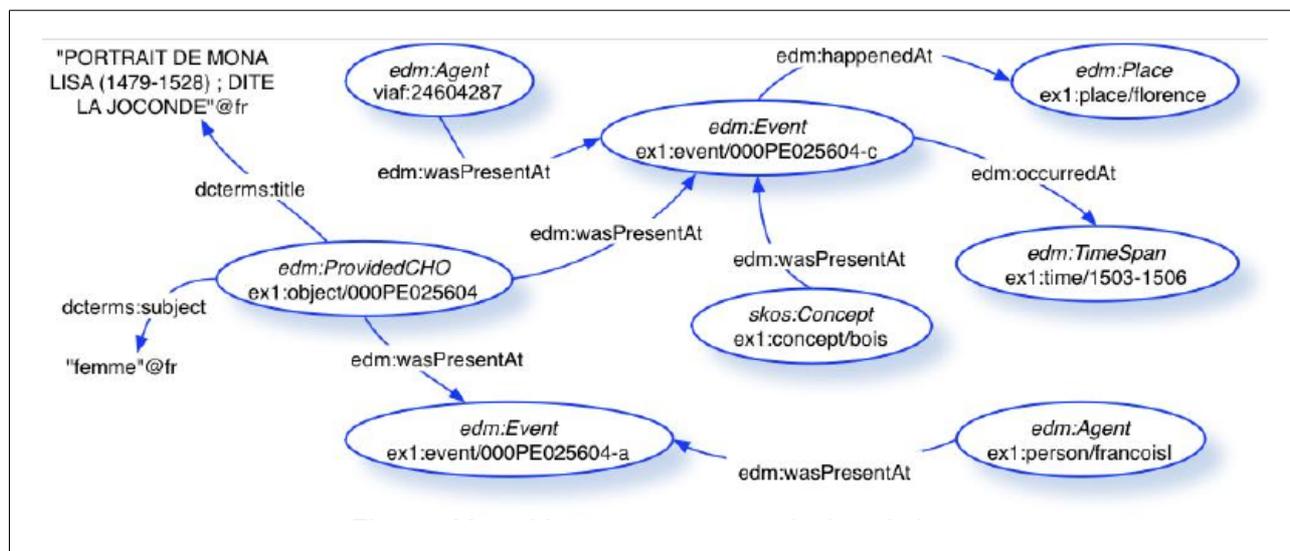


fig.18 Mona Lisa - an event-centric description

Isaac (2013, p. 16)

Nel momento in cui MIMO ha caricato i propri dati su Europeana, EDM era alla sua prima implementazione: è stata la prima volta in cui un aggregatore (e quindi non un semplice *data provider*) ha consegnato i propri dati ad Europeana usando il nuovo modello orientato al *semantic web*. In questa prima versione la classe `edm:Event` non era ancora contemplata, per cui tutta l'informazione legata al concetto di “evento” è andata perduta nella mappatura da LIDO a EDM; fa eccezione quella legata alla “creazione”, che è stata preservata nelle nuove descrizioni (Bailly, R. *EDM Case Study*).

Terza parte

La Scheda Guizzi e il nuovo contesto tecnologico

Il modello di una teoria organologica

Storia, funzioni, adattamenti

Il catalogo del Museo del Paesaggio Sonoro è costituito da circa 250 schede catalografiche in formato documentale (file di testo) realizzate nel 2007: esso ricopre quindi solo una parte dei circa 1000 oggetti che costituiscono il fondo del museo. Per le descrizioni è stata utilizzata la Scheda Guizzi, il modello ideato da Febo Guizzi e già sperimentato “sul campo” nel lavoro di catalogazione svolto per diversi musei e collezioni private¹.

In Italia non esiste uno standard nazionale per la catalogazione degli strumenti musicali: l'ICCD (Istituto Centrale per il Catalogo e la Documentazione) ha elaborato la scheda SM (Strumenti Musicali), che risulta però ancora tra le normative in fase di sperimentazione². L'assenza di uno standard *de jure* non è stato comunque il motivo che ha condotto all'adozione della Scheda Guizzi (che rappresenta a tutti gli effetti uno standard *de facto* per la catalogazione degli strumenti musicali in Italia, dal momento che è stata applicata in maniera uniforme in molteplici occasioni)³; siamo convinti che essa costituisca, oggi, lo strumento migliore disponibile in Italia per una corretta e ricca descrizione delle collezioni organologiche e di quelle etno-organologiche in particolare. A partire da questa constatazione, è nostra intenzione iniziare un percorso di evoluzione e allineamento della scheda alle più moderne tecnologie dell'informazione, in particolare a quelle relative a *semantic web* e *Linked Open Data* che stanno destando particolare interesse nell'ambito della gestione del patrimonio di biblioteche, archivi e musei.

Il lavoro di ammodernamento ha condotto all'elaborazione di un XML Schema (versione 1.0) che rappresenta la prima completa digitalizzazione della scheda. Sono inoltre state redatte le istruzioni per una mappatura della versione 1.0 nell'XML schema LIDO (in riferimento alla struttura LIDO delimitata dalla piattaforma MIMO).

1 Museo degli strumenti musicali del Castello Sforzesco di Milano, Raccolta teatrale Coccia di Romentino di Novara, Museo Ettore Guatelli di Ozzano Taro, Museo Cultura Musica Popolare dei Monti Peloritani di Gesso, Civiche raccolte di arte antica di Torino, Lascito Cuneo di Calvari, e altre raccolte catalogate nell'ambito di lavori condotti per tesi di laurea.

2 <http://www.iccd.beniculturali.it/index.php?it/211/sperimentazione-normative>

3 Si segnalano anche le esperienze di catalogazione degli strumenti musicali avviate in Sicilia, con una scheda ideata da Giovanni Di Stefano e Girolamo Garofalo (Università di Palermo). La scheda è stata pubblicata dal “Centro regionale per l'inventario, la catalogazione e la documentazione grafica, fotografica, aerofotografica, audiovisiva e Filmoteca regionale siciliana” a cura di Giuliano e Proto (2010).

I possibili sviluppi della versione 1.0 sono orientati verso una maggiore atomizzazione dell'informazione e l'inclusione di descrizioni evento-centriche⁴; tali orientamenti hanno guidato la stesura di una seconda versione dell'XML Schema (versione 1.1 *draft*) e di un *corpus* di linee-guida per future implementazioni.

Il percorso compiuto ha costituito l'occasione per rivedere la Scheda Guizzi sotto nuove prospettive. Come è ovvio, l'adozione di una nuova tecnologia non si esaurisce nella mera mappatura del contenuto: la conversione dei formati e l'adozione di nuovi applicativi, linguaggi e *policies* delle tecnologie dell'informazione richiedono di indagare ulteriormente la natura del modello originario e verificarne l'efficacia a confronto con problemi e, soprattutto, possibilità nuove.

D'altra parte anche le nuove tecnologie sono sottoposte a verifica, messe di fronte al compito di rappresentare un modello complesso e consolidato come quello elaborato da Guizzi; esse devono dimostrarsi in grado non solo di restituire una rappresentazione digitale della sequenza di campi che costituiscono la scheda, ma anche di sostenere l'intreccio dei molteplici ambiti disciplinari, competenze e funzioni di cui la scheda è il fulcro⁵.

Guizzi ha concepito la sua scheda come uno strumento utile a svolgere tutte le funzioni necessarie per una catalogazione efficace, ovvero come il risultato della convergenza di diverse attività (amministrative, conservative, documentarie, descrittive, classificatorie ecc.) rivolte alla cura, valorizzazione e interpretazione dello strumento musicale. Il catalogatore è chiamato a mettere in campo una serie di competenze organologiche che si attuano e definiscono inderogabilmente nell'ambito di un incontro fattuale con lo strumento, considerato nella sua materialità oggettiva. È necessario che nel processo di digitalizzazione della scheda l'integrità di questo complesso di funzioni sia preservata.

Lo scenario di sviluppo del *semantic web* si configura in un «web of things in the world, described by data on the Web» (Berners-Lee, Bizer e Heath, 2009, p. 2): i dati si combinano e si aggregano nel Web al fine di costituire una rete di descrizioni di oggetti appartenenti al mondo reale. Supportati dalla semantizzazione operata dalle tecnologie LOD, i dati atomizzati acquistano crescente autonomia in un processo di automatizzazione che comporta un rischio di eccessiva alienazione dell'informazione dall'oggetto reale. Per scongiurare questo rischio è necessario che i

4 Quest'ultimo aspetto scaturisce dall'interesse suscitato dal progetto MIMO e dalla conseguente intenzione di rendere la Scheda Guizzi compatibile con il modello evento-centrico LIDO.

5 Considerando, per esempio, che l'attività classificatoria è una delle funzioni della catalogazione, segnaliamo che l'analisi che abbiamo condotto sulle rappresentazioni in ambito LOD delle classificazioni organologiche hanno evidenziato alcuni aspetti critici. Si veda a questo proposito il capitolo "Il problema delle *keywords* e della sistematica Hornbostel-Sachs".

dati pubblicati nel Web abbiano origine nell'ambito di procedimenti a stretto contatto con l'oggetto e con tutte le questioni generate dalla sua presenza; ogni passo diretto verso l'autonomia dei dati richiede il raffinamento della conoscenza dell'oggetto reale. Il vantaggio è duplice: il Web si assicura dati affidabili e strutturati a partire dai quali la macchina può effettuare processi inferenziali efficaci; l'oggetto ritrova nel Web un veicolo di amplificazione della propria presenza e l'intensificazione dell'interesse nei confronti della propria oggettività.

In questo contesto il modello di descrizione elaborato da Guizzi riflette proprio il senso che lo strumento ha in quanto *oggetto sonoro*; in questa concezione si sommano le competenze del cultore di organologia, etnomusicologia e antropologia della musica; e, per estensione, del curatore museale che conserva l'oggetto di tale studio.

Il concetto di *oggetto sonoro*

L'*oggetto sonoro* è inteso da Guizzi in una duplice accezione; in primo luogo c'è la constatazione della sua natura di *cosa materiale*. Nell' "Introduzione" al suo volume sugli strumenti della musica popolare in Italia, Guizzi osserva che in molte occasioni in cui ci si occupa di musica di tradizione orale, gli strumenti sono «gli unici testimoni diretti [...] che ci consentono uno sguardo sulla storia» (Guizzi, 2002, p. XX). In virtù della loro capacità di sopravvivere allo scorrere del tempo, essi sono autonomi «dalla transitorietà della memoria e della presenza umane» e sono disponibili a «essere analizzati e interpretati anche in assenza di altre testimonianze sul contesto musicale della loro diretta utilizzazione» (*ibidem*).

D'altra parte è proprio l'apporto che l'approccio etnomusicologico ha dato a tutta la scienza organologica a far sì che quell'*oggetto sonoro* sia collocato in una più ampia "cornice" antropologica:

alcuni dei criteri utilizzati nella più compiuta e matura opera di classificazione organologica⁶ [...] in particolare quelli relativi alle condotte fondamentali messe in atto dall'uomo per produrre suoni con gli strumenti [...] o per rendere adatti all'uso musicale con modalità diverse materiali di base, sono il risultato di una considerazione che solo l'approccio etno-musicologico aveva consentito di maturare, e cioè quello che prende in esame l'evento musicale in atto, all'interno del quale il rapporto tra uomo e strumento è posto al centro dell'interesse quale momento squisitamente antropologico. (Guizzi, 2002, p. XXII)

La relazione tra uomo e *oggetto sonoro* si manifesta nell'attualizzazione di una più o meno consapevole dimensione del "fare" culturale. È grazie alla messa in atto (anche potenziale) di determinate tecniche corporee, che lo strumento diventa un *nodo oggettivo complesso* che incorpora aspetti «costruttivi, funzionali, acustici, musicali, comunicativi, formali, estetici, simbolici» (*ivi*, p. XXIII).

Per la loro natura di cose materiali, si può dire che gli strumenti [...] possano essere collocati a rappresentare una sorta di edizione riveduta e corretta del livello "neutro" di Jean Jacques Nattiez⁷, posto in situazione intermedia tra quello poetico (quello della realizzazione dell'evento) e quello estesico (o della ricezione); esiste infatti indubbiamente un'autonomia dello strumento in sé nella sua consistenza materiale e separata dall'uomo e dalla sua corporeità, così come peraltro è imprescindibile la sua discendenza di manufatto da un processo costruttivo

6 Guizzi si riferisce alla classificazione elaborata da Hornbostel e Sachs (1914) della quale fornisce, in appendice al testo (Guizzi, 2012, pp. 409-482), la traduzione italiana corredata di un apparato di note che propongono emendamenti e integrazioni. [NdR]

7 Qui Guizzi si riferisce alle tesi elaborate in Nattiez (1975). [NdR]

che ne determina la “fattura” e, a sua volta, la sua integrazione funzionale nella sfera del “fare” musicale culturalmente codificato ed espressivo dell’umana predisposizione alla musica.⁸

Il ragionamento di Guizzi può essere riconsiderato come una progressione di momenti in cui l’esperto incontra l’*oggetto sonoro* e lo osserva adottando alternativamente prospettive diverse.

All’inizio predomina la prospettiva del curatore, che esamina lo strumento in quanto oggetto custodito nell’ambito di una collezione museale. Intervengono poi altre competenze, quella più strettamente organologica (che osserva i dettagli, confronta le esperienze, interpreta e classifica) e quella storico-antropologica (che raccoglie le narrazioni intorno all’oggetto e lo inserisce nel contesto): esse descrivono, definiscono e ipotizzano le connessioni tra lo strumento e le dimensioni del “fare” culturale. La consapevolezza dell’esistenza di tale ricchezza antropologica restituisce una visione prospettica più ampia a chi si occupa della custodia dell’oggetto; ed è proprio a questo punto che è concepita la Scheda Guizzi: in essa infatti traspare e non è mai abbandonata la predilezione per quell’incontro originario tra strumento e organologo.

La seconda accezione di *oggetto sonoro* è un esplicito richiamo a quella categoria di «materiali eterogenei» (principalmente oggetti d’uso quotidiano propri del lavoro domestico e del lavoro esterno) che hanno in comune la caratteristica di «trovarsi in uno stato di temporanea rifunzionalizzazione a scopo musicale, che li sottrae in modo estemporaneo alla loro funzione permanente» (Guizzi, 2002, p. 337). Questa definizione di *oggetto sonoro* è centrale nel pensiero di Guizzi, perché

nel momento in cui consente [...] di tenere aperto il campo reale della pratica strumentale a ogni esperienza, anche la più occasionale, introduce riferimenti preziosi al rapporto tra soggetto e oggetto nel vivo di tale pratica, richiamando così l’attenzione su un nodo essenziale del processo di creazione e di fissazione di uno strumentario.

(*Ibidem*)

Proprio in quanto risultato di rifunzionalizzazioni estemporanee gli *oggetti sonori* portano «tracce rilevanti, a volte addirittura clamorose, dei processi “germinali” di individuazione di strumenti

8 Guizzi (2002, pp. XXIII-XXIV). Il Museo del Paesaggio Sonoro pone particolarmente in risalto questo duplice movimento dell’oggetto che, materialmente autonomo, è contemporaneamente il risultato di un processo costruttivo e il potenziale oggetto di una sollecitazione sonora per mezzo di tecniche esecutive.

Gli strumenti della sua collezione, sia i dispositivi sonori (i giochi infantili, i richiami per la caccia) che gli strumenti musicali “propriamente detti”, sembrano sollecitare nell’osservatore le domande relative a “come è stato fatto” e “come si suona”. In questo caso l’aspetto più rilevante del paesaggio sonoro non è tanto relativo alla sua dimensione acustica, quanto a quella competenza culturalmente diffusa e condivisa del “fare” sonoro, legata a *quel* paesaggio sonoro e ai materiali di cui esso dispone. Il Museo cerca di preservare una simile originalità e ricchezza antropologica, e per questo motivo ha attivato una serie di iniziative didattiche rivolte soprattutto a bambini e adolescenti in età scolare. [NdR]

musicali» (*ivi*, p. 341). Nel tentativo di definire il concetto di strumento musicale *tout court* Guizzi sottolinea ulteriormente l'essenzialità di quel *rapporto tra oggetto e soggetto*, che abbiamo precedentemente collocato nella dimensione del “fare” culturale:

affinché si possa parlare di strumento sonoro, non è né necessario né sufficiente che vi sia [...] una relazione reificata (cioè tradotta permanentemente in una “cosa”) tra un'intenzione progettuale e una finalizzazione, entrambe concepite in termini di adeguatezza tecnica a produrre suono. Occorre invece e basta che all'oggetto [...] sia conferita una attuale ma momentanea destinazione funzionale all'impiego acustico. (*Ivi*, p. 335)

Quindi, a condizione che si condivida un'idea “ampia” del concetto di “musica”⁹, sarà da considerarsi “strumento musicale”

ogni oggetto che sia volutamente utilizzato in tale produzione sonora, indipendentemente dalla maggiore o minore complessità della sua struttura, dalla maggiore o minore specializzazione e difficoltà del processo lavorativo necessario per realizzarlo, o dalla univoca e primaria potenzialità a emettere suoni che ne possa qualificare il funzionamento. (Guizzi, 2002, p. 335).

In conclusione si verifica una sorta di reciproca identificazione: se in base a una definizione ampia del concetto di “musica” ogni *oggetto sonoro* è uno *strumento musicale*, per la scienza organologica ogni *strumento musicale* è anche un *oggetto sonoro*, in quanto “cosa” funzionalmente destinata all'impiego acustico.

Le accezioni di *oggetto sonoro* qui riportate si concentrano su due aspetti opposti ma complementari degli strumenti musicali: il loro essere “cose” materiali e il loro essere definiti e riconosciuti in quanto tali esclusivamente nell'ambito di una “relazione” attuata nella dimensione del “fare”. Risulta evidente la complessità di prospettive e di dettagli che si dispiegano di fronte al catalogatore nel momento dell'incontro con l'oggetto del suo studio.

9 A questo proposito Guizzi parla di «qualunque attività umana in cui la produzione di suoni sia perseguita o percepita in forma dotata di senso» (*ivi*, p. 335). Il riferimento al concetto di «humanly organized sound» definito in Blacking (1973) è d'obbligo.

La precarietà dell'incontro con l'oggetto: tre motivi

Le indicazioni redatte da Guizzi per la compilazione della scheda catalografica rivelano la consapevolezza che non sempre questo incontro può essere reiterato.

I motivi dell'eventuale o parziale indisponibilità dello strumento ad essere osservato possono essere legati proprio alla sua natura di "cosa materiale" soggetta a un naturale decadimento; è questo il caso di strumenti antichi che si trovano in precario stato di conservazione. Un esempio ancor più significativo riguarda quegli strumenti "temporanei" la cui durata e funzionalità è consapevolmente limitata nel tempo. Si tratta, nei casi più "nobili" della cultura popolare italiana, di strumenti che sono il «risultato di un'estemporanea operazione costruttiva» all'interno di cerimoniali calendariali (il ciclo natalizio e pasquale, il Carnevale ecc.) o determinate circostanze stagionali o occasionali (macellazione del maiale, serenate derisorie ecc.); il loro uso circoscritto dalla ritualità, che comporta lo smontaggio o addirittura la distruzione dello strumento al termine della ricorrenza, amplifica la carica simbolica che tali *oggetti sonori* mostrano di possedere¹⁰.

Tale valore simbolico è condiviso da tutti gli strumenti "effimeri" in generale, anche se non sono circondati dall'aura della ritualità "ufficiale"¹¹; si tratta per la maggior parte dei casi di aerofoni (ma non mancano alcuni cordofoni) che sono accomunati dall'essere

costruiti con materiali naturali prelevati dal mondo vegetale e utilizzati immediatamente, e cioè senza stagionatura o essiccazione, ma anzi sfruttandone l'elasticità e il turgore assicurati dall'attività vegetativa appena interrotta. In tal modo materiali altrimenti non disponibili [...] sono sfruttati per un brevissimo lasso di tempo, quello durante il quale il rapido deterioramento cui tali materiali "verdi" sono sottoposti non ha ancora raggiunto e superato il punto critico della irreversibile degenerazione.¹²

È ovvio che la presenza di tali strumenti pone importanti questioni relative alla loro conservazione, documentazione e catalogazione¹³.

10 «Particolarmente significativi sono gli strumenti effimeri impiegati all'interno di importanti rituali, come è nel caso dei corni di corteccia di castagno costruiti a Ceriana (Im) il Giovedì Santo e utilizzati nei riti delle Tenebre del Venerdì» (Guizzi, 2002, p. 376).

11 Si vedano a questo proposito le considerazioni sugli strumenti del gioco come "residuo di funzioni ataviche" in Guizzi (2002, pp. 350-354).

12 *Ivi*, p. 374. L'uso di questi strumenti è attestato anche in Bonanni (1723).

13 La collezione del Museo del Paesaggio Sonoro di Riva presso Chieri, costituita in gran parte da strumenti appartenenti al cosiddetto "strumentario minore", contiene molti esempi di oggetti "effimeri" fatti di materiali semplici e ampiamente disponibili nel "paesaggio sonoro" rivese. Un "caso limite" è rappresentato dall'aerofono ad ancia quadrupla battente munito di fori digitali, costruito da Dino Tosco praticando due incisioni sul gambo appena raccolto di una pianta di zucchini; i processi di costruzione, rifinitura e uso dello strumento (la cui funzionalità è limitata a qualche decina di minuti) sono documentati da una registrazione audiovisiva realizzata da Guido Raschieri, Cristina Ghirardini e Domenico Torta.

La precarietà dell'incontro con lo strumento può essere dovuta anche alla natura *relazionale* del suo essere *oggetto sonoro*. Ci riferiamo qui alla seconda accezione del termine, che definisce l'*oggetto sonoro* nell'ambito esclusivo del "fare" culturale, le cui tracce possono essere rinvenute anche nell'incontro con le narrazioni che accompagnano l'oggetto¹⁴. La natura soggettiva, mutevole ed evanescente di tali narrazioni pone problemi analoghi a quelli relativi al naturale decadimento dei materiali di cui l'*oggetto sonoro* è fatto. Nel ciclo di vita degli strumenti ci sono occasioni in cui esse possono venire alla luce con maggiore facilità; ciò accade soprattutto nei momenti di rinnovato interesse in cui si intensificano le attenzioni di potenziali narratori verso gli strumenti (per esempio attività di catalogazione, mostre, restauri, acquisizioni e passaggi di proprietà, raccolta durante ricerca su campo, oltre, ovviamente, a tutte le circostanze di reale pratica "musicale" a cui si possa direttamente assistere). Nel caso di strumenti in disuso, soprattutto se appartenenti a culture di tradizione orale, tali informazioni tendono nel tempo a perdere in definizione e affidabilità; è quindi importante che siano perentoriamente registrate.

Infine bisogna constatare che, per motivi di carattere economico e meramente pratico, talvolta la redazione della prima scheda catalografica costituisce l'unica occasione in cui è possibile registrare le osservazioni che scaturiscono dall'incontro tra organologo e strumento.

Per tutte queste ragioni le note per la compilazione della scheda redatte da Guizzi sono così minuziose e ricche di dettagli; attraverso di esse egli non intende solo indicare al catalogatore in quali campi inserire i dati di cui già dispone, ma anche guidarlo in una corretta e meticolosa raccolta dei dati. La Scheda Guizzi sollecita chi la compila

a porsi domande, a valutare caratteristiche materiali e attribuzioni di significato conferite all'oggetto, a contestualizzarlo, sia che si tratti di uno strumento in uso, sia che invece si tratti di un esemplare uscito dall'uso, a indagarne la storia e a riconoscere nelle sue caratteristiche materiali i segni di intervento dell'uomo [...].
(Ghirardini, 2009, p. 242)

A titolo di esempio segnaliamo l'accuratezza con cui è descritto il campo "Misure", il cui contenuto «non può essere predeterminato una volta per tutte» pur essendo annoverato tra i campi «indispensabili per ogni tipo di scheda»¹⁵. Guizzi specifica i criteri generali che devono essere rispettati: le misurazioni devono avere natura descrittiva e non prescrittiva, all'interno della stessa collezione deve essere rispettata l'omogeneità di scelta e attuazione delle misurazioni per strumenti

14 Si tratta di informazioni che identificano l'*oggetto sonoro* in relazione all'attività dell'uomo: fabbricazione, uso, repertori, contesto e cultura di appartenenza ecc.

15 Cfr. *infra* p. 181.

analoghi, e il grado di approssimazione deve essere dichiarato. Sono presenti consigli per affrontare casi particolari¹⁶ e presentare in maniera efficace il risultato delle misurazioni¹⁷.

Altrove Guizzi chiede testualmente al catalogatore di "fare uno sforzo", affinare gli strumenti della ricerca e registrarne i risultati anche quando i dati sembrano indisponibili. Questo accade soprattutto nei campi relativi a "Fabbricazione" e "Uso" dello strumento laddove, nell'eventuale indisponibilità di informazioni, Guizzi suggerisce di riportare la dicitura "ignoto" o "sconosciuto", dichiarando apertamente di avere intrapreso (senza successo) la ricerca dell'informazione. In particolare, in caso di assenza di informazioni sull' "Epoca di fabbricazione", egli dichiara che «tuttavia occorrerebbe fare ogni volta uno sforzo di approssimazione (indicando che il dato è probabile) o di ipotesi (indicando che il dato è presunto)»¹⁸.

Infine le note sulle classificazioni relative a sistematica Hornbostel-Sachs e "Definizione tipologica" si configurano come brevi saggi che illustrano il punto di vista organologico dell'autore¹⁹.

16 «Alcuni dispositivi essenziali dal punto di vista acustico possono richiedere due tipi diversi di misurazione (ad es.: per le corde, "lunghezza totale" e "lunghezza vibrante", intendendo questa come distanza tra i due punti estremi della corda libera, come capotasto e ponticello o simili: lo stesso si può fare per i membranofoni rilevando il diametro totale della membrana — quindi considerando anche la parte che poggia sul bordo del supporto e quello che lo supera per consentirne l'ancoraggio — e il diametro vibrante — misurando nei limiti del possibile la parte libera). [...] Soprattutto per alcune classi di strumenti, in primis per gli idiofoni, è importante rilevare il peso dell'oggetto o di sue parti costitutive essenziali ai fini della produzione del suono (ad es.: le singole piastre o lamine di metallofoni, xilofoni e simili)» (cfr. *infra* p. 182).

17 «È utile disporre le misure di dispositivi seriali, specie se numerosi (come le corde o le piastre di xilofoni, metallofoni, ecc., o le lamelle di sanze e così via) in tabelle che rendano immediatamente confrontabili le variazioni mensurali di ciascuno (quindi in colonne e non in fila indiana entro la stessa riga)» (*ibidem*).

18 Cfr. *infra* p. 182.

19 Per un'analisi delle questioni sollevate da questi campi, centrali nell'attività di catalogazione, si veda il capitolo "Il problema delle *keywords* e della sistematica Hornbostel-Sachs".

La digitalizzazione della scheda

La prima implementazione in XML Schema (versione 1.0)

La versione originale della Scheda Guizzi è un file di testo costituito da una sequenza di campi posti sullo stesso livello gerarchico. Ogni campo è indicato con un titolo a cui segue una nota di testo che ne illustra la compilazione²⁰. La prima digitalizzazione completa della scheda è stata realizzata con la creazione di un XML Schema:

```
scheda guizzi 1_0.xsd
```

Dal file .xsd sono state derivate due rappresentazioni dello schema: una tabella²¹ e una visualizzazione ad albero in grafica vettoriale scalabile²².

Tutti i file .xsd .xml e .svg sono consultabili al seguente indirizzo:

<http://www.suonoemagine.unito.it/bevilacqua.html>

L'implementazione in XML Schema richiede l'elaborazione di una struttura ad albero i cui elementi siano raggruppati in annidamenti; è possibile attribuire vincoli di occorrenza (obbligatorietà e ripetibilità dei campi) a qualsiasi livello gerarchico della struttura. Inoltre le dichiarazioni degli elementi possono essere impostate in modo da consentire, tramite un sistema di rimandi interni, il riuso degli elementi o dei gruppi di elementi più comuni. Per sfruttare al meglio queste e altre possibilità messe a disposizione dal linguaggio XML è necessario rielaborare concettualmente alcuni aspetti della scheda.

La realizzazione della versione 1.0, che rappresenta il punto di partenza di questo percorso, è stata rivolta alla normalizzazione dei campi al fine di creare annidamenti omogenei e porre le basi per successive e più complesse articolazioni. I campi sono stati ripensati anche compatibilmente alle nuove funzioni attribuite alla scheda catalografica nel contesto digitale: i dati catalogati sono infatti destinati non solo all'uso interno al museo ma anche alla pubblicazione online e all'integrazione con altre tecnologie.

In riferimento alla Scheda Guizzi originale, la normalizzazione ha riguardato lievi ma significative modifiche dell'ordine dei campi; alcuni di essi sono stati eliminati, accorpati o annidati e ne sono

²⁰ La scheda originale è pubblicata integralmente in Appendice 1.

²¹ La tabella è pubblicata integralmente in Appendice 2.

²² I file in grafica vettoriale, in formato .svg (Scalable Vector Graphics), sono stati creati con l'applicazione *open source* XSDVi (<https://sourceforge.net/projects/xsdvi/>).

stati inseriti di nuovi. Le note sono state aggiornate contestualmente alla modifica della struttura generale e in relazione alle nuove funzioni della scheda.

Come illustrato in fig.1 tutti gli elementi dello schema XML sono dichiarati localmente, in una sequenza all'interno dell'elemento `record` (l'elemento che contiene tutte le informazioni della scheda)²³; a ciascun elemento sono associati gli attributi di occorrenza²⁴.

```

24 ☐ <xs:element name="record" maxOccurs="unbounded" minOccurs="0">
25 ☐   <xs:complexType>
26 ☐     <xs:sequence>
27 ☐       <xs:element name="scheda">
28 ☐       <xs:element name="collocazione">
37 ☐       <xs:element name="inventarioGenerale">
47 ☐       <xs:element name="inventarioParticolare" minOccurs="0">
57 ☐       <xs:element name="inventarioTopografico" minOccurs="0">
67 ☐       <xs:element name="acquisizione">
77 ☐       <xs:element name="collezione" minOccurs="0">
90 ☐       <xs:element name="strumento">
100 ☐       <xs:element name="taxonHS">
110 ☐       <xs:element name="numeroHS">
125 ☐       <xs:element name="definizioneTipologica">
138 ☐       <xs:element name="nomeOriginario" maxOccurs="unbounded" minOccurs="0">
154 ☐       <xs:element name="fabbricazione" maxOccurs="unbounded">
164 ☐       <xs:element name="uso" maxOccurs="unbounded">
203 ☐       <xs:element name="cultura" minOccurs="0">
243 ☐       <xs:element name="cartiglioliscrizioni">
253 ☐       <xs:element name="descrizione">
265 ☐       <xs:element name="statoDiConservazione">
276 ☐       <xs:element name="restauri" maxOccurs="unbounded">
286 ☐       <xs:element name="materiali">
296 ☐       <xs:element name="misure">
306 ☐       <xs:element name="annotazioni" minOccurs="0">
316 ☐       <xs:element name="esposizioni" maxOccurs="unbounded" minOccurs="0">
326 ☐       <xs:element name="bibliografia">
336 ☐       <xs:element name="risorse">
346 ☐       <xs:element name="compilatoreDellaScheda">
356 ☐       <xs:element name="aggiornamenti" maxOccurs="unbounded" minOccurs="0">
395 ☐     </xs:sequence>
434 |   </xs:complexType>
435 | </xs:element>
436 |

```

fig.1 scheda guizzi 1_0.xsd (estratto)

23 Nel file XML gli elementi contrassegnati dal segno "+" sono "collassati" e sono espandibili cliccandoci sopra.

24 Ci si riferisce ai vincoli di obligatorietà e ripetibilità definiti dagli attributi `maxOccurs` e `minOccurs`. I valori di default sono: `maxOccurs="1"` e `minOccurs="1"`. Ciò vuol dire che, se non diversamente segnalato, tutti gli elementi sono obligatori e non ripetibili.

Ogni campo (cfr. fig.2) contiene un elemento `documentation` che riporta le note redatte da Guizzi: in questo modo l'XML Schema include le istruzioni per il suo utilizzo. Sono inoltre stati impostati i valori massimi di lunghezza dei campi²⁵.

```

27 <xs:element name="scheda">
28 <xs:annotation>
29 <xs:documentation>Si indica qui il numero progressivo della
   scheda all'interno del lavoro che si sta svolgendo. A ciascun
   strumento deve corrispondere una scheda, salvo in casi di
   dispositivi sonori pervenuti insieme e necessari l'uno all'altro. Se
   per esempio uno strumento ad ancia è stato raccolto insieme
   all'ancia con cui è stato suonato, è utile catalogare l'ancia e lo
   strumento insieme e fare schede separate per eventuali altre ancie
   pervenute.</xs:documentation>
30 </xs:annotation>
31 <xs:simpleType>
32 <xs:restriction base="xs:string">
33 <xs:maxLength value="20"/>
34 </xs:restriction>
35 </xs:simpleType>
36 </xs:element>

```

fig.2 scheda guizzi 1_0.xsd (estratto)

Espansione dell'elemento scheda

Abbiamo modificato la nota del campo `scheda` escludendo la possibilità di usare la stessa scheda per catalogare gruppi di strumenti affini²⁶. Tale possibilità è stata introdotta da Guizzi per limitare la proliferazione di documenti in un contesto di catalogazione di derivazione “cartacea”; l’adozione di tecnologie digitali invece, consente la produzione di un elevato numero di documenti, anche seriali o ridondanti, gestibili in maniera efficace ai fini inventariali.

Abbiamo eliminato dalla nota di `acquisizione` il suggerimento di registrare la documentazione contabile-amministrativa (non sono informazioni adatte alla pubblicazione online), e dal campo `misure` abbiamo escluso la possibilità di organizzare i risultati delle misurazioni in tabelle (non ancora implementate nella versione 1.0)²⁷. La voce `categoria` è stata completamente rimossa²⁸.

²⁵ I valori di lunghezza massima non sono inclusi nella tabella pubblicata in Appendice 2: essi non sono ancora definitivi e sono pertanto suscettibili di variazione.

²⁶ Non rientrano in questo caso i dispositivi (per es. le ancie) pervenuti con lo strumento del quale possono essere considerati parte costitutiva nella produzione fonatoria.

²⁷ Sulla necessità di allegare tabelle delle misurazioni si veda il campo “Misure” della Scheda Guizzi (cfr. *infra* p. 180) e il cap. “Linee-guida per successive implementazioni” (cfr. *infra* p.182).

²⁸ Era utilizzata per distinguere gli strumenti musicali da altri tipi di oggetti catalogati con la stessa scheda (ad es. utensili per la costruzione dello strumento). La catalogazione digitale, soprattutto quella rivolta al *semantic web*, mette a disposizione una “libreria” di modelli integrabili per la descrizione di qualsiasi tipo di oggetto. Questa voce non è quindi più necessaria.

Il campo `inventarioTopografico` è stato spostato in prossimità di altri campi contenenti informazioni inventariali²⁹.

Gli elementi `costruttore`, `dataEpocaDiFabbricazione` e `luogoDiFabbricazione` sono raggruppati in un annidamento sotto il nuovo elemento `fabbricazione`. La stessa cosa avviene per le informazioni contenute nell'elemento `uso`; in quest'ultimo caso, per analogia con `fabbricazione` e in vista di una mappatura dei dati nel modello evento-centrico di LIDO, è stato aggiunto un nuovo campo `suonatore` che consente di inserire informazioni sugli "attori" dell'evento³⁰.

```

164 | <xs:element name="fabbricazione" maxOccurs="unbounded">
165 |   <xs:annotation>
168 |   <xs:complexType>
169 |     <xs:sequence>
170 |       <xs:element name="costruttore">
180 |       <xs:element name="luogoDiFabbricazione">
190 |       <xs:element name="dataEpocaDiFabbricazione">
200 |     </xs:sequence>
201 |   </xs:complexType>
202 | </xs:element>
203 | <xs:element name="uso" maxOccurs="unbounded">
204 |   <xs:annotation>
208 |   <xs:complexType>
209 |     <xs:sequence>
210 |       <xs:element name="suonatore" maxOccurs="unbounded" minOccurs="0">
220 |       <xs:element name="luogoUso" maxOccurs="unbounded">
230 |       <xs:element name="dataEpocaUso" maxOccurs="unbounded">
240 |     </xs:sequence>
241 |   </xs:complexType>
242 | </xs:element>

```

fig.3 scheda `quizzi 1_0.xsd` (estratto)

Espansione degli elementi `fabbricazione` e `uso`

Il campo `esposizioni` è stato spostato in modo da consentire l'adiacenza di `bibliografia` e `risorse`; quest'ultimo è un campo nuovo nel quale confluiscono tutte le informazioni originariamente contenute in "documentazione", "archivio fotografico" e "allegati".

²⁹ Il raggruppamento di tutte le informazioni inventariali è stato utile per la creazione dell'annidamento `identificazioneSet` della versione 1.1 *draft*.

³⁰ Abbiamo scelto "suonatore" e non "musicista" perché riteniamo che l'organologia debba avere per oggetto del suo studio ogni *oggetto sonoro* volutamente utilizzato in «qualunque attività umana in cui la produzione di suoni sia perseguita o percepita in forma dotata di senso» (Guizzi, 2002, p. 335). Rientrano in questo caso numerosi oggetti estranei a contesti socialmente riconosciuti come "musicali", e di conseguenza non può essere attribuito il ruolo di "musicista" a chi attiva sonoramente lo strumento. D'altra parte i criteri di definizione dello *status* di "musicista" sono molto diversi a seconda della cultura, e non possono essere descritti una volta per tutte. A questo proposito si veda la rassegna della letteratura etnomusicologica sull'argomento commentata in Merriam (2007, pp. 133-152).

I campi della scheda originale “compilatore della scheda”, “data” e “aggiornamenti” sono riuniti in due nuovi annidamenti che restituiscono descrizioni più granulari: `compilatoreDellaScheda` contiene elementi relativi ad autore, data e diritti di proprietà intellettuale e licenze d'uso della scheda; `aggiornamenti` contiene anch'esso elementi relativi ad autore e data dell'aggiornamento; il suo sotto-elemento `commenti` descrive quali aggiornamenti sono stati fatti e in quali campi, in modo che si possa tenerne traccia.

Sono state compilate *ex novo* le note degli elementi nuovi, mentre quelle degli elementi oggetto di annidamento o accorpamento sono state adattate o riformulate.

Alla nuova struttura sono stati assegnati gli attributi di occorrenza³¹.

L'attributo di ripetibilità è tanto più efficace quanto più lo schema è organizzato in annidamenti. Nel caso dell'elemento `uso` abbiamo consentito la ripetizione sia dell'elemento genitore che di tutti i suoi sotto-elementi; in questo modo sono possibili diverse configurazioni in cui i sotto-elementi sono efficacemente messi in relazione dall'elemento genitore `uso` (cfr. fig.4). Lo stesso discorso vale per `<fabbricazione>` e `<aggiornamenti>`, dove abbiamo attribuito la ripetibilità a diversi livelli dell'annidamento.

```

22 □ <uso>
23 | <suonatore>suonatore 1</suonatore>
24 | <luogoUso>torino</luogoUso>
25 | <dataEpocaUso>29/08/1980</dataEpocaUso>
26 | </uso>
27 □ <uso>
28 | <suonatore>suonatore 2</suonatore>
29 | <suonatore>suonatore 3</suonatore>
30 | <luogoUso>Rive presso Chieri</luogoUso>
31 | <dataEpocaUso>Anni '80 del XX secolo</dataEpocaUso>
32 | </uso>
33 □ <uso>
34 | <suonatore>suonatore 4</suonatore>
35 | <luogoUso>Piemonte</luogoUso>
36 | <dataEpocaUso>10/08/2014</dataEpocaUso>
37 | <dataEpocaUso>12/08/2016</dataEpocaUso>
38 | <dataEpocaUso>dal 2014 al 2016</dataEpocaUso>
39 | </uso>

```

fig.4 scheda `quizzi 1_0.xsd` (estratto)

Esempio di possibili annidamenti dell'elemento `uso`

Invece, nel caso di `nomeOriginario`, `restauri`, `esposizioni` e dei sotto-campi di `compilatoreDellaScheda`, la ripetibilità è attribuita ad un solo livello di annidamento, essendo

31 Per una loro consultazione si rimanda alla tabella pubblicata in Appendice 2.

il contenuto dei campi un testo libero. In questo caso ogni campo ripetuto serve solamente a individuare le singole occorrenze di uno stesso concetto all'interno di una lista, e non a disporre in una relazione sintattica un gruppo di sotto-elementi (cfr. fig.5).

```

32 <misure>Diametro minimo dei bottoni: 24,2 mm
33 Diametro massimo dei bottoni: 27,8 mm
34 Lunghezza totale del filo: 960 mm </misure>
35 <annotazioni>Costruito dalla nonna, Giuseppina Tamagnone, per la nipote Anna Torta, come giocattolo sonoro che produce suono
   < tendendo il filo e facendo scorrere i bottoni. Successivamente Anna ha imparato a contare con questo strumento.</annotazioni>
36 <esposizioni>Agricoltore Artigiano Artista, Riva presso Chieri, Palazzo Grosso, 16-30 aprile 1995</esposizioni>
37 <esposizioni>Seconda edizione del Festival Chierinmusica, Chieri, 1995</esposizioni>
38 <esposizioni>X Rescontre Occitan, Limone Piemonte, in collaborazione con Roberto Tombesi, 1996</esposizioni>
39 <esposizioni>Identità e differenza, Torino, Sermig, 1996</esposizioni>
40 <esposizioni>Virtuosi senza pretese, Entracque, Centro Visite del Parco, 24 luglio – 5 settembre 2004</esposizioni>
41 <esposizioni>Virtuosi senza pretese, Rovigo, 2004</esposizioni>
42 <bibliografia>Guizzi, Febo (2002). Gli strumenti della musica popolare in Italia. Lucca: Lim Editrice.</bibliografia>
43 <risorse>0202_SM.tiff</risorse>
44 <compilatoreDellaScheda>
45 <autoreCompilazione>Cristina Ghirardini</autoreCompilazione>
46 <dataCompilazione>Aprile 2007</dataCompilazione>
47 <dirittiScheda>Copyright, Civico Museo del Paesaggio Sonoro, 2007. This work is licensed under the Creative Commons Attribution 4.0
   < International License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/ or send a letter to Creative
   < Commons, PO Box 1866, Mountain View, CA 94042, USA</dirittiScheda>
48 </compilatoreDellaScheda>

```

fig.5 sonaglio_per_bambini_2_guizzi.xml (estratto)

Riguardo l'attribuzione delle obbligarietà, la scelta di definire obbligatorio un numero elevato di elementi, e quindi vincolare la validità del record XML all'inserimento di numerose informazioni, vuole rendere effettivo quell'intento di "guida" al reperimento dei dati espresso da Guizzi nelle note della scheda.

La scheda deve obbligatoriamente fornire informazioni relative ad acquisizione, classificazione, fabbricazione, uso, descrizione, stato di conservazione, materiali e misure dello strumento. Per quei campi obbligatori, come quelli relativi a fabbricazione ed uso, per i quali c'è la possibilità che i dati non siano disponibili, abbiamo escluso la possibilità di lasciare il campo vuoto, suggerendo l'inserimento di diciture standard come "assenti" e "non rilevabili".

Abbiamo attribuito l'obbligatorietà anche agli elementi cartiglioIscrizioni e restauri³², bibliografia³³, risorse³⁴ e compilatoreDellaScheda.

Il catalogatore è così sollecitato a indagare molteplici aspetti dello strumento, a registrare le informazioni disponibili, a raffinare la ricerca in caso di indisponibilità dei dati e a collocare lo strumento in un quadro concettuale e disciplinare documentato. Le schede compilate secondo questo modello sono ricche di informazioni raccolte e organizzate con precisione, e riferite a classificazioni e bibliografie scientifiche. Questa è secondo noi la competenza organologica che può essere integrata con le tecnologie del *web semantico*.

32 È molto probabile, per esempio, che gli strumenti, soprattutto quelli più antichi, siano stati sottoposti a interventi di restauro o segnati in vario modo da iscrizioni: Guizzi chiede di cercarne le tracce portate dallo strumento.

33 La bibliografia necessita di un campo a sé stante distinto da quello più generale delle risorse: essa è fondamentale perché induce il catalogatore a condividere i documenti consultati per la compilazione di quei campi che richiedono un'interpretazione dello strumento; in questo modo anche il consultatore della scheda avrà la possibilità di estendere la propria competenza organologica a riguardo.

34 È necessario allegare almeno una immagine recente e rappresentativa dello strumento, possibilmente acquisita in fase di catalogazione

La mappatura Guizzi-MIMO

Il lavoro di confronto tra la Scheda Guizzi versione 1.0 e il modello di organizzazione dei dati di MIMO ha portato alla stesura delle “Istruzioni per una mappatura Guizzi-MIMO” pubblicate in Appendice 3. MIMO, punto di riferimento tecnologico per la catalogazione degli strumenti musicali rivolta al *semantic web*, adotta lo schema LIDO e struttura le informazioni in campi molto granulari. La Scheda Guizzi invece, anche nella sua implementazione in XML Schema versione 1.0, presenta l'informazione sotto forma di testo libero; di conseguenza nella mappatura abbiamo utilizzato molti campi *display* predisposti da LIDO proprio per le informazioni strutturate in questo modo. In alcuni casi non è stato possibile individuare una corrispondenza univoca tra i due modelli; di conseguenza il testo libero dei campi Guizzi è stato diviso e destinato a più campi LIDO³⁵. Altrove abbiamo estrapolato dal testo libero le informazioni atomizzate e le abbiamo destinate a descrizioni LIDO granulari da affiancare all'elemento *display*³⁶.

La mappatura è avvenuta secondo l'impostazione evento-centrica di LIDO. Per alcuni campi della Scheda Guizzi già strutturati ad “evento”, come *fabbricazione* e *uso*, il processo è avvenuto in modo naturale; altrove invece si è percepita una forzatura concettuale, come nel caso di *materiali* e *cultura* che abbiamo rispettivamente destinato a eventi di tipo `lido:type="production"`³⁷ e `lido:type="use"`³⁸. Più in generale bisogna entrare nell'ottica che molte delle informazioni atomizzate possono essere inserite in un record LIDO solo in una strutturazione ad eventi.

35 Per esempio nel caso delle *annotazioni*, elemento contenente informazioni di natura molto varia, si è scelto di destinare tali informazioni a diversi elementi `lido:objectDescriptionSet` con attributi `lido:type` differenti: “general description”, “context”, “decorative features” e “performance characteristics”.

36 Questo è avvenuto in modo spontaneo per quei campi in cui il testo libero è costituito dalla concatenazione di informazioni di genere diverso; si veda il modo in cui sono state mappate le *esposizioni* nell’ “esempio di record MIMO” riportati in Appendice 5.

37 Ciò non esclude che si possano inserire informazioni sui materiali in *ogni* tipo di evento. Anzi, una descrizione rigorosa dovrebbe risalire a tutti gli interventi che hanno apportato modifiche ai materiali che costituiscono l'oggetto, e descriverli in “eventi” di tipo restaurativo; un lavoro del genere diventa un’impresa proibitiva se svolto per strumenti di liuteria antica, che sono stati sottoposti a molteplici e continui interventi, anche lievi, di manutenzione. Inoltre bisognerebbe almeno tenere in considerazione che, come sostiene Guizzi (2002), ci troviamo di fronte anche a un problema di natura ontologica, dal momento che «suonare uno strumento significa sottoporlo ogni volta a un ciclo poetico che influenza la sua condizione materiale» (p. XXIV).

38 Anche nel caso di `lido:culture` l'elemento può essere definito in qualsiasi tipo di “evento”; si possono descrivere per esempio casi di strumenti fabbricati in paesi di cultura diversa da quella in cui sono stabilmente utilizzati. `lido:culture` può essere espresso però solo da un elemento `lido:term` (che indica un termine appartenente a una lista controllata) e non da un elemento *display*; ciò non facilita la descrizione di situazioni che possono essere di natura complessa.

In merito alle misurazioni abbiamo deciso di usare il campo *display*, anche se LIDO consente di fornire descrizioni più granulari.

Siccome in LIDO non ci sono campi analoghi a `collezione` e `statoDiConservazione` (presenti nella Scheda Guizzi), abbiamo inserito queste informazioni nella descrizione generale; lo stesso problema si è riscontrato con `compilatoreDellaScheda` e `aggiornamenti`: LIDO richiede la compilazione obbligatoria dell'elemento `lido:recordSource`, che è destinato principalmente ad identificare l'istituzione che fornisce i dati, non la persona che compila la scheda. Tutte le informazioni su autori e aggiornamenti sono state perciò inserite in `lido:recordMetadataDate`, un campo di testo libero appartenente all'elemento `lido:recordInfoSet`³⁹.

Infine, LIDO prevede un ampio apparato di elementi per la descrizione delle risorse e della loro rappresentazione. Purtroppo MIMO prevede che le risorse siano solo di tipo fotografico o audiovisivo; la bibliografia, le didascalie e altri testi documentari vanno inserite in un elemento di testo libero all'interno di `lido:objectDescriptionSet` e non possono essere identificate univocamente né collegate in maniera ipertestuale o in modalità LOD a database esterni.

39 `recordInfoSet` consente anche di indicare il link al record nel dominio locale del *data provider*.

Linee-guida per successive implementazioni

La versione 1.0 dell'XML Schema rappresenta il primo passo verso una digitalizzazione della Scheda Guizzi compatibile con le tecnologie del *semantic web*. I successivi sviluppi devono orientarsi verso una maggiore articolazione degli annidamenti finalizzata all'atomizzazione dei dati: una strutturazione fine del livello sintattico consentirebbe alla scheda di rendersi autonoma nella predisposizione dei dati al collegamento semantico in modalità LOD.

Nell'elaborazione della versione 1.1 *draft*⁴⁰ abbiamo lavorato in questa direzione: tutti gli elementi della versione 1.0 sono annidati all'interno di 6 Set principali (cfr. fig.5a e fig5b) che descrivono un percorso coerente attraverso il quale il catalogatore è indotto ad osservare lo strumento adottando prospettive e applicando competenze diverse:

- inizialmente lo strumento è idealmente collocato all'interno del museo: `identificazioneSet` contiene i campi relativi a collocazione fisica, numeri di inventario, notizie sull'acquisizione e al nome con cui lo strumento è univocamente identificato nel museo;
- in `classificazioneSet` è richiesta una specifica competenza organologica che interviene a interpretare lo strumento e a collocarlo all'interno delle classificazioni (Sistematica Hornbostel-Sachs e Definizione tipologica). Esse permettono al museo di organizzare praticamente insiemi di strumenti affini: per questo motivo il Set è posto in posizione intermedia tra una prima parte "amministrativa" e quelle seguenti che esaminano lo strumento più nel dettaglio⁴¹;
- `contestoSet` raccoglie le informazioni relative a fabbricazione, uso e cultura di appartenenza. In questi campi si inseriscono osservazioni dirette o, più probabilmente, narrazioni sottoposte a verifica; nel caso in cui non si disponga di questi dati e si elaborino ipotesi la competenza organologica sostiene il lavoro dello storico;
- `descrizioneSet` contiene tutte le informazioni relative alla descrizione dello strumento: iscrizioni, descrizione generale e di aspetti particolari, stato di conservazione, restauri, materiali e misure;
- infine `risorseSet` e `datiAmministrativi` raccolgono rispettivamente documentazione e risorse relative allo strumento, e informazioni su autori e date di compilazione della scheda.

40 Il file `scheda guizzi 1_1 draft.xsd` è pubblicato, insieme alle sue rappresentazioni in grafica vettoriale (cfr. fig.5a e fig.5b) all'indirizzo: <http://www.suonoemagine.unito.it/bevilacqua.html>

41 A rigor di logica, il lavoro di classificazione è possibile solo *dopo* l'analisi dettagliata dello strumento.

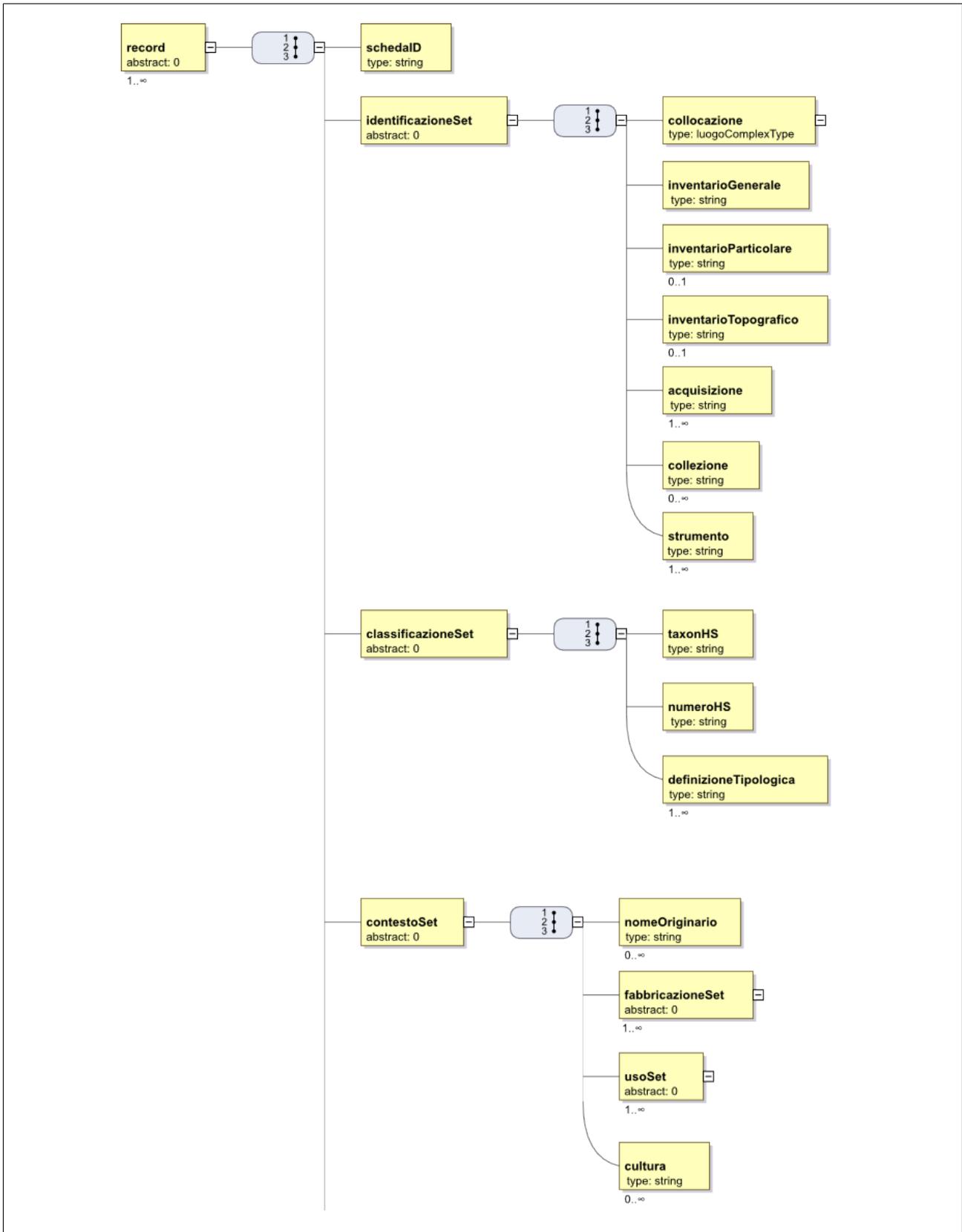


fig.5a Scheda Guizzi 1.1 *draft* – schema generale 1/2

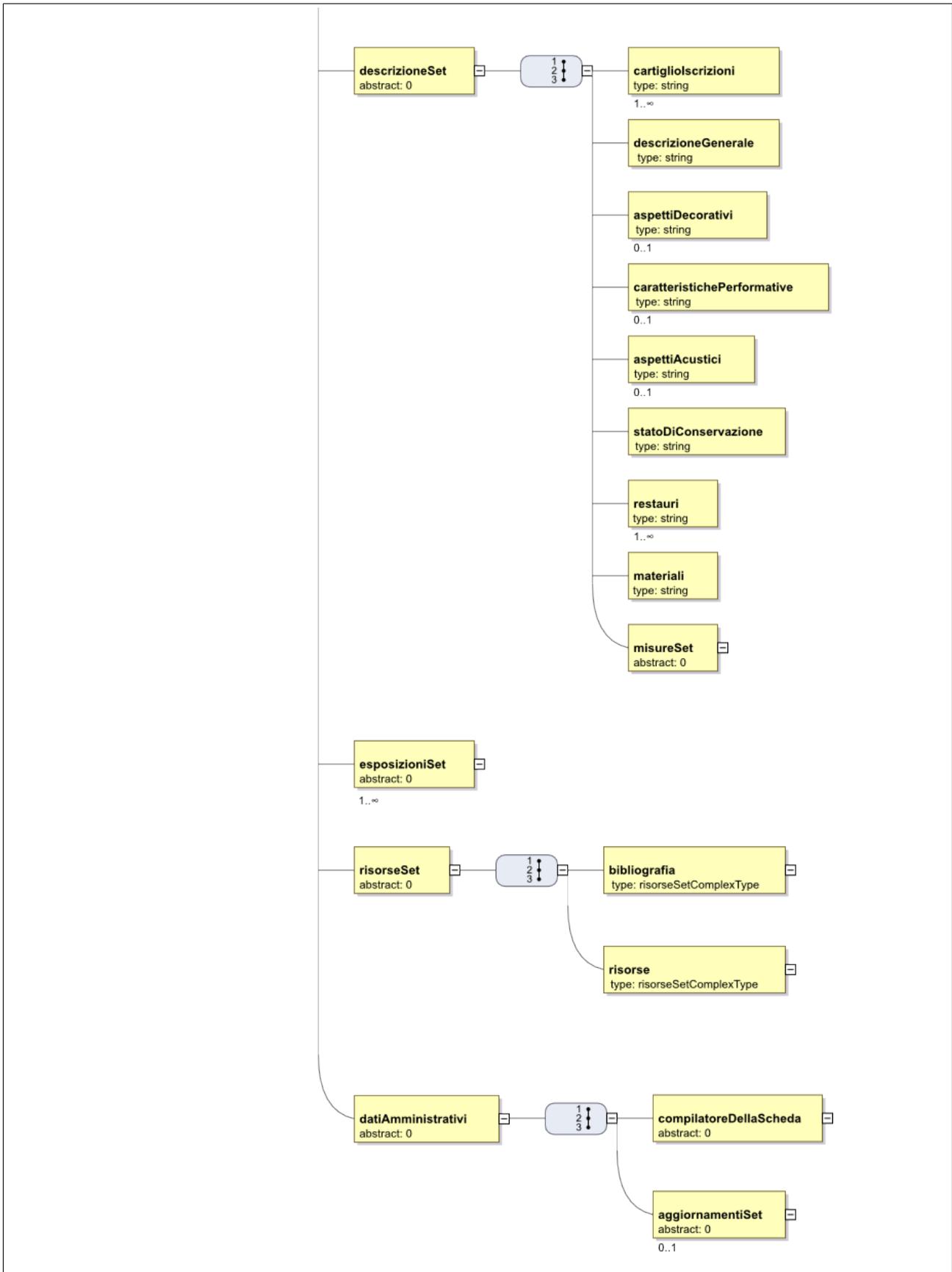


fig.5b Scheda Guizzi 1.1 *draft* – schema generale 2/2

La fig.6 mostra un esempio di articolazione interna di un elemento complesso: sono presenti sia descrizioni granulari che elementi *display* contenenti testo libero; gli annidamenti sono regolamentati attraverso un meccanismo di relazione simile a quello degli elementi *Wrap* e *Set* usati in LIDO; per agevolare il processo di mappatura in LIDO le informazioni relative al concetto di “evento” sono incapsulate all’interno dei campi previsti da Guizzi, il cui ordine è preservato (cfr. fig.5), e le date possono essere registrate come *time-span*.

In questo modo la struttura della scheda conserva in superficie il riferimento al contenuto organologico dei campi mentre le informazioni relative al concetto “generale” di evento sono contenute all’interno⁴².

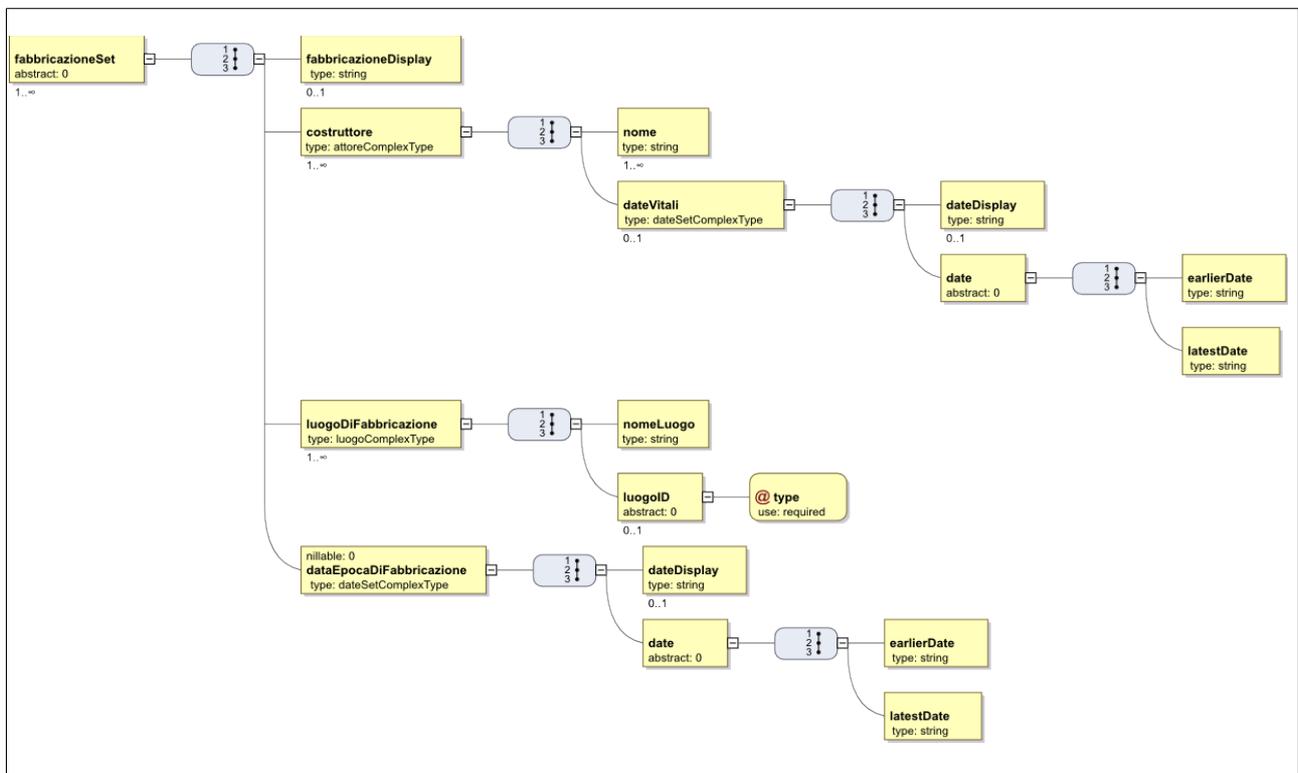


fig.6 Scheda Guizzi 1.1 *draft* – Struttura dell’elemento *fabbricazioneSet*

42 In LIDO accade esattamente il contrario: il concetto “generale” di “evento” è in superficie mentre il tipo “particolare” di evento è definito solo all’interno della struttura del campo stesso.

Nella mappatura Guizzi-MIMO abbiamo sottolineato che il modello ontologico evento-centrico alla base di LIDO impone alcune forzature concettuali nella rappresentazione delle informazioni che ci interessano. Il caso più evidente riguarda il campo `materiali` che nel nostro modello ha una propria autonomia ontologica, direttamente connessa alla primarietà che Guizzi attribuisce all'*incontro* tra organologo e oggetto.

Circoscrivere, come fa LIDO, la descrizione dei materiali all'ambito della descrizione degli "eventi" comporta uno slittamento di significati perché si confondono le informazioni connesse all'osservazione diretta dell'oggetto (così come *esso si presenta* all'osservatore) con quelle riportate da narrazioni di eventi che hanno coinvolto l'oggetto⁴³.

Se il modello evento-centrico non è il più diffuso nella rappresentazione dei dati appartenenti al dominio del patrimonio culturale, e altre iniziative che implementano tecnologie di trasferimento dei dati, come OAI-PMH ed Europea, richiedono che i *data provider* garantiscano che almeno un nucleo dei dati forniti siano compatibili con il modello oggetto-centrico (in entrambi i casi rappresentato dal formato *Dubin Core*), sarà opportuno prendere in considerazione questo modello anche per lo sviluppo della Scheda Guizzi. In particolare bisognerà verificare che il lavoro di atomizzazione dei dati, finora concepito in vista della mappatura verso un modello evento-centrico, non crei problemi di incompatibilità nei confronti della mappatura verso modelli oggetto-centrici.

43 Anche l'analisi dello strumento può essere considerata un "evento" del quale si riportano narrazioni. Come abbiamo visto, l'ontologia CIDOC-CRM prevede una specifica classe E13 Attribute Assignment mediante la quale è possibile registrare alcuni "eventi" connessi all'attività di conservazione e catalogazione degli oggetti; tra questi non è compreso il riconoscimento dei materiali che costituiscono l'oggetto. Il CRM descrive l'entità E57 Material come una sottoclasse di E55 Type che è a sua volta sottoclasse di E28 Conceptual Object. Anche se può prendere parte ad "eventi" come E11 Modification e E12 Production, E57 Material è considerata un'entità persistente direttamente collegata alla cosa fisica, infatti:

Physical Thing (E18) *consists of* (P45) Material (E57)

Di conseguenza la connessione esclusiva del concetto di "materiali" con quello di "evento" sembra una caratteristica distintiva che LIDO non deriva dal CRM.

Il campo relativo alle misurazioni dello strumento musicale pone questioni particolari: data l'imprevedibilità delle misure, dovuta all'eterogeneità degli oggetti e dei criteri di volta in volta adottati per misurarli, non è possibile predeterminarne la struttura. Per questo motivo nella mappatura Guizzi-MIMO abbiamo deciso di utilizzare solamente l'elemento *display*.

In realtà LIDO dispone di ulteriori strumenti che descrivono le misure in modo più granulare; oltre agli elementi usati da MIMO, che indicano tipo, unità di misura e valore della misurazione, l'elemento `lido:objectMeasurements` contiene i seguenti elementi:

- `lido:extentMeasurements` che indica la parte dell'oggetto descritta;
- `lido:qualifierMeasurements` che descrive la qualità della misurazione, indicandone per esempio l'approssimazione.

Questi strumenti sono utili per la descrizione di misurazioni semplici, in cui pochi elementi sono sintatticamente collegati all'interno di una struttura predefinita. Nel caso di misurazioni più complesse essi si dimostrano meno adeguati; per descrivere l'esempio della "müsa" riportato in fig.8 LIDO deve ricorrere a lunghe stringhe di caratteri, più vicine all'idea di "testo libero" che a quella di atomizzazione dei dati (cfr. fig.7).

```

5 | <lido:objectMeasurementsSet>
6 |   <lido:objectMeasurements>
7 |     <lido:measurementsSet>
8 |       <lido:measurementType xml:lang="it">Distanza dall'estremità prossimale al centro del I foro</lido:measurementType>
9 |       <lido:measurementUnit>mm</lido:measurementUnit>
10 |      <lido:measurementValue>58</lido:measurementValue>
11 |     </lido:measurementsSet>
12 |     <lido:measurementsSet>
13 |       <lido:measurementType xml:lang="it">Distanza dall'estremità prossimale al centro del II foro</lido:measurementType>
14 |       <lido:measurementUnit>mm</lido:measurementUnit>
15 |      <lido:measurementValue>85</lido:measurementValue>
16 |     </lido:measurementsSet>
17 |     <lido:extentMeasurements xml:lang="it">Chanter</lido:extentMeasurements>
18 |   </lido:objectMeasurements>
19 | </lido:objectMeasurementsSet>

```

fig.7 Esempio di descrizione delle misurazioni in LIDO

A questo punto bisogna considerare che quello delle misurazioni non è un campo particolarmente significativo nel processo di indicizzazione dei dati per i motori di ricerca: l'atomizzazione dei dati delle misurazioni e la loro conseguente interpretazione non sono, per ora, fondamentali nel contesto del *semantic web*.

Le misurazioni sono invece molto importanti nel lavoro di ricerca organologica. Sarà quindi opportuno che il catalogatore possa disporre di strumenti più duttili per la loro registrazione; in particolare bisognerà poter allegare al record delle tabelle XML, la cui strutturazione deve essere libera da vincoli dovuti a percorsi di semantizzazione.

La struttura della scheda dovrà inoltre disporre di campi identificativi di concetti: si potrà qui seguire l'esempio di LIDO che riutilizza in numerosi punti dello schema il tipo complesso `lido:conceptComplexType`, che può essere espresso da un `lido:conceptID` e da un `lido:term`.

LIDO utilizza anche altri strumenti che sarebbe utile integrare nello sviluppo della Scheda Guizzi; si tratta principalmente di attributi, anch'essi utilizzabili diffusamente nello schema:

- `xml:lang` per il riconoscimento del linguaggio;
- `lido:pref` per stabilire il livello di preferenza di dati analoghi;
- `lido:type` per collegare i concetti ai vocabolari LOD di pertinenza.

In questo modo la scheda predisporrebbe una struttura sintattica capace di collocare autonomamente i propri dati nel contesto delle tecnologie del *semantic web*.

La compilazione dei nuovi record e la conversione delle schede catalografiche già redatte in formato testuale si dovranno svolgere mediante l'uso di un'interfaccia (DBMS, DataBase Management System) che agevoli e ottimizzi il lavoro del catalogatore. La compilazione delle schede catalografiche tramite il DBMS (configurato sulla base dell'XML Schema) genera automaticamente dei record XML contenenti dati atomizzati, la cui mappatura verso nuovi formati può avvenire mediante procedimenti anch'essi automatici

Chanter

Lunghezza totale: 313.

Lunghezza del tenone: 25.

Distanza dall'estremità prossimale al centro	del I foro	58
	del II foro	85
	del III foro	112
	del IV foro	138
	del V foro	164
	del VI foro	191
	del VII foro	215
	dei fori di intonazione laterali	237
	del foro di intonazione anteriore	247
	dei fori di sfiato	266

Diametro all'estremità prossimale	interno:	9
	esterno:	20

La lunghezza dell'alloggiamento dell'ancia all'interno del tenone non è misurabile

Diametro del rigonfiamento all'estremità prossimale: 30

Diametro esterno della canna all'altezza	del I foro	19,5
	del VII foro	23
	dei fori di sfiato	25

Diametro all'estremità distale	interno	20
	esterno	53

Diametro	del I foro	5,7
	del II foro	6,5
	del III foro	6,9
	del IV foro	7,4
	del V foro	8,5
	del VI foro	8,7
	del VII foro	8,5
	dei fori di intonazione laterali	5
	del foro di intonazione anteriore	5,3
	dei fori di sfiato	8,5

Bordone

Lunghezza totale: 405

Distanza dall'estremità prossimale	alla prima coppia di fori	292
	alla seconda coppia di fori	323
	alla terza coppia di fori	343

Diametro	della prima coppia di fori	5,7; 6
	della seconda coppia di fori	6; 6
	della terza coppia di fori	6 ca (entrambi i fori sono tappati con cera)

Segmento iniziale

Lunghezza: 232

Lunghezza del tenone all'estremità prossimale: 26

Lunghezza del tenone all'estremità distale: 65

Diametro all'estremità prossimale (tenone)	interno	11,5
	esterno	23,5
Diametro all'estremità distale (tenone)	interno	14
	esterno	19

Segmento terminale

Lunghezza: 237

Diametro all'estremità prossimale (mortasa)	interno	22
	esterno	32,5
Diametro all'estremità distale	interno	9,5
	esterno	31,5 (diametro massimo del padiglione)

Blocco

Lunghezza totale: 113

Diametro all'estremità prossimale	interno	28
	esterno	46 (minimo), 47 (massimo)
Diametro all'estremità distale	interno	22
	esterno	35

fig.8 Scheda 232 del Museo Guatelli di Ozzano Taro - Ghirardini (2009, pp. 252-253)

Il problema delle *keywords* e della sistematica Hornbostel-Sachs

La funzione classificatoria è uno dei compiti primari dell'attività di catalogazione ed il suo valore è amplificato nel contesto tecnologico del *semantic web*; le istituzioni della memoria sono infatti chiamate a elaborare le forme di strutturazione dei dati, fornire quantità di dati strutturati e, a un livello più alto dell'architettura, redigere i vocabolari attraverso i quali i concetti fondamentali del dominio di pertinenza sono rappresentati.

Nell'ambito degli strumenti musicali MIMO ha elaborato i vocabolari delle *keywords* e della sistematica HS secondo le tecnologie LOD: la loro struttura gerarchica e associativa è definita tramite il linguaggio SKOS; ogni concetto (`skos:Concept`) è identificato da un HTTP URI.

Dal punto di vista scientifico, l'adozione della sistematica Hornbostel-Sachs (HS) per definire una struttura gerarchica di URI comporta alcuni aspetti critici. Nella concezione degli autori essa è intesa come una gerarchia ad albero di categorie suddivise sulla base di un principio analogo a quello della classificazione decimale Dewey; i concetti espressi dai numeri Hornbostel-Sachs non sono unità a sé stanti, ma possono essere combinati tra di loro per descrivere oggetti complessi; o, in casi molto frequenti, può essere loro apposto un suffisso che esprime un concetto applicabile in modo trasversale a molteplici categorie (appartenenti alla medesima classe principale).

Questo uso "sintattico" della numerazione Hornbostel-Sachs si avvale di una specifica punteggiatura: " - " (trattino), " . " (punto), " : " (due punti), "] " (parentesi quadra) e " + " (più). Ecco un esempio tratto dalla Scheda Guizzi che descrive la *müsa*, aerofono poliorganico tipico della cosiddetta "area delle quattro province":

422-62:.2]112.2+211.1

Questo numero HS è ricco di informazioni: la prima parte, quella che precede la parentesi, esprime il fatto che si tratta di "tubi ad ancia in serie con serbatoio d'aria flessibile" (422-62:.2), mentre nella seconda si differenziano i tubi come "un oboe a caneggio conico munito di fori digitali" (il chanter: 112.2) e "un clarinetto cilindrico privo di fori digitali" (il bordone: 211.1)⁴⁴.

44 Per una descrizione più dettagliata del significato della punteggiatura, si veda la nota del campo "HSN" redatta da Guizzi (cfr. *infra* p. 177).

La revisione della classificazione HS elaborata da MIMO si avvale dell'uso dei suffissi ma non di quello delle parentesi e dei due punti per la descrizione degli strumenti poliorganici:

Since the numerical codes must be used consistently within the databases of the different MIMO partners, in the practical application of Hornbostel Sachs numbers to multicategory instruments within this digital context none of the abbreviations suggested by Hornbostel and Sachs have been used, rather, the codes have been used in full, without colons or brackets. As a general principle, the numerical code for any suffix that applies (or suffixes that apply) to all the categories appears at the end of the series of numbers. Thus the full numerical code for the Highland bagpipe would appear as 422.112-7+422.22-62 Double-reed chanter, conical bore (-7 with fingerholes) + set of single-reed drones with cylindrical bore (-62) flexible air reservoir for all pipes. (MIMO, 2011, p. 3)

Un uso simile della numerazione HS comporta una perdita di informazione: l'assenza della parentesi quadra e dei due punti non consente di dire che sia il chanter che i bordoni sono canne in serie installate sullo stesso blocco. Inoltre la nuova norma secondo cui il suffisso elencato in fondo al codice si applica a tutte le categorie elencate può generare ambiguità: non è contemplato il caso in cui esso possa riferirsi *solo* all'ultimo codice numerico.

MIMO ha optato per una semplificazione della sintassi perché i codici numerici devono essere usati in modo consistente in diversi database, in un contesto digitale. L'analisi che abbiamo condotto sul linguaggio SKOS impiegato per la definizione del vocabolario HS non ha però evidenziato limiti o ostacoli all'uso di tutta la punteggiatura HS: questi ci sarebbero stati se essa fosse stata impiegata direttamente nella stringa che compone gli HTTP URI, perché la punteggiatura contiene caratteri notoriamente "ambigui" per i calcolatori; inoltre avrebbe interrotto la cardinalità di una progressione numerica che supponiamo potesse aiutare a conferire un ordine nella successione degli URI. Ma, come abbiamo visto in precedenza⁴⁵, la parte finale degli URI è formata da semplici numeri primi e il numero HS è inserito all'interno di un elemento `skos:prefLabel`, che può contenere qualsiasi successione di caratteri.

45 Cfr. *supra* p. 105 (nota 54).

La tassonomia HS di MIMO associa un nuovo concetto ad ogni codice numerico individuato nella sua composizione. Questo modello non prevede una rappresentazione sintattica della punteggiatura HS; tutti i codici numerici sono identificati come entità autonome sullo stesso piano ontologico, senza riguardo per la presenza di punteggiatura. L'unica relazione che struttura la tassonomia è quella “padre-figlio” con vincolo di cardinalità “1:N”.

ConceptScheme : <http://www.mimo-db.eu/HornbostelAndSachs>

skos:Concept
314.122-4 True board zithers with resonator box (box zither) sounded by hammers or beaters < >
<http://www.mimo-db.eu/HornbostelAndSachs/6456>

skos:definition
 One or more strings are stretched between fixed points. The string bearer is a rigid board, with a box-shaped resonator made of slats which is not integral and can be detached without destroying the sound-producing apparatus. The plane of the strings is parallel with that of the string bearer. Sounded by hammers or beaters.

skos:prefLabel
 • 314.122-4 True board zithers with resonator box (box zither) sounded by hammers or beaters (pivot)

skos:broader
 ► [Simbing](#)
 ►► [314.122 True board zithers with resonator box \(box zither\)](#)

skos:exactMatch

- <http://www.mimo-db.eu/InstrumentsKeywords/3705>
- <http://www.mimo-db.eu/InstrumentsKeywords/3644>
- <http://www.mimo-db.eu/InstrumentsKeywords/4798>
- <http://www.mimo-db.eu/InstrumentsKeywords/4962>
- <http://www.mimo-db.eu/InstrumentsKeywords/3689>
- <http://www.mimo-db.eu/InstrumentsKeywords/3633>
- <http://www.mimo-db.eu/InstrumentsKeywords/3713>

fig.9 Classe HS “314.122-4”: rappresentazione grafica dello schema concettuale

<http://www.mimo-db.eu/>

Prendiamo il caso del *cimbalom* di fig.9; esso è classificato da un *taxon* e un numero HS:

314.122-4 True board zithers with resonator box (box zither) sounded by hammers or beaters

Il concetto così espresso è identificato da un URI:

<http://www.mimo-db.eu/HornbostelAndSachs/6456>

Il linguaggio SKOS definisce il rapporto tra l'URI del *cimbalom* e quello del suo “genitore”, che è:
<http://www.mimo-db.eu/HornbostelAndSachs/182>

che identifica la categoria:

314.122 True board zithers with resonator box (box zither)

Abbiamo così a che fare con due concetti che appartengono allo stesso livello ontologico (con la differenza che uno dei due, 314.122-4, contiene un suffisso) e non, come dovrebbe essere, con un unico concetto a cui è stato applicato un suffisso.

La struttura ad albero elaborata da MIMO rappresenta la sistematica HS in modo rigido, ovvero come la cristallizzazione di un modello dinamico di classificazione applicato a uno specifico *dataset* in uno specifico momento.

Questa non-completa corrispondenza tra i due modelli genera un'altra anomalia; consideriamo il concetto di un aerofono poliorganico (cfr. fig.10):

422.122+422.22-62

Set of conical double-reed chanters + set of cylindrical single-reed drones; flexible air reservoir

Esso è “figlio” di uno solo dei due concetti che lo definiscono, ovvero:

422.122 Sets of oboes with conical bore

La tassonomia di MIMO non è poligerarchica. Tutti gli elementi hanno solo un “padre”, anche se il linguaggio SKOS consente la creazione di strutture poligerarchiche: l'elemento `skos:broader` è infatti ripetibile⁴⁶.

In conclusione, siamo consapevoli che quelle adottate sono soluzioni di compromesso che possono comunque essere efficaci e funzionali nel contesto tecnologico per cui sono state elaborate. Gli esempi riportati sono però utili per comprendere a quali tipi di problemi si va incontro nella rappresentazione della conoscenza applicata alle nuove tecnologie dell'informazione.

46 «As is often the case in KOS, a SKOS concept can be attached to several broader concepts at the same time. For example, a concept `ex:dog` could have both `ex:mammals` and `ex:domesticatedAnimals` as broader concepts.»
<https://www.w3.org/TR/skos-primer/>

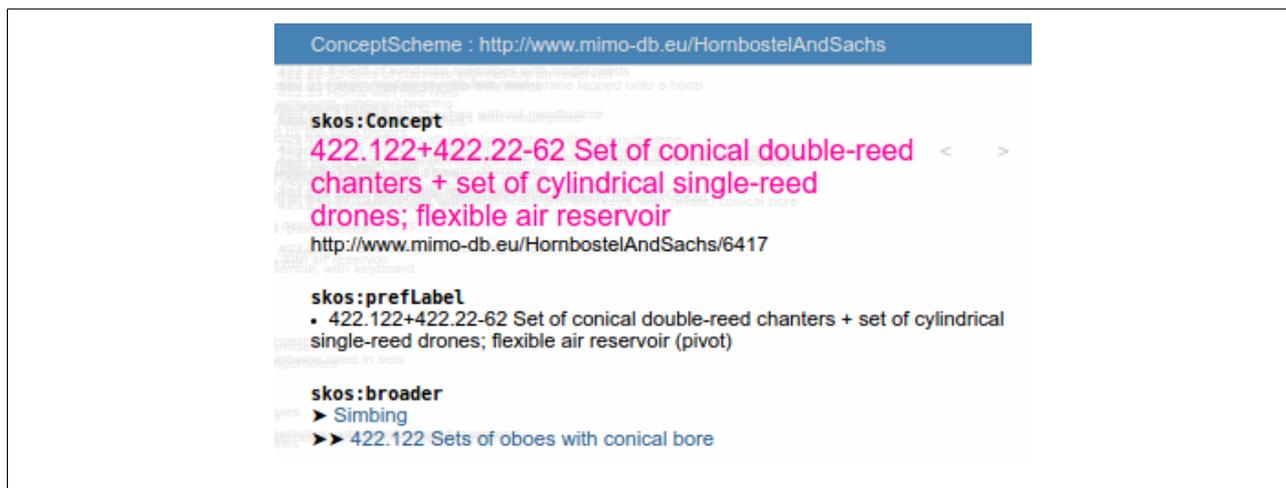


fig.10 Classe HS “422.122+422.22-62”: rappresentazione grafica dello schema concettuale

<http://www.mimo-db.eu/>

Anche il vocabolario delle *keywords* è strutturato in una gerarchia (di tre livelli). Ogni strumento inserito nel database MIMO deve obbligatoriamente essere associato a una *keyword*.

A differenza della sistematica HS, che è una classificazione specialistica, le *keywords* sono rivolte anche alle esigenze di un’utenza meno esperta. L’articolazione dei tre livelli gerarchici è stata fatta sulla base di un criterio non scientifico, che rispecchia l’uso linguistico e le competenze organologiche di molti gruppi sociali occidentali.

Il livello gerarchico superiore suddivide gli strumenti in 9 categorie principali:

- parti di strumenti;
- elettronici;
- a tastiera;
- meccanici;
- mirliton;
- a percussione;
- a corda;
- a fiato;
- altri strumenti.

Questa classificazione accorpa in modo disomogeneo criteri diversi di classificazione; per questo motivo non può essere applicata coerentemente e alcuni strumenti devono essere collocati in categorie create *ad hoc*⁴⁷ o in categorie sbagliate dal punto di vista organologico. È questo il caso dello *scacciapensieri*, che è classificato tra gli strumenti a percussione, mentre per la HS è un “lamellafono (idiofono a pizzico)”; siccome i vocabolari di *keywords* e sistematica HS sono collegati tra di loro⁴⁸, accade che due classificazioni non corrispondenti e ingannevoli siano riferite allo stesso strumento. Crediamo che non sia questo il tipo di informazione strutturata che i musei possono fornire al mondo dei *Linked Open Data*; in questo caso sarebbe meglio redigere una semplice lista controllata di termini piuttosto che un *thesaurus* di difficile applicazione perché non fondato su criteri scientifici.

Le *keywords* sono *parole*-chiave connesse alla dimensione linguistica dello strumento, cioè al modo in cui esso è o è stato nominato. A questo proposito è opportuno riflettere sulla funzione svolta dai vocabolari nelle tecnologie LOD; essi permettono di connotare semanticamente i dati: definire una *keyword* significa definire un concetto applicabile a una o possibilmente a più istanze del dominio descritto. Affidarsi troppo all’ambito delle nomenclature locali degli strumenti significa rischiare di strutturare un *corpus* di nomi di scarsa efficacia classificatoria. Lo stesso “tipo” di strumento può essere conosciuto con un numero troppo elevato di varianti locali perché si possa giungere facilmente a un consenso su quale sia la *keyword* adatta; tali varianti si esprimono spesso *solo* sotto forma orale, e la loro trascrizione comporta un problema scientifico e uno di consistenza dei dati. Il rischio è quello di avere una lista di termini troppo contestualizzati nelle lingue locali per poter essere riusati nelle descrizioni di altri strumenti, se non rischiando di forzare il significato delle parole.

È quindi necessaria una definizione di carattere più generale che cerchi di risolvere le ambiguità linguistiche e garantire la possibilità di riuso dei concetti⁴⁹. Una possibile soluzione potrebbe essere l'uso a questo scopo della “Definizione tipologica” descritta da Anthony Baines⁵⁰ e Febo Guizzi⁵¹.

47 Ad esempio i mirliton e tutti gli “Altri strumenti” che comprendono, tra l’altro, gli idiofoni a frizione, i bicchieri e le seghe musicali.

48 Il collegamento avviene tramite l’elemento `skos:exactMatch`. La definizione del W3C afferma che «the property `skos:exactMatch` is used to link two concepts, indicating a high degree of confidence that the concepts can be used interchangeably across a wide range of information retrieval applications». - <https://www.w3.org/TR/skos-reference/>

49 Le varianti locali non andrebbero ignorate ma bensì ricondotte alla *keyword* sotto l’etichetta di “nomi alternativi” (usando `skos:altLabel`).

50 Baines (1960).

51 Guizzi descrive il concetto di “Definizione tipologica” e le sue applicazioni pratiche nelle note della scheda catalografica. Pubblicata integralmente in Appendice 1, se ne riportano qui alcuni passaggi.

Essa consente di inserire lo strumento in un insieme di strumenti aventi caratteristiche morfologiche e funzionali omogenee. Talvolta i nomi con cui vengono designati alcuni strumenti musicali possono costituire anche definizioni tipologiche; è questo il caso della “müsa” descritta da Ghirardini (2009, pp. 250-255), laddove la scelta del termine locale intende specificare l’unicità dello strumento nel panorama italiano. In altri casi la definizione tipologica

può essere [...] ricavata dalla stessa sistematica Hornbostel-Sachs [...] quando il taxon corrispondente all’apertura del “secondo livello” gerarchico-definitorio (quello stampato in corsivo, che si forma con l’apposizione del primo punto di separazione della numerazione Dewey) esprime una definizione che assume i caratteri individuanti un tipo.

Esempi:

111.1 Idiofoni a percussione reciproca ovvero *crotali*

ove la definizione tipologica è costituita da “crotali”

321.2 Liuti a giogo ovvero *Lire*

ove la definizione tipologica è costituita da “Lire”

La definizione tipologica tuttavia può essere ricavata anche da altri livelli della tassonomia. [...] In alcuni casi la si deve scovare a livelli molto avanzati della gerarchia, come ad es. nel caso della “campana”⁵².

La derivazione della definizione tipologica dalla sistematica HS garantisce un approccio scientifico, e per questo essa va consultata di volta in volta per verificare che ci sia questa opportunità.

In ogni caso il criterio fondamentale per la formulazione di una corretta e utile definizione tipologica è contemporaneamente di carattere empirico, astrattivo e linguistico: si tratta cioè di esaminare bene lo strumento per ricondurlo a caratteristiche generali che lo accomunino ad altri analoghi o all’interno dello stesso taxon [...] e considerare poi gli usi linguistici correnti e quelli più specialmente organologici⁵³.

Sarebbe quindi opportuno formulare, per ogni strumentario locale, una lista (non gerarchica) di definizioni tipologiche sulla base di un criterio condiviso; le liste, redatte a partire dalle competenze sui domini locali, possono essere il punto di partenza per un confronto (internazionale) tra i diversi modi di suddividere gli strumenti e per una conseguente traduzione dei termini nelle lingue nazionali.

In questo modo le istituzioni museali perseguirebbero il duplice scopo di fornire nuovi strumenti alla tecnologia del *semantic web* e raffinare la conoscenza del proprio patrimonio mediante la ricerca organologica.

52 Cfr. *infra* pp. 178-179.

53 *Ibidem*.

Conclusioni e sviluppi futuri

Il confronto tra le teorie organologiche, le prassi catalografiche che ne derivano e i più recenti sviluppi tecnologici legati al *semantic web* può risultare vantaggioso per tutti gli ambiti e i progetti di ricerca.

Per identificare chiaramente il proprio ruolo ed orientare consapevolmente le proprie scelte, le istituzioni della memoria devono comprendere a fondo le novità introdotte dal nuovo contesto tecnologico: l'accostamento dei paradigmi deve svolgersi in un contesto chiaro, di reciproca conoscenza e comprensione delle politiche di sviluppo, e di esigenze, limiti e possibilità tecniche.

Le direzioni di sviluppo che abbiamo identificato nel presente lavoro possono essere molteplici: per alcune si tratta di approfondire percorsi già avviati, mentre altre comportano un ripensamento radicale di alcuni aspetti (come il ricorso ad un'ontologia oggetto-centrica, o la redazione di una lista controllata di *keywords* non gerarchica).

Nel nuovo contesto tecnologico degli *open data* l'aggregazione dei dati assume un valore di tipo conoscitivo *in sé*, e va per questo motivo sfruttata e valorizzata; le tecnologie digitali a disposizione consentono, tramite le mappature concettuali e gli strumenti di recupero e indirizzamento puntuale dei dati, di rivedere e correggere i percorsi già intrapresi o di scegliere consapevolmente di percorrere strade parallele. Nella molteplicità di strade possibili le soluzioni sono di carattere collaborativo.

Le presenti "conclusioni" elencano una sequenza di possibili sviluppi del lavoro, organizzata in maniera modulare in base a criteri di natura pratica e di reale implementazione.

Tali proposte intendono seguire un duplice intento di ricerca: affinare la conoscenza del dominio organologico e rendere più efficace lo sviluppo del *semantic web*; allo stesso tempo esse vanno costantemente relazionate all'interesse dovuto agli "oggetti" di tale ricerca, e più in particolare a quell'intento di valorizzazione della collezione del Museo del Paesaggio Sonoro che è stato lo spunto per il presente lavoro e ha rappresentato un costante punto di riferimento. La sua originalità, compattezza e pregnanza organologica hanno costituito la misura di molte osservazioni. Si auspica che i possibili sviluppi della ricerca possano contribuire ad una ulteriore valorizzazione della collezione, e soprattutto che l'adozione di strumenti tecnologici d'avanguardia conduca ad un potenziamento della sua presenza per tutta la comunità, non solo quella accademica.

Sviluppi futuri:

1. Redazione di una versione 2.0 dell'XML Schema della Scheda Guizzi

Il progetto sviluppa le “Linee-guida per successive implementazioni”. Si tratta di elaborare uno schema che preservi l'efficacia catalografica originaria della Scheda Guizzi e che nello stesso tempo sia predisposto alla resa di descrizioni evento-centriche (la direzione intrapresa dall'ICOM e, in campo più specificamente organologico, da MIMO).

A questo scopo lo schema dovrà implementare le caratteristiche individuate nelle linee-guida (ad es. l'incremento degli annidamenti, l'uso del meccanismo *Wrap/Set* e quello degli elementi *display*) e sarà inoltre necessario sviluppare un nuovo sistema per la descrizione delle misurazioni degli strumenti (il sistema dovrà supportare l'implementazione di tabelle XML).

Fondamentale è predisporre una struttura che porti ad una efficace atomizzazione dell'informazione, ovvero una scomposizione dei campi di testo libero della Scheda Guizzi in *dati discreti*. I dati risultanti dalla scomposizione potranno poi essere mappati anche in sistemi ontologici diversi da quello evento-centrico¹.

Infine, per un'ottimizzazione del lavoro, lo sviluppo dell'XML Schema dovrebbe procedere contestualmente alla realizzazione del DBMS (come descritto al punto successivo). In questo modo l'elaborazione dello schema si gioverebbe del *feedback* generato dalla sua messa in pratica (nella realizzazione dell'interfaccia e nella mappatura effettiva delle schede catalografiche) e del supporto di una *software house*.

2. Realizzazione di un DataBase Management System (DBMS) basato sulla versione 2.0 dello schema XML

La realizzazione tecnica del progetto si affida al contributo di una *software house*. L'idea è quella di personalizzare un DBMS *open source*; in questo modo, oltre a contenere i costi, ci si può avvalere del supporto tecnico a lungo termine della comunità di sviluppatori informatici.

Per essere davvero efficiente, il DBMS ha bisogno di essere testato in un progetto di catalogazione di una (o parte di una) collezione², che generi un *feedback* utile per il lavoro di *debugging*.

1 Una prima implementazione del modello oggetto-centrico può già essere prevista nella versione 2.0, tramite il ricorso ai noti sistemi di metadati *Dublin Core* e *DCMI Terms*.

2 Nel caso della collezione del Museo del Paesaggio Sonoro, si consideri che, oltre alle tradizionali funzionalità di *back-office*, alcuni DBMS *open source*, offrono un'implementazione *front-end* per la pubblicazione diretta dei dati catalogati sul sito web.

Il DBMS è uno strumento agile e preciso per l'attività di catalogazione: si avvale di procedimenti automatizzati per l'inserimento dei dati e genera automaticamente dei record in formato XML (circostanza che rende molto più agevoli successive mappature in nuovi modelli ontologici).

3. Valutazione di modelli di descrizione di impostazione oggetto-centrica (alternativi al modello evento-centrico)

Si è rilevato che il CIDOC-CRM, l'ontologia predisposta dall'ICOM per la rappresentazione del dominio del "patrimonio culturale", ha alcune incompatibilità con il modello elaborato da Guizzi. Il CRM infatti adotta un'impostazione evento-centrica molto diversa da quella della Scheda Guizzi (che dimostra una particolare *attenzione per l'oggetto* della catalogazione).

Sarà opportuno approfondire la ricerca da un punto di vista ontologico al fine di verificare le possibilità di coesistenza dei due modelli³ e l'eventuale esistenza di ontologie oggetto-centriche specifiche per il dominio del patrimonio culturale.

Si potrà inoltre compilare una libreria di ontologie adatta al dominio specifico degli strumenti musicali (applicando il sistema delle ontologie "ibride" e prevedendo l'eventuale inclusione di nuove classi e proprietà di carattere esclusivamente "organologico").

Le nuove osservazioni sul modello oggetto-centrico dovranno essere integrate nelle successive versioni dello schema XML della Scheda Guizzi.

4. Realizzazione di strumenti per la mappatura automatica dei dati dai record Guizzi (versione XML 2.0) ai record MIMO

Il lavoro prevede l'aggiornamento delle "Istruzioni per una mappatura Guizzi-MIMO" pubblicate in Appendice 3 che vanno riviste alla luce della versione 2.0 dello schema XML della Scheda Guizzi.

Sarà quindi necessaria la programmazione di un *tool* per la mappatura automatica dei dati, avvalendosi della collaborazione di una *software house*.

La disponibilità di record catalogafici in formato XML generati da un DBMS renderebbe l'operazione molto più agevole.

³ Il progetto Europeana ha predisposto una struttura (EDM – Europeana Data Model) capace di rendere parallelamente descrizioni evento-centriche ed oggetto-centriche dello stesso oggetto (cfr. *supra* pp. 131-134).

5. Proposte di emendamento da presentare al progetto MIMO

Nel capitolo “Il problema delle *keywords* e della sistematica Hornbostel-Sachs” alcuni aspetti dei vocabolari redatti da MIMO sono stati sottoposti a una critica di carattere organologico che potrebbe evolvere nella forma di proposte di emendamento. MIMO si è rivelato particolarmente aperto alla condivisione di idee e proposte per il suo sviluppo e affinamento, soprattutto in merito alla redazione dei vocabolari, strutturalmente predisposta ed improntata ad uno spirito collaborativo.

Le prime osservazioni riguardano il vocabolario della classificazione Hornbostel-Sachs. Le proposte di emendamento possono così essere riassunte:

- l’uso dei numeri HS andrebbe ricondotto alle indicazioni originali degli autori, ovvero utilizzando la punteggiatura (la parentesi quadra e i due punti) prevista per la descrizione degli strumenti poliorganici;
- la struttura del vocabolario dovrebbe implementare la poligerarchia al fine di migliorare la descrizione degli strumenti poliorganici;
- la struttura del vocabolario dovrebbe predisporre un sistema per rappresentare correttamente il meccanismo di apposizione dei suffissi HS.

Il vocabolario è strutturato tramite il linguaggio SKOS. Se per i primi due punti (codici HS e poligerarchia) la nostra analisi ha verificato che con SKOS è possibile attuare le proposte, per la rappresentazione dei suffissi è necessaria una ricerca di fattibilità.

Le proposte relative al *thesaurus* delle *keywords* riguardano:

- la conversione del *thesaurus* da struttura gerarchica articolata su tre livelli a lista controllata di termini posti sullo stesso livello gerarchico (la fenomenologia linguistica connessa agli strumenti musicali è troppo varia perché si possa individuare un criterio scientifico valido per definire una struttura gerarchica coerente);
- l’adozione di un criterio condiviso per l’individuazione dei termini da inserire nella lista di *keywords*; a tal proposito si può prendere in considerazione la “Definizione tipologica” descritta da Guizzi e Baines⁴.

4 Cfr *supra* pp. 167-168.

6. Elaborazione di una lista di definizioni tipologiche

La proposta riguarda un lavoro di ricerca di ambito organologico, i cui esiti possono essere sfruttati anche nel nuovo contesto tecnologico della catalogazione. Si tratta di redigere liste di definizioni tipologiche a partire dall'analisi di strumentari locali di dimensioni ridotte che, una volta confrontate, andranno a confluire in una lista che possa accogliere strumentari sempre più ampi ed eterogenei. Oltre al beneficio di definire una serie di termini riusabili, il progetto servirebbe a chiarire e condividere, tra i diversi gruppi di lavoro, i criteri per una corretta definizione tipologica.

Appendice

1. La Scheda Guizzi

Scheda organologica elaborata da Febo Guizzi
(Università di Torino)

Organological file worked out by Febo Guizzi
(Torino University)

SCHEDA [FILE]: Si indica qui il numero progressivo della scheda all'interno del lavoro che si sta svolgendo. Nel caso di un nucleo di strumenti simili per caratteristiche formali e dimensioni e che non pongano particolari problemi di descrizione, è possibile riunire più strumenti in una sola scheda: ciò avviene in pratica soprattutto per oggetti di piccole dimensioni e/o di produzione industriale (es. bubboli, ancie di produzione industriale). Si può avere il caso in cui ad uno stesso numero di inventario corrispondano più schede, ad esempio se un oboe è corredato da ancie di riserva e da un'apposita custodia si può scegliere di descrivere l'oboe in una scheda, e gli accessori in un'altra: ciò risulta pratico poiché consente una maggior chiarezza nella distribuzione dei dati tra le voci della scheda.

COLLOCAZIONE [PLACING]: Per collocazione si intende il luogo in cui lo strumento è conservato. Si può intendere in due modi: l'istituzione che possiede lo strumento, senza indicare il luogo preciso dove lo strumento è collocato; oppure nel caso in cui appaia scontata la proprietà dell'oggetto (vale a dire è già specificata in una premessa generale alle schede) si può indicare in questa voce il luogo preciso in cui è conservato lo strumento all'interno dell'edificio dove si trova la raccolta. Nel caso di musei e raccolte aperte al pubblico tale dato consente di sapere immediatamente se lo strumento è esposto o se per esempio si trova in magazzino.

INVENTARIO GENERALE [BROAD INVENTORY]: Si indica qui il numero di inventario dell'oggetto.

INVENTARIO PARTICOLARE [SPECIAL INVENTORY]: In alcuni casi può essere presente un inventario per categoria di oggetti di proprietà del museo o della raccolta. Nel caso di musei non specializzati in strumenti, e che quindi possiedono raccolte di vario genere e non dispongono di un inventario generale, è opportuno compilare contestualmente alla schedatura un inventario apposito per gli strumenti musicali.

ACQUISIZIONE [ACQUISITION]: Si indicano qui le modalità attraverso cui lo strumento è diventato di proprietà del museo o della raccolta: ad esempio, acquisto, dono, raccolta durante una ricerca sul campo, ecc. In ogni caso è utile indicare tutte le circostanze che si conoscono relativamente all'acquisizione (dove è stato acquistato, quando, il prezzo pagato, chi è il donatore, quando è stato donato lo strumento, in che circostanze è stato raccolto ecc.). Se vi sono stati passaggi di proprietà e si conoscono i dati relativi, è utile riportarli. Nel caso di istituzioni pubbliche

dotate di una documentazione contabile-amministrativa, è utile indicare l'esistenza e il tipo dell'eventuale documentazione relativa all'acquisizione (delibere, stime, fatture, mandati di pagamento, ecc.) e la collocazione dei documenti originali.

COLLEZIONE [COLLECTION]: Nel caso in cui un nucleo di strumenti musicali si sia costituito tramite l'acquisizione di più collezioni, si indica qui la collezione di appartenenza del singolo strumento.

INVENTARIO TOPOGRAFICO [TOPOGRAPHIC INVENTORY]: Se esiste si indica qui la sigla corrispondente al luogo (o alla struttura espositiva/conservativa: ad es. "vetrina n°.., scaffale n°...", ecc.) occupato dallo strumento in questione all'interno della raccolta.

CATEGORIA [CATEGORY]: SM (sta per Strumento Musicale). Tale voce è utile nel caso la raccolta di cui gli strumenti musicali fanno parte comprenda altre categorie di oggetti.

STRUMENTO [INSTRUMENT]: Si indica qui il nome italiano con il quale si designa lo strumento in questione, lasciando le varianti locali o dialettali alla voce NOME ORIGINARIO. Nel caso in cui non esista un nome in italiano, si usa il nome dialettale più diffuso (es. *launeddas*, indicando poi in NOME ORIGINARIO il tipo specifico). Nel caso di strumenti di provenienza non italiana si indica qui il nome in italiano e si riporta il nome nella lingua originaria alla voce NOME ORIGINARIO.

TAXON HORNBOSTEL-SACHS [HORNBOSTEL-TAXON TAXON]: Si indica qui il taxon corrispondente allo strumento nella classificazione Hornbostel Sachs, utilizzando la traduzione pubblicata in GUIZZI 2002¹. Dato che le varie suddivisioni della classificazione HS si diramano in definizioni sempre più specifiche a partire da livelli superficiali più generali, si è ritenuto opportuno indicare non solo il taxon specifico (con gli eventuali suffissi) ma anche i livelli superiori che indicano il percorso logico che porta a quel taxon.

1 FEBO GUIZZI, *Gli strumenti della musica popolare in Italia*, Lucca, LIM, 2002: 409-482. [NdA]

Esempi:

- Per le **castagnette** si indica: Idiofoni – Idiofoni a percussione non mediata – Idiofoni a percussione reciproca ovvero crotali – Castagnette.
- Per una **fisarmonica** si indica: Aerofoni liberi – Aerofoni ad interruzione autofonici ovvero ancie – Ancie libere in serie, con serbatoio d'aria flessibile, con tastiera.

HSN [HSN: HORNBOSTEL-SACHS NUMBER]: Si indica il numero della classificazione Hornbostel Sachs corrispondente al taxon. Per gli strumenti poliorganici si compone il numero seguendo le indicazioni date dagli autori nella prefazione alla sistematica.

Esempi:

- **Castagnette:** 111.141
- **Fisarmonica:** 421.132 -62 -8
- **Zampogna a chiave:** 422-62:.2]122-71+122+121 (422 indica che si tratta di tubi ad ancia; -62 è il suffisso comune che indica il serbatoio d'aria flessibile; i due punti si mettono per indicare che la suddivisione continua oltre il suffisso; il punto indica che manca una cifra e che sarà riportata dopo la parentesi quadra, in questo caso si tratta della cifra che indica che si tratta di oboi; 2 indica che si tratta di tubi ad ancia in serie; 122-71 si riferisce al chanter con la chiave e indica gli oboi a canneggio conico muniti di fori digitali: è composto dalla cifra 1 tralasciata per prima, da 2 che indica il canneggio conico e dal 2 che indica la presenza dei fori digitali; -71 è il suffisso che indica la presenza di una meccanica a chiavi. 122 si riferisce al chanter senza chiave (oboe a canneggio conico con fori digitali); 121 si riferisce al bordone (oboe conico privo di fori digitali).
- **Müsa:** 422-62:.2]112.2+211.1 (422 indica che si tratta di tubi ad ancia; -62 è il suffisso comune che indica il serbatoio d'aria flessibile; i due punti si mettono per indicare che la suddivisione continua oltre il suffisso; il punto indica che manca una cifra e che sarà riportata dopo la parentesi quadra, in questo caso si tratta della cifra che specifica il tipo di ancia (doppia, singola o libera); 2 indica che si tratta di tubi ad ancia in serie; 112.2 si riferisce al chanter: un oboe a canneggio conico munito di fori digitali; 211.1 si riferisce al bordone: un clarinetto cilindrico privo di fori digitali.

DEFINIZIONE TIPOLOGICA [TYPOLGICAL DEFINITION]: Per definizione tipologica si intende una definizione che consente di inserire lo strumento musicale in questione in un insieme di strumenti aventi caratteristiche morfologiche e funzionali omogenee. Le classi di strumenti così ottenute integrano la sistematica Hornbostel Sachs poiché consentono accorpamenti o distinzioni di strumenti non previsti da essa. Per esempio la ciaramella dell'Italia centro-meridionale e il piffero delle Quattro Province, pur essendo entrambi oboi a caneggio conico e privi di chiavi, rientrano in definizioni tipologiche diverse: la ciaramella rientra nella definizione tipologica “oboe conico”, il piffero rientra nella definizione “oboe conico con pirouette”. Talvolta i nomi con cui vengono designati alcuni strumenti musicali possono costituire anche definizioni tipologiche, ad esempio “organetto” e “fisarmonica”². La definizione tipologica può essere anche ricavata dalla stessa sistematica Hornbostel-Sachs (che va comunque consultata di volta in volta anche allo scopo di verificare questa opportunità), quando il taxon corrispondente all'apertura del “secondo livello” gerarchico-definitorio (quello stampato in corsivo, che si forma con l'apposizione del primo punto di separazione della numerazione Dewey) esprime una definizione che assume i caratteri individuanti un tipo.

Esempi:

111.1 Idiofoni a percussione reciproca ovvero crotali

ove la definizione tipologica è costituita da “crotali”

321.2 Liuti a giogo ovvero Lire

ove la definizione tipologica è costituita da “Lire”

La definizione tipologica tuttavia può essere ricavata anche da altri livelli della tassonomia, soprattutto quando la raccolta di cui ci si sta occupando comprenda varianti morfologiche specifiche di strumenti che richiedano una più precisa determinazione — ad es. “castagnette” in luogo di “crotali” — oppure varianti locali con precisa determinazione tipologica (esempio: il già citato caso di ciaramella e piffero, distinti dall'assenza o dalla presenza di una pirouette). In alcuni casi la si deve scovare a livelli molto avanzati della gerarchia, come ad es. nel caso della “campana”. In ogni caso il criterio fondamentale per la formulazione di una corretta e utile definizione tipologica è contemporaneamente di carattere empirico, astrattivo e linguistico: si tratta cioè di esaminare bene lo strumento per ricondurlo a caratteristiche generali che lo accomunino ad altri analoghi o all'interno dello stesso taxon (si veda il caso degli oboi sopra richiamati o quello di “traccola” che può comprendere la specificazione “a martelli”, “a martelli oscillanti”, “a piastra oscillante”, ecc., o

2 Si veda: ANTHONY BAINES, “Organology and European Folk Music Instruments”, in *International Folk Music Journal*, XII, 1960: 10-13. [NdA]

può riferirsi trasversalmente a taxa diversi, come nel caso della “traccola mista” che comprenda sia piastre a battente che raschiatoi a ruota, azionati tutti dalla stessa manovella) e considerare poi gli usi linguistici correnti e quelli più specialmente organologici.

Al fine di distinguere la definizione tipologica da quella sistematica, si deve intendere la prima come definizione dello strumento secondo un criterio individuante universale concreto, per lo più rispondente agli usi linguistici comuni (es.: lo strumento di Arthur Rubinstein: “pianoforte”) e la seconda, ovviamente, come definizione dello strumento secondo la classificazione organologica di Hornbostel - Sachs (es.: lo strumento di Arthur Rubinstein: “cetra a cassa di risonanza suonata a mezzo di martelletti munita di tastiera”), ma anche la distinzione tra “pianoforte”, “fortepiano”, “clavicordo” come casi specifici e rilevanti all’interno dello stesso taxon qui sopra riportato, quando non si sia proceduto all’integrazione della classificazione con la sua articolazione in profondità entro la stessa definizione sistematica (si veda l’esempio degli xilofoni riportato da Hornbostel-Sachs nella loro introduzione o si pensi ad analoghe procedure riferite alle suddette “cetre a cassa di risonanza suonate a mezzo di martelletti munite di tastiera” o alle “cetre a cassa di risonanza suonate a mezzo di plettri munite di tastiera”, distinguendovi clavicembali, spinette, virginali e così via).

LUOGO DI FABBRICAZIONE [*PLACE OF CONSTRUCTION*]: Si indica qui, quando è noto, il luogo dove lo strumento è stato costruito. Se non si conosce il dato, si può scegliere di lasciare vuoto il campo o di inserirvi “sconosciuto” oppure “ignoto”.

LUOGO D'USO [*PLACE OF UTILIZATION*]: Si indica qui, quando è noto, il luogo oppure l'area in cui lo strumento è stato utilizzato. Se lo strumento è stato ceduto da un suonatore ad un altro è utile indicare dove ciascuno lo ha utilizzato, ove si tratti di zone diverse.

CULTURA [*CULTURE*]: Si indica la cultura di appartenenza dello strumento. Quando è possibile è utile indicare l’etnia. Se non si conosce il dato, si può scegliere di lasciare vuoto il campo o di inserirvi “sconosciuta” oppure “ignota”.

COSTRUTTORE [*MAKER*]: Si indica il costruttore dello strumento, se noto. Se non si conosce il dato, si può scegliere di lasciare vuoto il campo o di inserirvi “sconosciuto” oppure “ignoto”.

NOME ORIGINARIO [*NATIVE DENOMINATION*]: Si indica qui il nome dello strumento nella lingua originale, le varianti locali e/o dialettali. Nel caso si tratti di uno strumento costituito da più parti, è utile indicare anche il nome specifico con cui si indica ciascuna delle parti nelle zone di costruzione ed utilizzo dello strumento in questione. Se non si conosce il dato, si può scegliere di lasciare vuoto il campo o di inserirvi “sconosciuto” oppure “ignoto”.

EPOCA DI FABBRICAZIONE [*AGE OF CONSTRUCTION*]: Si indica, se nota, l'epoca in cui lo strumento è stato costruito. Se non si conosce il dato, si può scegliere di lasciare vuoto il campo o di inserirvi “sconosciuta” oppure “ignota”. Tuttavia occorrerebbe fare ogni volta uno sforzo di approssimazione (indicando che il dato è probabile) o di ipotesi (indicando che il dato è presunto).

EPOCA D'USO [*AGE OF UTILIZATION*]: Si indica, se nota, l'epoca in cui è stato utilizzato. Se lo strumento è stato ceduto da un suonatore ad un altro si deve riportare il periodo in cui è stato utilizzato da ciascuno dei due.

CARTIGLIO O ISCRIZIONI [*CARTOUCHE OR INSCRIPTIONS*]: Si riporta in trascrizione diplomatica il contenuto di eventuali cartigli o iscrizioni presenti sullo strumento. Per “cartiglio o iscrizioni” si intende: marchi, etichette, targhette, bolli, timbri, incisioni relative al costruttore, a colui che ha venduto lo strumento, al suonatore, ad eventuali riparatori e/o restauratori e alla storia museale dell'oggetto. Eventuali cifre o segni riscontrabili all'interno o all'esterno dello strumento dai quali non si ricavano informazioni sull'origine, sulla proprietà e/o sulla storia dello strumento vanno indicati alla voce DESCRIZIONE, oppure alla voce RESTAURI, se si è certi che si tratta di segni tracciati da eventuali riparatori o restauratori per necessità pratiche. Il numero di inventario corrente, redatto dal museo che possiede l'oggetto, in genere si omette.

DESCRIZIONE [*DESCRIPTION*]: Si redige qui una descrizione il più possibile dettagliata dell'oggetto, in modo da metterne in evidenza la morfologia, gli elementi costitutivi, le eventuali particolarità riscontrate. Vanno descritti con particolare attenzione i dettagli significativi non evidenti da una fotografia. Dalla descrizione, comunque, deve essere possibile riconoscere l'oggetto anche in assenza della fotografia. È opportuno attenersi il più possibile a termini descrittivi ricorrenti di alto contenuto tecnico – derivati dagli usi specifici, ad esempio delle tradizioni liutaria, o ricavabili dal glossario tecnico descrittivo di strutture e morfologie, quale sono quelli dell'architettura o dell'anatomia – e comunque resi omogenei e non equivoci da un'esplicita scelta

convenzionale operata dai responsabili scientifici della schedatura (in tal caso l'uso della terminologia convenzionale è da ritenersi obbligatoria).

Esempi: per oggetti o parti di oggetti sviluppati in lunghezza le estremità vanno distinte in “prossimale” e “distale” in relazione alla loro collocazione rispettiva, nell'uso, in prossimità o a distanza dal corpo dell'utilizzatore. In caso di strumenti curvi, la collocazione di specifici dettagli tecnici lungo la curvatura va distinta utilizzando i termini alternativi di “intradosso” e “estradosso”.

STATO DI CONSERVAZIONE [*CONDITIONS OF PRESERVATION*]: Si riportano qui le osservazioni sullo stato di conservazione dello strumento, evitando di dare semplicemente un giudizio ma specificando le condizioni in cui lo strumento si trova al momento della schedatura, soprattutto se esso non è in buono stato di conservazione. In questa voce è utile indicare gli eventuali interventi di restauro che si ritengono necessari.

RESTAURI [*RESTORATIONS*]: Si indicano in questa voce sia i restauri veri e propri, sia eventuali riparazioni eseguite dal costruttore, dal suonatore o da altri. Quando si conosce è necessario specificare chi ha eseguito l'intervento e l'epoca in cui è stato effettuato.

MATERIALI [*MATERIALS*]: Si indicano qui i materiali di cui è fatto lo strumento, specificando se possibile i tipi di legno, di canna, di metallo, di stoffa ecc.

MISURE [*MEASURES*]: Si indicano qui le misure in millimetri dello strumento: si devono prendere misure utili ad identificare correttamente l'oggetto e a eventuali confronti con altri strumenti analoghi. Si tratta di misure descrittive e non prescrittive: non si tratta cioè di misure destinate a guidare il lavoro dei costruttori dello strumento; esse in ogni caso devono rendere un'idea almeno sommaria delle dimensioni delle varie parti di cui è costituito lo strumento. Questo campo, che è tra quelli considerati indispensabili in qualunque tipo di scheda – anche quando sono molto diverse tra loro –, non può essere predeterminato una volta per tutte nel suo contenuto, sia perché la varietà di misure riportate cambia ovviamente da strumento a strumento, sia perché in ogni specifica schedatura si decide quanto analitiche o sommarie debbano essere le misure da effettuare per ciascun tipo di strumento. Di regola ogni collezione deve mostrare gli stessi criteri oggettivi e soggettivi di scelta e di attuazione delle misure per strumenti analoghi. L'approssimazione nella precisione delle misurazioni deve essere dichiarata: in genere ci si attiene a quanto normali strumenti di misura consentono di rilevare e quindi non è il caso di spingersi al di

sotto del limite di precisione pari a 0,5mm. Alcuni dispositivi essenziali dal punto di vista acustico possono richiedere due tipi diversi di misurazione (ad es.: per le corde, “lunghezza totale” e “lunghezza vibrante”, intendendo questa come distanza tra i due punti estremi della corda libera, come capotasto e ponticello o simili: lo stesso si può fare per i membranofoni rilevando il diametro totale della membrana — quindi considerando anche la parte che poggia sul bordo del supporto e quello che lo supera per consentirne l’ancoraggio — e il diametro vibrante — misurando nei limiti del possibile la parte libera). È utile disporre le misure di dispositivi seriali, specie se numerosi (come le corde o le piastre di xilofoni, metallofoni, ecc., o le lamelle di sanze e così via) in tabelle che rendano immediatamente confrontabili le variazioni mensurali di ciascuno (quindi in colonne e non in fila indiana entro la stessa riga). Soprattutto per alcune classi di strumenti, in primis per gli idiofoni, è importante rilevare il peso dell’oggetto o di sue parti costitutive essenziali ai fini della produzione del suono (ad es.: le singole piastre o lamine di metallofoni, xilofoni e simili). In genere i valori di peso non sono facilmente rilevabili per l’indisponibilità di bilance adatte o per l’ingombro dello strumento o delle sue parti o per la difficoltà a separare dall’insieme ciascuna parte suscettibile di pesatura. In questi casi è opportuno segnalare nella scheda la mancanza forzata di questo tipo di misure con l’intento di segnalarne comunque la rilevanza e di sollecitare pertanto i responsabili della conservazione a provvedere alla misurazione quando si rendesse possibile (ad es. in occasione di un restauro che comporti il temporaneo smontaggio). In questo campo possono essere inserite anche considerazioni valutative sullo stato di conservazione, anche se esiste il campo apposito, soprattutto se si vogliono descrivere particolari situazioni di intervento pregresso e spurio sullo strumento (es.: corde non originali, o membrana sostituita con materiale spurio o montato in modo non conforme, cavigliere incollato con angolo non rispondente alla norma tradizionale, ecc.), nonché considerazioni valutative, o rinvii alle stesse, indicate entro altro campo, sull’autenticità dello strumento stesso o sul suo livello qualitativo, poiché queste caratteristiche possono legittimamente scongiurare l’accuratezza di misurazioni che altrimenti sarebbero di prammatica (es.: strumento falso, simulacro o modellino, versione per turisti o comunque per un mercato estraneo alla funzione culturale primaria, strumento di basso livello per uso di materiali di scarso pregio relativo, ecc.).

ANNOTAZIONI [NOTES AND REMARKS]: Si tratta di un campo libero destinato ad eventuali annotazioni dello schedatore, che servano a dare conto di notizie sulla storia e sulle vicende dello specifico strumento schedato, delle persone cui esso è stato legato nell’uso, delle circostanze dell’utilizzazione, su aspetti significativi della sua costruzione, ecc. (è appartenuto a..., è un tipico

esemplare della tradizione costruttiva di..., è riconducibile al modello cosiddetto “Trimarco”, dal nome del suonatore che per primo lo ha utilizzato in questa particolare forma, è stato usato durante i festeggiamenti per l’indipendenza del paese d’origine alla presenza di, ecc.), sia sull’uso cui strumenti analoghi sono di regola destinati, sul contesto, sulle potenzialità musicali, sulla tecnica esecutiva, sul repertorio, sulla simbologia ad essi legata, su particolarità musicali o extramusicali, su specifici aspetti della struttura costruttiva, sul nome e i particolari significati cui esso si ricollega, ecc.

DOCUMENTAZIONE [DOCUMENTATION]: Si rende conto qui, quando possibile, della presenza, presso il museo o altrove, di altre schede relative allo strumento, eventuali disegni, radiografie, fotografie, misurazioni dettagliate, ecc. Si rammenta qui che è buona norma per musei e collezioni che mettano a disposizione di terzi strumenti delle loro raccolte per la rilevazione di misure, disegni tecnici o altro, esigere che copia della documentazione stessa sia depositata presso l’ente o il soggetto proprietario.

ARCHIVIO FOTOGRAFICO [PHOTOGRAPHIC ARCHIVE]: Se il museo o la raccolta possiedono un archivio fotografico contenente le foto degli strumenti di loro proprietà si rinvia qui alla foto corrispondente allo strumento in questione. Si rammenta qui che è buona norma per musei e collezioni che mettano a disposizione di terzi strumenti delle loro raccolte per rilevazioni fotografiche di qualunque tipo, anche eventualmente di taglio non tecnico-scientifico, esigere che copia delle fotografie stesse sia depositata presso l’ente o il soggetto proprietario.

ESPOSIZIONI [EXHIBITIONS]: Si indicano qui le varie mostre nelle quali lo strumento è stato inserito ovvero alle quali è stato eventualmente prestato, specificando il luogo, la durata di ciascuna mostra e i relativi curatori.

BIBLIOGRAFIA [BIBLIOGRAPHY]: Si indica qui la bibliografia specifica sullo strumento in questione, inteso sia come oggetto specifico della raccolta, se esiste qualche scritto che lo riguardi, sia come tipo di strumento in generale. Il rinvio ai testi può essere effettuato in forma abbreviata, o ricorrendo ad acronimi per indicare riviste o enciclopedie di chiara fama, o usando il metodo AUTORE DATA (esempio GUIZZI 2002: 322) che ovviamente si integrino con apposita legenda degli acronimi e con un elenco bibliografico completo in tutte le sue parti dei testi utilizzati, da mettere in calce alla raccolta delle schede compilate.

COMPILATORE DELLA SCHEDA [*FILE AUTHOR*]: Nome del compilatore della scheda, utile se più schedatori collaborano alla schedatura di una raccolta di oggetti. Eventuali successive integrazioni, correzioni o simili devono essere riferibili al diverso schedatore che le abbia effettuate.

DATA [*DATE*]: Data in cui è stata redatta la scheda. Eventuali successive integrazioni, correzioni o simili devono essere riferibili al momento in cui sono state effettuate.

FIRMA [*AUTHOR'S SIGN*]: Firma del compilatore della scheda, utile se più schedatori hanno collaborato alla schedatura di una raccolta di oggetti.

ALLEGATI [*ATTACHEMENTS*]: Se sono disponibili, si indicano qui eventuali disegni o qualsiasi altro documento che venga allegato alla scheda.

AGGIORNAMENTO [*UPDATES*]: Si può indicare qui la data di eventuali aggiornamenti alla scheda e in sintesi cosa è stato modificato o aggiunto, anche se si è già indicato il dato nel campo specifico.

La scheda, in ogni caso se è redatta con programmi di testo e se è tecnicamente possibile qualora se ne usi una versione redatta con programmi di archiviazione dati, può contenere note (meglio se a pie' di pagina per facilità di consultazione) che chiariscano eventuali scelte, opzioni, criteri adottati dallo schedatore o riferiscano di problemi incontrati e della soluzione scelta per risolverli o aggirarli.

2. scheda guizzi 1_0.xsd

Scheda per la catalogazione di strumenti musicali, redatta dal prof. Febo Guizzi

Revisione e adattamento a cura di Giorgio Bevilacqua, Cristina Ghirardini e Ilario Meandri
Università di Torino

Rappresentazione in forma tabellare dell'XML Schema:

`scheda guizzi 1_0.xsd`

La scheda è pubblicata in formato .xsd e .svg all'indirizzo:

<http://www.suonoeimmagine.unito.it/bevilacqua.html>

La scheda è costituita da 35 campi, per i quali sono indicati descrizione, ripetibilità e obbligatorietà. Alcuni campi sono raggruppati in sovracampi che non vanno compilati (hanno lunghezza del testo uguale a zero e servono a creare "annidamenti").

La dicitura "flag" presente a fianco di alcuni titoli indica il testo da inserire in un campo obbligatorio nella circostanza di assenza del dato.

SOVRACAMPI		Descrizione	Ripet.	Obbl.
CAMPI				
/	SCHEDA	<p>Si indica qui il numero progressivo della scheda all'interno del lavoro che si sta svolgendo. A ciascun strumento deve corrispondere una scheda, salvo in casi di dispositivi sonori pervenuti insieme e necessari l'uno all'altro.</p> <p>Se per esempio uno strumento ad ancia è stato raccolto insieme all'ancia con cui è stato suonato, è utile catalogare l'ancia e lo strumento insieme e fare schede separate per eventuali altre ancie pervenute.</p>	NO	SI
/	COLLOCAZIONE	<p>Per collocazione si intende il luogo in cui lo strumento è conservato. Si può intendere in due modi: l'istituzione che possiede lo strumento, senza indicare il luogo preciso dove lo strumento è collocato; oppure nel caso in cui appaia scontata la proprietà dell'oggetto (vale a dire è già specificata in una premessa generale alle schede) si può indicare in questa voce il luogo preciso in cui è conservato lo strumento all'interno dell'edificio dove si trova la raccolta.</p> <p>Nel caso di musei e raccolte aperte al pubblico tale dato consente di sapere immediatamente se lo strumento è esposto o se per esempio si trova in magazzino.</p>	NO	SI
/	INVENTARIO GENERALE	<p>Si indica qui il numero di inventario dell'oggetto.</p>	NO	SI
/	INVENTARIO PARTICOLARE	<p>In alcuni casi può essere presente un inventario per categoria di oggetti di proprietà del museo o della raccolta. Nel caso di musei non specializzati in strumenti, e che quindi possiedono raccolte di vario genere e non dispongono di un inventario generale, è opportuno compilare contestualmente alla schedatura un inventario apposito per gli strumenti musicali.</p>	NO	NO
/	INVENTARIO TOPOGRAFICO	<p>Se esiste si indica qui la sigla corrispondente al luogo (o alla struttura espositiva/conservativa: ad es. "vetrina n°.., scaffale n°..." , ecc.) occupato dallo strumento in questione all'interno della raccolta.</p>	NO	NO
/	ACQUISIZIONE	<p>Si indicano qui le modalità attraverso cui lo strumento è diventato di proprietà del museo o della raccolta: ad esempio, acquisto, dono, raccolta durante una ricerca sul campo, ecc. Nel caso in cui si tratti di informazioni riservate, è necessario assicurarsi che esse possano essere rese pubbliche.</p> <p>In ogni caso è utile indicare tutte le circostanze che si conoscono relativamente all'acquisizione, (dove è stato acquistato, quando, il prezzo pagato, chi è il donatore, quando è stato donato lo strumento, in che circostanze è stato raccolto ecc.). Se vi sono stati passaggi di proprietà e si conoscono i dati relativi, è utile riportarli. Nel caso di istituzioni pubbliche dotate di una documentazione contabile-amministrativa, è utile indicare l'esistenza e il tipo dell'eventuale documentazione relativa all'acquisizione (delibere, stime, fatture, mandati di pagamento, ecc.) e la collocazione dei documenti originali.</p>	NO	SI

SOVRACAMPI		Descrizione	Ripet.	Obbl.
	CAMPI			
/	COLLEZIONE	Nel caso in cui un nucleo di strumenti musicali si sia costituito tramite l'acquisizione di più collezioni, si indica qui la collezione di appartenenza del singolo strumento.	NO	NO
/	STRUMENTO	Si indica qui il nome italiano con il quale si designa lo strumento in questione, lasciando le varianti locali o dialettali alla voce NOME ORIGINARIO. Nel caso in cui non esista un nome in italiano, si usa il nome dialettale più diffuso (es. launeddas, indicando poi in NOME ORIGINARIO il tipo specifico). Nel caso di strumenti di provenienza non italiana si indica qui il nome in italiano e si riporta il nome nella lingua originaria alla voce NOME ORIGINARIO.	NO	SI
/	TAXON H-S	Si indica qui il taxon corrispondente allo strumento nella classificazione Hornbostel Sachs, utilizzando la traduzione pubblicata in Guizzi 2002 (Febo Guizzi, Gli strumenti della musica popolare in Italia, Lucca, LIM, 2002: 409-482). Dato che le varie suddivisioni della classificazione HS si diramano in definizioni sempre più specifiche a partire da livelli superficiali più generali, si è ritenuto opportuno indicare non solo il taxon specifico (con gli eventuali suffissi) ma anche i livelli superiori che indicano il percorso logico che porta a quel taxon. Esempi: Per le castagnette si indica: Idiofoni – Idiofoni a percussione non mediata – Idiofoni a percussione reciproca ovvero crotali – Castagnette. Per una fisarmonica si indica: Aerofoni liberi – Aerofoni ad interruzione autofonici ovvero ancie – Ancie libere in serie, con serbatoio d'aria flessibile, con tastiera.	NO	SI
/	NUMERO H-S	Si indica il numero della classificazione Hornbostel Sachs corrispondente al taxon. Per gli strumenti poliorganici si compone il numero seguendo le indicazioni date dagli autori nella prefazione alla sistematica. Esempi: Castagnette: 111.141 - Fisarmonica: 421.132 -62 -8 Zampogna a chiave: 422-62:.2]122-71+122+121 (422 indica che si tratta di tubi ad ancia; -62 è il suffisso comune che indica il serbatoio d'aria flessibile; i due punti si mettono per indicare che la suddivisione continua oltre il suffisso; il punto indica che manca una cifra e che sarà riportata dopo la parentesi quadra, in questo caso si tratta della cifra che indica che si tratta di oboi; 2 indica che si tratta di tubi ad ancia in serie; 122-71 si riferisce al chanter con la chiave e indica gli oboi a caneggio conico muniti di fori digitali: è composto dalla cifra 1 tralasciata per prima, da 2 che indica il caneggio conico e dal 2 che indica la presenza dei fori digitali; -71 è il suffisso che indica la presenza di una meccanica a chiavi. 122 si riferisce al chanter senza chiave (oboe a	NO	SI

SOVRACAMPI		Descrizione	Ripet.	Obbl.
	CAMPI			
		<p>canneggio conico con fori digitali); 121 si riferisce al bordone (oboe conico privo di fori digitali).</p> <p>Müsa: 422-62:.2]112.2+211.1 (422 indica che si tratta di tubi ad ancia; -62 è il suffisso comune che indica il serbatoio d'aria flessibile; i due punti si mettono per indicare che la suddivisione continua oltre il suffisso; il punto indica che manca una cifra e che sarà riportata dopo la parentesi quadra, in questo caso si tratta della cifra che specifica il tipo di ancia (doppia, singola o libera); 2 indica che si tratta di tubi ad ancia in serie; 112.2 si riferisce al chanter: un oboe a canneggio conico munito di fori digitali; 211.1 si riferisce al bordone: un clarinetto cilindrico privo di fori digitali.</p>		
/	DEFINIZIONE TIPOLOGICA	<p>Per definizione tipologica si intende una definizione che consente di inserire lo strumento musicale in questione in un insieme di strumenti aventi caratteristiche morfologiche e funzionali omogenee.</p> <p>Le classi di strumenti così ottenute integrano la sistematica Hornbostel Sachs poiché consentono accorpamenti o distinzioni di strumenti non previsti da essa. Per esempio la ciaramella dell'Italia centro-meridionale e il piffero delle Quattro Province, pur essendo entrambi oboi a canneggio conico e privi di chiavi, rientrano in definizioni tipologiche diverse: la ciaramella rientra nella definizione tipologica "oboe conico", il piffero rientra nella definizione "oboe conico con pirouette".</p> <p>Talvolta i nomi con cui vengono designati alcuni strumenti musicali possono costituire anche definizioni tipologiche, ad esempio "organetto" e "fisarmonica" (si veda: Anthony Baines, "Organology and European Folk Music Instruments", in International Folk Music Journal, XII, 1960: 10-13).</p> <p>La definizione tipologica può essere anche ricavata dalla stessa sistematica Hornbostel-Sachs (che va comunque consultata di volta in volta anche allo scopo di verificare questa opportunità), quando il taxon corrispondente all'apertura del "secondo livello" gerarchico-definitorio (quello stampato in corsivo, che si forma con l'apposizione del primo punto di separazione della numerazione Dewey) esprime una definizione che assume i caratteri individuanti un tipo. Esempi:</p> <p>111.1 Idiofoni a percussione reciproca ovvero crotali ove la definizione tipologica è costituita da "crotali"</p> <p>321.2 Liuti a giogo ovvero Lire ove la definizione tipologica è costituita da "Lire"</p> <p>La definizione tipologica tuttavia può essere ricavata anche da altri livelli della tassonomia, soprattutto quando la raccolta di cui ci si sta occupando</p>	NO	SI

SOVRACAMPI		Descrizione	Ripet.	Obbl.
	CAMPI			
		<p>comprenda varianti morfologiche specifiche di strumenti che richiedano una più precisa determinazione — ad es. “castagnette” in luogo di “crotali” — oppure varianti locali con precisa determinazione tipologica (esempio: il già citato caso di ciaramella e piffero, distinti dall’assenza o dalla presenza di una pirouette).</p> <p>In alcuni casi la si deve scovare a livelli molto avanzati della gerarchia, come ad es. nel caso della “campana”. In ogni caso il criterio fondamentale per la formulazione di una corretta e utile definizione tipologica è contemporaneamente di carattere empirico, astrattivo e linguistico: si tratta cioè di esaminare bene lo strumento per ricondurlo a caratteristiche generali che lo accomunino ad altri analoghi o all’interno dello stesso taxon (si veda il caso degli oboi sopra richiamati o quello di “traccola” che può comprendere la specificazione “a martelli”, “a martelli oscillanti”, “a piastra oscillante”, ecc., o può riferirsi trasversalmente a taxa diversi, come nel caso della “traccola mista” che comprenda sia piastre a battente che raschiatoi a ruota, azionati tutti dalla stessa manovella) e considerare poi gli usi linguistici correnti e quelli più specialmente organologici. Al fine di distinguere la definizione tipologica da quella sistematica, si deve intendere la prima come definizione dello strumento secondo un criterio individuante universale concreto, per lo più rispondente agli usi linguistici comuni (es.: lo strumento di Arthur Rubinstein: “pianoforte”) e la seconda, ovviamente, come definizione dello strumento secondo la classificazione organologica di Hornbostel - Sachs (es.: lo strumento di Arthur Rubinstein: “cetra a cassa di risonanza suonata a mezzo di martelletti munita di tastiera”), ma anche la distinzione tra “pianoforte”, “fortepiano”, “clavicordo” come casi specifici e rilevanti all’interno dello stesso taxon qui sopra riportato, quando non si sia proceduto all’integrazione della classificazione con la sua articolazione in profondità entro la stessa definizione sistematica (si veda l’esempio degli xilofoni riportato da Hornbostel-Sachs nella loro introduzione o si pensi ad analoghe procedure riferite alle suddette “cetre a cassa di risonanza suonate a mezzo di martelletti munite di tastiera” o alle “cetre a cassa di risonanza suonate a mezzo di plettri munite di tastiera”, distinguendovi clavicembali, spinette, virginali e così via).</p>		
/	NOME ORIGINARIO	<p>Si indica qui il nome dello strumento nella lingua originale, le varianti locali e/o dialettali. Nel caso si tratti di uno strumento costituito da più parti, è utile indicare anche il nome specifico con cui si indica ciascuna delle parti nelle zone di costruzione ed utilizzo dello strumento in questione. Ripetere il campo se si vogliono indicare diverse varianti del nome.</p>	SI	NO

SOVRACAMPI		Descrizione	Ripet.	Obbl.
CAMPI				
FABBRICAZIONE		Si indicano, se note, le informazioni relative alla costruzione dello strumento. Nel caso di strumenti costituiti dall'assemblaggio di parti fabbricate separatamente ripetere il campo, indicando la parte a cui ci si riferisce nella voce COSTRUTTORE.	SI	SI
	COSTRUTTORE flag: "ignoto"	Si indica il costruttore dello strumento, se noto. Se non si conosce il dato, si può scegliere di inserire nel campo "ignoto".	NO	SI
	LUOGO DI FABBRICAZIONE flag: "ignoto"	Si indica qui, quando è noto, il luogo dove lo strumento è stato costruito. Se non si conosce il dato, si può scegliere di inserire nel campo "ignoto".	NO	SI
	DATA/EPOCA DI FABBRICAZIONE flag: "ignota"	Si indica, se nota, l'epoca in cui lo strumento è stato costruito. Se non si conosce il dato, si può scegliere di inserire nel campo "ignota". Tuttavia occorrerebbe fare ogni volta uno sforzo di approssimazione (indicando che il dato è probabile) o di ipotesi (indicando che il dato è presunto).	NO	SI
USO		Si indicano qui, se note, le informazioni relative all'uso dello strumento. Nel caso in cui si disponga di dati relativi a molteplici usi dello strumento, è possibile ripetere il campo per raggruppare nel modo corretto i dati relativi a musicista, luogo e data/epoca d'uso. Se, per esempio, lo strumento è stato ceduto da un suonatore ad un altro è utile indicare dove ciascuno lo ha utilizzato, ove si tratti di zone diverse, e il periodo di utilizzo.	SI	SI
	SUONATORE	Si indicano il nome e i dati del suonatore o del gruppo musicale che hanno utilizzato lo strumento. Ripetere il campo se necessario.	SI	NO
	LUOGO D'USO flag: "ignoto"	Si indica qui, quando è noto, il luogo oppure l'area in cui lo strumento è stato utilizzato. Ripetere il campo se necessario. Se non si conosce il dato, si può scegliere di inserire nel campo "ignoto".	SI	SI
	DATA/EPOCA D'USO flag: "ignota"	Si indica, se nota, l'epoca in cui è stato utilizzato. Ripetere il campo se necessario. Se non si conosce il dato, si può scegliere di inserire nel campo "ignota".	SI	SI
/	CULTURA	Si indica la cultura di appartenenza dello strumento. Quando è possibile è utile indicare l'etnia. Se non si conosce il dato, si può scegliere di inserire nel campo "sconosciuta" oppure "ignota".	NO	NO
/	CARTIGLIO ISCRIZIONI flag: "assenti"	Si riporta in trascrizione diplomatica il contenuto di eventuali cartigli o iscrizioni presenti sullo strumento. Per "cartiglio o iscrizioni" si intende: marchi, etichette, targhette, bolli, timbri, incisioni relative al costruttore, a colui che ha venduto lo strumento, al suonatore, ad eventuali riparatori e/o restauratori e alla storia museale dell'oggetto. Nel caso in cui non si disponga del dato, inserire "assenti". Eventuali cifre o segni riscontrabili all'interno o	NO	SI

SOVRACAMPI		Descrizione	Ripet.	Obbl.
	CAMPI			
		all'esterno dello strumento dai quali non si ricavano informazioni sull'origine, sulla proprietà e/o sulla storia dello strumento vanno indicati alla voce DESCRIZIONE, oppure alla voce RESTAURI, se si è certi che si tratta di segni tracciati da eventuali riparatori o restauratori per necessità pratiche. Il numero di inventario corrente, redatto dal museo che possiede l'oggetto, in genere si omette.		
/	DESCRIZIONE	<p>Si redige qui una descrizione il più possibile dettagliata dell'oggetto, in modo da metterne in evidenza la morfologia, gli elementi costitutivi, le eventuali particolarità riscontrate. Vanno descritti con particolare attenzione i dettagli significativi non evidenti da una fotografia. Dalla descrizione, comunque, deve essere possibile riconoscere l'oggetto anche in assenza della fotografia. È opportuno attenersi il più possibile a termini descrittivi ricorrenti di alto contenuto tecnico – derivati dagli usi specifici, ad esempio delle tradizioni liutaria, o ricavabili dal glossario tecnico descrittivo di strutture e morfologie, quale sono quelli dell'architettura o dell'anatomia – e comunque resi omogenei e non equivoci da un'esplicita scelta convenzionale operata dai responsabili scientifici della schedatura (in tal caso l'uso della terminologia convenzionale è da ritenersi obbligatoria).</p> <p>Esempi: per oggetti o parti di oggetti sviluppati in lunghezza le estremità vanno distinte in “prossimale” e “distale” in relazione alla loro collocazione rispettiva, nell'uso, in prossimità o a distanza dal corpo dell'utilizzatore. In caso di strumenti curvi, la collocazione di specifici dettagli tecnici lungo la curvatura va distinta utilizzando i termini alternativi di “intradosso” e “estradosso”.</p>	NO	SI
/	STATO DI CONSERVAZIONE	Si riportano qui le osservazioni sullo stato di conservazione dello strumento, evitando di dare semplicemente un giudizio ma specificando le condizioni in cui lo strumento si trova al momento della schedatura, soprattutto se esso non è in buono stato di conservazione. In questa voce è utile indicare gli eventuali interventi di restauro che si ritengono necessari.	NO	SI
/	RESTAURI flag:”non rilevabili”	Si indicano in questa voce sia i restauri veri e propri, sia eventuali riparazioni eseguite dal costruttore, dal suonatore o da altri. Quando si conosce è necessario specificare chi ha eseguito l'intervento e l'epoca in cui è stato effettuato. Ripetere il campo se utile a descrivere molteplici interventi di restauro. Se non si dispone del dato inserire “non rilevabili”.	SI	SI
/	MATERIALI	Si indicano qui i materiali di cui è fatto lo strumento, specificando se possibile i tipi di legno, di canna, di metallo, di stoffa ecc.	NO	SI

SOVRACAMPI		Descrizione	Ripet.	Obbl.
	CAMPI			
/	MISURE	<p>Si indicano qui le misure in millimetri dello strumento: si devono prendere misure utili ad identificare correttamente l'oggetto e a eventuali confronti con altri strumenti analoghi. Si tratta di misure descrittive e non prescrittive: non si tratta cioè di misure destinate a guidare il lavoro dei costruttori dello strumento; esse in ogni caso devono rendere un'idea almeno sommaria delle dimensioni delle varie parti di cui è costituito lo strumento. Questo campo, che è tra quelli considerati indispensabili in qualunque tipo di scheda – anche quando sono molto diverse tra loro –, non può essere predeterminato una volta per tutte nel suo contenuto, sia perché la varietà di misure riportate cambia ovviamente da strumento a strumento, sia perché in ogni specifica schedatura si decide quanto analitiche o sommarie debbano essere le misure da effettuare per ciascun tipo di strumento. Di regola ogni collezione deve mostrare gli stessi criteri oggettivi e soggettivi di scelta e di attuazione delle misure per strumenti analoghi. L'approssimazione nella precisione delle misurazioni deve essere dichiarata: in genere ci si attiene a quanto normali strumenti di misura consentono di rilevare e quindi non è il caso di spingersi al di sotto del limite di precisione pari a 0,5mm. Alcuni dispositivi essenziali dal punto di vista acustico possono richiedere due tipi diversi di misurazione (ad es.: per le corde, “lunghezza totale” e “lunghezza vibrante”, intendendo questa come distanza tra i due punti estremi della corda libera, come capotasto e ponticello o simili: lo stesso si può fare per i membranofoni rilevando il diametro totale della membrana — quindi considerando anche la parte che poggia sul bordo del supporto e quello che lo supera per consentirne l'ancoraggio — e il diametro vibrante — misurando nei limiti del possibile la parte libera). Soprattutto per alcune classi di strumenti, in primis per gli idiofoni, è importante rilevare il peso dell'oggetto o di sue parti costitutive essenziali ai fini della produzione del suono (ad es.: le singole piastre o lamine di metallofoni, xilofoni e simili). In genere i valori di peso non sono facilmente rilevabili per l'indisponibilità di bilance adatte o per l'ingombro dello strumento o delle sue parti o per la difficoltà a separare dall'insieme ciascuna parte suscettibile di pesatura. In questi casi è opportuno segnalare nella scheda la mancanza forzata di questo tipo di misure con l'intento di segnalarne comunque la rilevanza e di sollecitare pertanto i responsabili della conservazione a provvedere alla misurazione quando si rendesse possibile (ad es. in occasione di un restauro che comporti il temporaneo smontaggio). In questo campo possono essere inserite anche considerazioni valutative sullo stato di conservazione, anche se esiste il campo apposito, soprattutto se si</p>	NO	SI

SOVRACAMPI		Descrizione	Ripet.	Obbl.
	CAMPI			
		vogliono descrivere particolari situazioni di intervento pregresso e spurio sullo strumento (es.: corde non originali, o membrana sostituita con materiale spurio o montato in modo non conforme, cavigliere incollato con angolo non rispondente alla norma tradizionale, ecc.), nonché considerazioni valutative, o rinvii alle stesse, indicate entro altro campo, sull'autenticità dello strumento stesso o sul suo livello qualitativo, poiché queste caratteristiche possono legittimamente sconsigliare l'accuratezza di misurazioni che altrimenti sarebbero di prammatica (es.: strumento falso, simulacro o modellino, versione per turisti o comunque per un mercato estraneo alla funzione culturale primaria, strumento di basso livello per uso di materiali di scarso pregio relativo, ecc.).		
/	ANNOTAZIONI	Si tratta di un campo libero destinato ad eventuali annotazioni dello schedatore, che servano a dare conto di notizie sulla storia e sulle vicende dello specifico strumento schedato, delle persone cui esso è stato legato nell'uso, delle circostanze dell'utilizzazione, su aspetti significativi della sua costruzione, ecc. (è appartenuto a..., è un tipico esemplare della tradizione costruttiva di..., è riconducibile al modello cosiddetto "Trimarco", dal nome del suonatore che per primo lo ha utilizzato in questa particolare forma, è stato usato durante i festeggiamenti per l'indipendenza del paese d'origine alla presenza di, ecc.), sia sull'uso cui strumenti analoghi sono di regola destinati, sul contesto, sulle potenzialità musicali, sulla tecnica esecutiva, sul repertorio, sulla simbologia ad essi legata, su particolarità musicali o extramusicali, su specifici aspetti della struttura costruttiva, sul nome e i particolari significati cui esso si ricollega, ecc.	NO	NO
/	ESPOSIZIONI	Si indicano qui le varie mostre nelle quali lo strumento è stato inserito ovvero alle quali è stato eventualmente prestato, specificando il luogo, la durata di ciascuna mostra e i relativi curatori. Ripetere il campo se utile a descrivere molteplici esposizioni.	SI	NO
/	BIBLIOGRAFIA	Si indica qui la bibliografia specifica sullo strumento in questione, inteso sia come oggetto specifico della raccolta, se esiste qualche scritto che lo riguardi, sia come tipo di strumento in generale.	NO	SI
/	RISORSE	Si rende conto qui, quando possibile, della presenza, presso il museo o altrove, di altre schede relative allo strumento, eventuali disegni, radiografie, fotografie, misurazioni dettagliate, ecc. Se il museo o la raccolta possiedono un archivio fotografico contenente le foto degli strumenti di loro proprietà si rinvia qui alla foto corrispondente allo strumento in questione. È obbligatorio allegare almeno un'immagine recente e rappresentativa dello strumento,	NO	SI

SOVRACAMPI		Descrizione	Ripet.	Obbl.
CAMPI				
		possibilmente acquisita in fase di catalogazione. Si rammenta qui che è buona norma per musei e collezioni che mettano a disposizione di terzi strumenti delle loro raccolte, esigere che copia della documentazione eventualmente prodotta sia depositata presso l'ente o il soggetto proprietario.		
COMPILATORE DELLA SCHEDA		Si riportano qui i dati relativi alla creazione della scheda.	NO	SI
	AUTORE COMPILAZIONE	Nome del compilatore della scheda, utile se più schedatori collaborano alla schedatura di una raccolta di oggetti. Ripetere il campo nel caso in cui più autori abbiano collaborato alla compilazione della scheda.	SI	SI
	DATA COMPILAZIONE	Data in cui è stata redatta la scheda.	NO	SI
	DIRITTI SCHEDA	Si inseriscono le informazioni su diritti d'autore e tutela della proprietà intellettuale relative alla scheda catalografica (ad esempio: Copyright, licenze Creative Commons, pubblico dominio ecc...).	NO	SI
AGGIORNAMENTI		Si riportano qui i dati relativi agli aggiornamenti e alle modifiche apportate, nel corso del tempo, alle successive versioni della scheda.	SI	NO
	AUTORE AGGIORNAMENTI	Nome dell'autore dell'aggiornamento. Ripetere il campo nel caso in cui più autori abbiano collaborato all'aggiornamento della scheda.	SI	SI*
	DATA AGGIORNAMENTI	Data in cui sono state apportate le modifiche.	NO	SI*
	COMMENTI	Si elencano le aggiunte e le modifiche apportate alla versione precedente della scheda: per ogni modifica si inserisce il nome del campo in questione e una breve descrizione dell'aggiornamento.	NO	SI*
SI*: campi obbligatori relativi (solo se si apre un sovracampo "aggiornamenti")				
Copyright Giorgio Bevilacqua, Cristina Ghirardini e Ilario Meandri, 2016. Quest'opera è stata rilasciata con licenza Creative Commons Attribuzione - Non commerciale 4.0 Internazionale. Per leggere una copia della licenza visita il sito web http://creativecommons.org/licenses/by-nc/4.0/ .				

3. Istruzioni per una mappatura Guizzi-MIMO

Le istruzioni riportate nella tabella descrivono come organizzare una mappatura dei contenuti dal modello `scheda guizzi 1_0.xsd` al modello `lido-v1.0.xsd` adottato da MIMO.

Si ricorda che per formare un record MIMO valido è necessario compilare una serie di elementi obbligatori, alcuni dei quali non sono presenti nella Scheda Guizzi, e quindi non compaiono nella mappatura. Per la consultazione degli elementi obbligatori si rimanda al capitolo “MIMO data model organization” e soprattutto alle specifiche tecniche ufficialmente pubblicate da MIMO (Le Meur, 2010; Bailly e Le Meur, 2010).

Nella redazione della tabella sono state seguite alcune convenzioni:

- la colonna “SCHEDA GUIZZI” indica i nomi dei campi di `scheda guizzi 1_0.xsd` di cui è stato mantenuto quasi integralmente l’ordine originale;
- “MIMO data model organization” descrive l’intero percorso di destinazione delle informazioni nel modello LIDO;
- “CAMPI GENERALI” si riferisce al “MIMO data model organization” (cfr. *infra* p. 86, fig. 3);
- “PERCORSO XML” contiene gli elementi che costituiscono la parte finale dell’annidamento XML (al termine del quale vanno inseriti i contenuti). I percorsi proseguono quelli dichiarati in “CAMPI GENERALI”; sono privi di namespace `lido:` e sono separati da uno *slash*;
- per alcuni elementi sono stati indicati gli attributi, indicati convenzionalmente tra parentesi quadre, ad es. `[type=”current”]`;
- il contenuto di alcuni elementi, scelto da una lista controllata, è indicato tra virgolette, ad es. “acquisition”;
- la colonna “obbl. MIMO” indica gli elementi della Scheda Guizzi obbligatori per MIMO.

Percorso dei “CAMPI GENERALI”:

- records: `lidoWrap/lido/administrativeMetadata/recordWrap`
- repository: `lidoWrap/lido/descriptiveMetadata/objectIdentificationWrap/repositoryWrap`
- events: `lidoWrap/lido/descriptiveMetadata/eventWrap`
- instrument description:
 - `lidoWrap/lido/descriptiveMetadata/objectIdentificationWrap/objectDescriptionWrap`
- title/names: `lidoWrap/lido/descriptiveMetadata/objectIdentificationWrap/titleWrap`
- classifications: `lidoWrap/lido/descriptiveMetadata/objectClassificationWrap/classificationWrap`
- inscriptions: `lidoWrap/lido/descriptiveMetadata/objectIdentificationWrap/inscriptionWrap`
- measurements: `lidoWrap/lido/descriptiveMetadata/objectIdentificationWrap/objectMeasurementsWrap`
- resources: `lidoWrap/lido/administrativeMetadata/resourceWrap`

SCHEDE GUIZZI	MIMO data model organization		obbl. MIMO
	CAMPO GENERALE	PERCORSO XML	
SCHEDE	records	.../recordID[type="local"]	X
COLLOCAZIONE	repository	.../repositorySet[type="current"]/repositoryName/legalBodyName/appellationValue	X
		.../repositorySet[type="current"]/repositoryName/legalBodyWeblink	
		.../repositorySet[type="current"]/repositoryLocation/appellationValue	
INVENTARIO GENERALE	repository	.../repositorySet/workID[type="inventory number"]	X
INVENTARIO PARTICOLARE	repository	.../repositorySet/workID[type="special inventory number"]	
INVENTARIO TOPOGRAFICO	repository	.../repositorySet/workID[type="topographic inventory number"]	
ACQUISIZIONE ³	events	.../eventSet/displayEvent	
		.../eventSet/event/eventType/term "acquisition" [xml:lang="en"]	
COLLEZIONE	instrument description	.../objectDescriptionSet[type="general description"]/descriptiveNoteValue	
STRUMENTO	title/names	.../titleSet/appellationValue[pref="preferred"][xml:lang="it"]	X
		.../titleSet/sourceAppellation	
TAXON H-S	classifications	.../classification/term[pref="alternate"][xml:lang="it"]	
NUMERO H-S	classifications	.../classification/conceptID[type="SH_Class"]	
DEFINIZIONE TIPOLOGICA ⁴	classifications	.../classification/term[pref="preferred"]	X
NOME ORIGINARIO ⁵	title/names	.../titleSet/appellationValue[pref="alternate"][xml:lang=""]	
		.../titleSet/sourceAppellation	
FABBRICAZIONE	events "production"	.../eventSet/displayEvent	
COSTRUTTORE		.../eventSet/event/eventType/term "production" [xml:lang="en"]	
DATA/EPOCA DI FABBRICAZIONE		.../eventSet/event/eventType/eventActor/displayActorInRole	
LUOGO DI FABBRICAZIONE		.../eventSet/event/eventDate/displayDate	
		.../eventSet/event/eventPlace/displayPlace	
MATERIALI ⁶		.../eventSet/event/eventMaterialsTech/displayMaterialsTech	

3 Tutti i campi MIMO "events" possono utilizzare sia l'elemento *display* che tutti gli elementi messi a disposizione da LIDO per le descrizioni granulari degli eventi. In queste istruzioni abbiamo indicato solo i campi MIMO che, al momento, ci sembra opportuno compilare. Questo non esclude che si possa fare altrimenti; nell'esempio di record MIMO riportato in Appendice 5, molti elementi XML delle descrizioni di eventi sono dichiarati e non compilati.

4 Il `lido:term` indicato deve coincidere con un termine del vocabolario di *keyword* redatto MIMO; è possibile proporre a MIMO nuove *keywords* che soddisfino i criteri della definizione tipologica.

5 Nel caso di nomi in lingua diversa da quella impostata generalmente per i descriptive metadata, specificarlo con l'attributo `xml:lang`.

6 Il campo è stato spostato in questa posizione, diversa rispetto al modello originale, e va inserito all'interno di un `lido:eventSet` di tipo "production".

SCHEDE GUIZZI	MIMO data model organization		obbl. MIMO
	CAMPO GENERALE	PERCORSO XML	
USO	events "use"	.../eventSet/displayEvent	
SUONATORE		.../eventSet/event/eventType/term "use" [xml:lang="en"]	
DATA/EPOCA D'USO		.../eventSet/event/eventType/eventActor/displayActorInRole	
LUOGO D'USO		.../eventSet/event/eventDate/displayDate	
CULTURA ⁷		.../eventSet/event/eventPlace/displayPlace	
		.../eventSet/event/culture/term	
ISCRIZIONI	inscriptions	.../inscriptions/inscriptionDescription/descriptiveNoteValue	
		.../inscriptions/inscriptionTranscription	
DESCRIZIONE	instrument description	.../objectDescriptionSet[type="general description"]/descriptiveNoteValue	
STATO DI CONSERVAZIONE	instrument description	.../objectDescriptionSet[type="general description"]/descriptiveNoteValue	
RESTAURI	events "restoration"	.../eventSet/displayEvent	
		.../eventSet/event/eventType/term "restoration" [xml:lang="en"]	
MISURE	measurements	.../objectMeasurementsSet/displayObjectMeasurements	
ANNOTAZIONI ⁸	instrument description	.../objectDescriptionSet[type="general description"]/descriptiveNoteValue	
		.../objectDescriptionSet[type="context"]/descriptiveNoteValue	
		.../objectDescriptionSet[type="decorative features"]/descriptiveNoteValue	
		.../objectDescriptionSet[type="performance characteristics"]/descriptiveNoteValue	
ESPOSIZIONI	events "exhibition"	.../eventSet/displayEvent	
		.../eventSet/event/eventType/term "exhibition" [xml:lang="en"]	
		.../eventSet/event/eventName/appellationValue	
		.../eventSet/event/eventDate/displayDate	
		.../eventSet/event/eventPlace/displayPlace	
BIBLIOGRAFIA ⁹	instrument description	.../objectDescriptionSet[type="literature references"]/descriptiveNoteValue	

7 Il campo è stato spostato in questa posizione, diversa rispetto al modello originale, e va inserito all'interno di un `lido:eventSet` di tipo "use".

8 Le annotazioni vanno inserite in uno o più campi `lido:objectDescriptionSet` con attributo `lido:type` diverso a seconda del contenuto.

9 La bibliografia va indicata in un campo di testo libero all'interno di `lido:objectDescriptionSet`.

SCHEDA GUIZZI	MIMO data model organization		obbl. MIMO
	CAMPO GENERALE	PERCORSO XML	
RISORSE ¹⁰	resources	.../resourceSet/resourceID[type="local"]	
		.../resourceSet/resourceRepresentation/linkResource	
		.../resourceSet/resourceType/term	
		.../resourceSet/resourceDescription	
		.../resourceSet/rightsresource/rightsType/term	
		.../resourceSet/rightsresource/rightsDate/earliestDate	
		.../resourceSet/rightsresource/rightsDate/latestDate	
		.../resourceSet/rightsResource/rightsHolder/legalBodyName/appellationValue	
		.../resourceSet/rightsResource/rightsHolder/legalBodyWeblink	
	instrument description ¹¹	.../objectDescriptionSet[type="literature references"]/descriptiveNoteValue	
AUTORE COMPILAZIONE	records ¹²	.../recordSource/legalBodyName	
		.../recordInfoSet/recordMetadataDate	
DATA COMPILAZIONE	records	.../recordInfoSet/recordMetadataDate	
AUTORE AGGIORNAMENTI		.../recordInfoSet/recordMetadataDate	
DATA AGGIORNAMENTI		.../recordInfoSet/recordMetadataDate	
COMMENTI		.../recordInfoSet/recordMetadataDate	
DIRITTI SCHEDA	records	.../recordRights/rightsType/term	
		.../recordRights/rightsDate/earliestDate	
		.../recordRights/rightsDate/latestDate	
		.../recordRights/rightsHolder/legalBodyName/appellationValue	
		.../recordRights/rightsHolder/legalBodyWeblink	

10 MIMO richiede che si fornisca almeno un'immagine dello strumento, considerata l'immagine principale. Essa deve essere contrassegnata dall'attributo `lido:pref="preferred"` assegnato a `lido:resourceID`. È possibile allegare altre risorse del tipo "image", "sound", "video" (da specificare nell'elemento `lido:resourceType`).

11 Nel caso in cui si voglia allegare un testo, per esempio una didascalia, come risorsa documentale, è possibile inserirlo nel campo di testo libero in cui si inserisce la bibliografia.

12 LIDO richiede obbligatoriamente la compilazione del campo `lido:recordSource`, che è espresso da un `lido:legalBody` e indica l'istituzione che fornisce il record (che è di solito la stessa istituzione che custodisce lo strumento). `lido:recordSource` è un campo ripetibile e può essere usato per segnalare, sempre tramite `lido:legalBodyName`, i nomi degli autori (della scheda originale e degli aggiornamenti).

Tutte le informazioni relative ai nomi degli autori, alle date e al tipo di modifiche apportate negli aggiornamenti vanno inserite, sotto forma di testo libero, anche nel campo `lido:recordMetadataDate`.

4. Esempio di record Guizzi (versione 1.0)

Si riporta qui il testo di:

sonaglio_per_bambini_2_guizzi.xml

Il file XML è pubblicato all'indirizzo:

<http://www.suonoemagine.unito.it/bevilacqua.html>

```
<record>
  <scheda>2</scheda>
  <collocazione>Civico Museo del Paesaggio Sonoro di Riva presso Chieri. Sala n.
  1</collocazione>
  <inventarioGenerale>0202SM</inventarioGenerale>
  <inventarioParticolare></inventarioParticolare>
  <inventarioTopografico>Sala n.1, vetrina n. 1.1</inventarioTopografico>
  <acquisizione>Dono di Anna Torta</acquisizione>
  <collezione>Collezione di Domenico Torta</collezione>
  <strumento>Sonaglio per bambini</strumento>
  <taxonHS>Idiofoni – Idiofoni a percussione mediata – Idiofoni a scuotimento o
  crepitacoli – Crepitacoli in filze – Crepitacoli su corda</taxonHS>
  <numeroHS>112.111</numeroHS>
  <definizioneTipologica>Crepitacoli in filze</definizioneTipologica>
  <nomeOriginario>La scheda redatta da Anna Torta e conservata presso il Civico
  Museo del Paesaggio sonoro riporta la denominazione giargiatora (sistema di
  trascrizione semplificato secondo la grafia italiana di Glauco Sanga:
  giargiàtuja). Il termine non indica uno strumento musicale preciso, ma significa
  “gingillo”, “giocattolo” nel senso di cosa di poca importanza.</nomeOriginario>
  <fabbricazione>
    <costruttore>Giuseppina Tamagnone (Riva presso Chieri, 1925) l’ha costruito
    per la nipote Anna Torta, nata il 25 febbraio 1983</costruttore>
    <luogoDiFabbricazione>Riva presso Chieri</luogoDiFabbricazione>
    <dataEpocaDiFabbricazione>Anni '80 del XX secolo</dataEpocaDiFabbricazione>
  </fabbricazione>
  <uso>
    <suonatore></suonatore>
    <luogoUso>Riva presso Chieri</luogoUso>
    <dataEpocaUso>Anni '80 del XX secolo</dataEpocaUso>
  </uso>
  <cultura></cultura>
  <cartiglioIscrizioni>Assenti</cartiglioIscrizioni>
```

```
<descrizione>Crepitacolo costituito da una serie di bottoni di plastica infilati in
un filo di spago, le cui estremità sono state annodate. I bottoni sono di colore
nero, bianco, marrone e blu e sono di diametro leggermente diverso</descrizione>
<statoDiConservazione>Buono</statoDiConservazione>
<restauri>Non rilevabili</restauri>
<materiali>Spago, plastica</materiali>
<misure>Diametro minimo dei bottoni: 24,2 mm
Diametro massimo dei bottoni: 27,8 mm
Lunghezza totale del filo: 960 mm ca</misure>
<annotazioni>Costruito dalla nonna, Giuseppina Tamagnone, per la nipote Anna
Torta, come giocattolo sonoro che produce suono tendendo il filo e facendo scorrere
i bottoni. Successivamente Anna ha imparato a contare con questo
strumento.</annotazioni>
<esposizioni>Agricoltore Artigiano Artista, Riva presso Chieri, Palazzo Grosso,
16-30 aprile 1995</esposizioni>
<esposizioni>Seconda edizione del Festival Chierinmusica, Chieri,
1995</esposizioni>
<esposizioni>X Rescontre Occitan, Limone Piemonte, in collaborazione con Roberto
Tombesi, 1996</esposizioni>
<esposizioni>Identità e differenza, Torino, Sermig, 1996</esposizioni>
<esposizioni>Virtuosi senza pretese, Entracque, Centro Visite del Parco, 24 luglio
- 5 settembre 2004</esposizioni>
<esposizioni>Virtuosi senza pretese, Rovigo, 2004</esposizioni>
<bibliografia>Guizzi, Febo (2002). Gli strumenti della musica popolare in Italia.
Lucca: Lim Editrice.</bibliografia>
<risorse>0202_SM.tiff</risorse>
<compilatoreDellaScheda>
  <autoreCompilazione>Cristina Ghirardini</autoreCompilazione>
  <dataCompilazione>Aprile 2007</dataCompilazione>
  <dirittiScheda>This work is licensed under the Creative Commons Attribution 4.0
  International License. To view a copy of this license, visit
  http://creativecommons.org/licenses/by/4.0/ or send a letter to Creative
  Commons, PO Box 1866, Mountain View, CA 94042, USA</dirittiScheda>
</compilatoreDellaScheda>
<aggiornamenti>
<autoreAggiornamenti></autoreAggiornamenti>
<dataAggiornamenti></dataAggiornamenti>
<commenti></commenti>
</aggiornamenti>
</record>
```

5. Esempio di record MIMO

Si riporta qui il testo di:

sonaglio_per_bambini_2_MIMO.xml

Il file XML è pubblicato all'indirizzo:

<http://www.suonoemagine.unito.it/bevilacqua.html>

```
<lido:lidoWrap xsi:schemaLocation="http://www.lido-schema.org http://www.lido-
schema.org/schema/v1.0/lido-v1.0.xsd" xmlns:lido="http://www.lido-schema.org"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <lido:lido>
    <lido:lidoRecID lido:type="local">MPS:2</lido:lidoRecID>
    <lido:descriptiveMetadata xml:lang="it">
      <lido:objectClassificationWrap>
        <lido:objectWorkTypeWrap>
          <lido:objectWorkType>
            <lido:term xml:lang="en">musical instruments</lido:term>
          </lido:objectWorkType>
        </lido:objectWorkTypeWrap>
      <lido:classificationWrap>
        <lido:classification>
          <lido:term lido:pref="preferred">crepitacoli in filze</lido:term>
        </lido:classification>
        <lido:classification>
          <lido:conceptID lido:type="SH_Class">112.111</lido:conceptID>
          <lido:term lido:pref="alternate" xml:lang="it">Idiofoni – Idiofoni a
percussione mediata – Idiofoni a scuotimento o crepitacoli – Crepitacoli in filze –
Crepitacoli su corda</lido:term>
        </lido:classification>
      </lido:classificationWrap>
    </lido:objectClassificationWrap>
    <lido:objectIdentificationWrap>
      <lido:titleWrap>
        <lido:titleSet>
          <lido:appellationValue lido:pref="preferred" xml:lang="it">sonaglio
per bambini</lido:appellationValue>
        </lido:titleSet>
      </lido:titleWrap>
    </lido:objectIdentificationWrap>
  </lido:lido>
</lido:lidoWrap>
```

```

        <lido:appellationValue
lido:pref="alternate">giargiatore</lido:appellationValue>
        <lido:sourceAppellation>La scheda redatta da Anna Torta e conservata
presso il Civico Museo del Paesaggio sonoro riporta la denominazione giargiatore
(sistema di trascrizione semplificato secondo la grafia italiana di Glauco Sanga:
giargiàtuJa). Il termine non indica uno strumento musicale preciso, ma significa
"gingillo", "giocattolo" nel senso di cosa di poca importanza.</lido:sourceAppellation>
        </lido:titleSet>
</lido:titleWrap>
<lido:inscriptionsWrap>
    <lido:inscriptions>
        <lido:inscriptionTranscription></lido:inscriptionTranscription>
        <lido:inscriptionDescription>
            <lido:descriptiveNoteValue>iscrizioni
assenti</lido:descriptiveNoteValue>
        </lido:inscriptionDescription>
    </lido:inscriptions>
</lido:inscriptionsWrap>
<lido:repositoryWrap>
    <lido:repositorySet lido:type="current">
        <lido:repositoryName>
            <lido:legalBodyName>
                <lido:appellationValue>Civico Museo del Paesaggio Sonoro di Riva
presso Chieri</lido:appellationValue>
            </lido:legalBodyName>
<lido:legalBodyWeblink>http://www.museopaesaggiosonoro.org/</lido:legalBodyWeblink>
        </lido:repositoryName>
        <lido:workID lido:type="inventory number">0202SM</lido:workID>
        <lido:workID lido:type="special inventory number"></lido:workID>
        <lido:workID lido:type="topographic inventory number">Sala n.1,
vetrina n. 1.1</lido:workID>
    <lido:repositoryLocation>
        <lido:namePlaceSet>
            <lido:appellationValue>Riva presso Chieri</lido:appellationValue>
        </lido:namePlaceSet>
    </lido:repositoryLocation>
</lido:repositorySet>
</lido:repositoryWrap>
<lido:objectDescriptionWrap>
    <lido:objectDescriptionSet lido:type="general description">

```

```

        <lido:descriptiveNoteValue>Strumento appartenente alla collezione di
Domenico Torta. Crepitacolo costituito da una serie di bottoni di plastica infilati in
un filo di spago, le cui estremità sono state annodate. I bottoni sono di colore nero,
bianco, marrone e blu e sono di diametro leggermente diverso. Stato di conservazione
buono.</lido:descriptiveNoteValue>
    </lido:objectDescriptionSet>
    <lido:objectDescriptionSet lido:type="context">
        <lido:descriptiveNoteValue>Costruito dalla nonna, Giuseppina
Tamagnone, per la nipote Anna Torta. Successivamente Anna ha imparato a contare con
questo strumento.</lido:descriptiveNoteValue>
    </lido:objectDescriptionSet>
    <lido:objectDescriptionSet lido:type="decorative features">
        <lido:descriptiveNoteValue></lido:descriptiveNoteValue>
    </lido:objectDescriptionSet>
    <lido:objectDescriptionSet lido:type="performance characteristics">
        <lido:descriptiveNoteValue>Giocattolo sonoro che produce suono
tendendo il filo e facendo scorrere i bottoni.</lido:descriptiveNoteValue>
    </lido:objectDescriptionSet>
    <lido:objectDescriptionSet lido:type="literature references">
        <lido:descriptiveNoteValue>Guizzi, Febo (2002). Gli strumenti della
musica popolare in Italia. Lucca: Lim Editrice</lido:descriptiveNoteValue>
    </lido:objectDescriptionSet>
</lido:objectDescriptionWrap>
<lido:objectMeasurementsWrap>
    <lido:objectMeasurementsSet>
        <lido:displayObjectMeasurements>Diametro minimo dei bottoni: 24,2 mm /
Diametro massimo dei bottoni: 27,8 mm / Lunghezza totale del filo: 960 mm
ca</lido:displayObjectMeasurements>
    </lido:objectMeasurementsSet>
</lido:objectMeasurementsWrap>
</lido:objectIdentificationWrap>
<lido:eventWrap>
    <lido:eventSet>
        <lido:displayEvent>Dono di Anna Torta</lido:displayEvent>
    <lido:event>
        <lido:eventType>
            <lido:term xml:lang="en">acquisition</lido:term>
        </lido:eventType>
        <lido:eventName>
            <lido:appellationValue></lido:appellationValue>
        </lido:eventName>
        <lido:eventActor>

```

```

    <lido:displayActorInRole></lido:displayActorInRole>
  </lido:eventActor>
  <lido:culture>
    <lido:term></lido:term>
  </lido:culture>
<lido:eventDate>
  <lido:displayDate></lido:displayDate>
</lido:eventDate>
<lido:eventPlace>
  <lido:displayPlace></lido:displayPlace>
</lido:eventPlace>
<lido:eventMaterialsTech>
  <lido:displayMaterialsTech></lido:displayMaterialsTech>
</lido:eventMaterialsTech>
</lido:event>
</lido:eventSet>
<lido:eventSet>
  <lido:displayEvent>Giuseppina Tamagnone (Riva presso Chieri, 1925) l'ha
costruito per la nipote Anna Torta, nata il 25 febbraio 1983.</lido:displayEvent>
  <lido:event>
    <lido:eventType>
      <lido:term xml:lang="en">production</lido:term>
    </lido:eventType>
    <lido:eventName>
      <lido:appellationValue></lido:appellationValue>
    </lido:eventName>
    <lido:eventActor>
      <lido:displayActorInRole>Giuseppina Tamagnone (Riva presso Chieri,
1925)</lido:displayActorInRole>
    </lido:eventActor>
    <lido:culture>
      <lido:term></lido:term>
    </lido:culture>
    <lido:eventDate>
      <lido:displayDate>Anni '80 del XX secolo</lido:displayDate>
    </lido:eventDate>
    <lido:eventPlace>
      <lido:displayPlace>Riva presso Chieri</lido:displayPlace>
    </lido:eventPlace>
    <lido:eventMaterialsTech>
      <lido:displayMaterialsTech>spago,
plastica</lido:displayMaterialsTech>
  </lido:event>
</lido:eventSet>

```

```

        </lido:eventMaterialsTech>
    </lido:event>
</lido:eventSet>
<lido:eventSet>
    <lido:displayEvent></lido:displayEvent>
    <lido:event>
        <lido:eventType>
            <lido:term xml:lang="en">use</lido:term>
        </lido:eventType>
        <lido:eventName>
            <lido:appellationValue></lido:appellationValue>
        </lido:eventName>
        <lido:eventActor>
            <lido:displayActorInRole></lido:displayActorInRole>
        </lido:eventActor>
        <lido:culture>
            <lido:term></lido:term>
        </lido:culture>
        <lido:eventDate>
            <lido:displayDate>Anni '80 del XX secolo</lido:displayDate>
        </lido:eventDate>
        <lido:eventPlace>
            <lido:displayPlace>Riva presso Chieri</lido:displayPlace>
        </lido:eventPlace>
        <lido:eventMaterialsTech>
            <lido:displayMaterialsTech></lido:displayMaterialsTech>
        </lido:eventMaterialsTech>
    </lido:event>
</lido:eventSet>
<lido:eventSet>
    <lido:displayEvent>restauri non rilevabili</lido:displayEvent>
    <lido:event>
        <lido:eventType>
            <lido:term xml:lang="en">restoration</lido:term>
        </lido:eventType>
        <lido:eventName>
            <lido:appellationValue></lido:appellationValue>
        </lido:eventName>
        <lido:eventActor>
            <lido:displayActorInRole></lido:displayActorInRole>
        </lido:eventActor>
        <lido:culture>

```

```

        <lido:term></lido:term>
    </lido:culture>
<lido:eventDate>
    <lido:displayDate></lido:displayDate>
</lido:eventDate>
<lido:eventPlace>
    <lido:displayPlace></lido:displayPlace>
</lido:eventPlace>
<lido:eventMaterialsTech>
    <lido:displayMaterialsTech></lido:displayMaterialsTech>
</lido:eventMaterialsTech>
</lido:event>
</lido:eventSet>
<lido:eventSet>
    <lido:displayEvent>Agricoltore Artigiano Artista, Riva presso Chieri,
Palazzo Grosso, 16-30 aprile 1995</lido:displayEvent>
    <lido:event>
        <lido:eventType>
            <lido:term xml:lang="en">exhibition</lido:term>
        </lido:eventType>
        <lido:eventName>
            <lido:appellationValue>Agricoltore Artigiano
Artista</lido:appellationValue>
        </lido:eventName>
        <lido:eventActor>
            <lido:displayActorInRole></lido:displayActorInRole>
        </lido:eventActor>
        <lido:culture>
            <lido:term></lido:term>
        </lido:culture>
        <lido:eventDate>
            <lido:displayDate>16-30 aprile 1995</lido:displayDate>
        </lido:eventDate>
        <lido:eventPlace>
            <lido:displayPlace>Riva presso Chieri, Palazzo
Grosso</lido:displayPlace>
        </lido:eventPlace>
        <lido:eventMaterialsTech>
            <lido:displayMaterialsTech></lido:displayMaterialsTech>
        </lido:eventMaterialsTech>
    </lido:event>
</lido:eventSet>

```

```

    <lido:eventSet>
      <lido:displayEvent>Seconda edizione del Festival Chierinmusica, Chieri,
1995</lido:displayEvent>
      <lido:event>
        <lido:eventType>
          <lido:term xml:lang="en">exhibition</lido:term>
        </lido:eventType>
        <lido:eventName>
          <lido:appellationValue>Seconda edizione del Festival
Chierinmusica</lido:appellationValue>
        </lido:eventName>
        <lido:eventActor>
          <lido:displayActorInRole></lido:displayActorInRole>
        </lido:eventActor>
        <lido:culture>
          <lido:term></lido:term>
        </lido:culture>
        <lido:eventDate>
          <lido:displayDate>1995</lido:displayDate>
        </lido:eventDate>
        <lido:eventPlace>
          <lido:displayPlace>Chieri</lido:displayPlace>
        </lido:eventPlace>
        <lido:eventMaterialsTech>
          <lido:displayMaterialsTech></lido:displayMaterialsTech>
        </lido:eventMaterialsTech>
      </lido:event>
    </lido:eventSet>
    <lido:eventSet>
      <lido:displayEvent>X Rescontre Occitan, Limone Piemonte, in
collaborazione con Roberto Tombesi, 1996</lido:displayEvent>
      <lido:event>
        <lido:eventType>
          <lido:term xml:lang="en">exhibition</lido:term>
        </lido:eventType>
        <lido:eventName>
          <lido:appellationValue>X Rescontre Occitan</lido:appellationValue>
        </lido:eventName>
        <lido:eventActor>
          <lido:displayActorInRole></lido:displayActorInRole>
        </lido:eventActor>
        <lido:culture>

```

```
        <lido:term></lido:term>
    </lido:culture>
<lido:eventDate>
    <lido:displayDate>1996</lido:displayDate>
</lido:eventDate>
<lido:eventPlace>
    <lido:displayPlace>Limone Piemonte</lido:displayPlace>
</lido:eventPlace>
<lido:eventMaterialsTech>
    <lido:displayMaterialsTech></lido:displayMaterialsTech>
</lido:eventMaterialsTech>
</lido:event>
</lido:eventSet>
<lido:eventSet>
    <lido:displayEvent>Identità e differenza, Torino, Sermig,
1996</lido:displayEvent>
    <lido:event>
        <lido:eventType>
            <lido:term xml:lang="en">exhibition</lido:term>
        </lido:eventType>
        <lido:eventName>
            <lido:appellationValue>Identità e differenza</lido:appellationValue>
        </lido:eventName>
        <lido:eventActor>
            <lido:displayActorInRole></lido:displayActorInRole>
        </lido:eventActor>
        <lido:culture>
            <lido:term></lido:term>
        </lido:culture>
        <lido:eventDate>
            <lido:displayDate>1996</lido:displayDate>
        </lido:eventDate>
        <lido:eventPlace>
            <lido:displayPlace>Torino, Sermig</lido:displayPlace>
        </lido:eventPlace>
        <lido:eventMaterialsTech>
            <lido:displayMaterialsTech></lido:displayMaterialsTech>
        </lido:eventMaterialsTech>
    </lido:event>
</lido:eventSet>
<lido:eventSet>
```

<lido:displayEvent>Virtuosi senza pretese, Entracque, Centro Visite del Parco, 24 luglio – 5 settembre 2004</lido:displayEvent>

<lido:event>

<lido:eventType>

<lido:term xml:lang="en">exhibition</lido:term>

</lido:eventType>

<lido:eventName>

<lido:appellationValue>Virtuosi senza pretese</lido:appellationValue>

</lido:eventName>

<lido:eventActor>

<lido:displayActorInRole></lido:displayActorInRole>

</lido:eventActor>

<lido:culture>

<lido:term></lido:term>

</lido:culture>

<lido:eventDate>

<lido:displayDate>24 luglio - 5 settembre 2004</lido:displayDate>

</lido:eventDate>

<lido:eventPlace>

<lido:displayPlace>Entracque, Centro Visite del Parco</lido:displayPlace>

</lido:eventPlace>

<lido:eventMaterialsTech>

<lido:displayMaterialsTech></lido:displayMaterialsTech>

</lido:eventMaterialsTech>

</lido:event>

</lido:eventSet>

<lido:eventSet>

<lido:displayEvent>Virtuosi senza pretese, Rovigo, 2004</lido:displayEvent>

<lido:event>

<lido:eventType>

<lido:term xml:lang="en">exhibition</lido:term>

</lido:eventType>

<lido:eventName>

<lido:appellationValue>Virtuosi senza pretese</lido:appellationValue>

</lido:eventName>

<lido:eventActor>

<lido:displayActorInRole></lido:displayActorInRole>

</lido:eventActor>

```

        <lido:culture>
            <lido:term></lido:term>
        </lido:culture>
    <lido:eventDate>
        <lido:displayDate>2004</lido:displayDate>
    </lido:eventDate>
    <lido:eventPlace>
        <lido:displayPlace>Rovigo</lido:displayPlace>
    </lido:eventPlace>
    <lido:eventMaterialsTech>
        <lido:displayMaterialsTech></lido:displayMaterialsTech>
    </lido:eventMaterialsTech>
</lido:event>
</lido:eventSet>
</lido:eventWrap>
</lido:descriptiveMetadata>
<lido:administrativeMetadata xml:lang="it">
    <lido:rightsWorkWrap>
        <lido:rightsWorkSet>
            <lido:rightsType>
                <lido:term>copyrights</lido:term>
            </lido:rightsType>
        </lido:rightsWorkSet>
    </lido:rightsWorkWrap>
<lido:recordWrap>
    <lido:recordID lido:type="local">2</lido:recordID>
    <lido:recordType>
        <lido:term>item</lido:term>
    </lido:recordType>
    <lido:recordSource>
        <lido:legalBodyID lido:type="local">MPS</lido:legalBodyID>
        <lido:legalBodyName>
            <lido:appellationValue>Civico Museo del Paesaggio Sonoro di Riva
presso Chieri</lido:appellationValue>
        </lido:legalBodyName>

<lido:legalBodyWeblink>http://www.museopaesaggiosonoro.org/</lido:legalBodyWeblink>
    </lido:recordSource>
<lido:recordSource>
    <lido:legalBodyID lido:type=""></lido:legalBodyID>
    <lido:legalBodyName>
        <lido:appellationValue>Cristina Ghirardini</lido:appellationValue>

```

```

        </lido:legalBodyName>
        <lido:legalBodyWeblink></lido:legalBodyWeblink>
    </lido:recordSource>
    <lido:recordRights>
        <lido:rightsType>
            <lido:term>Copyright, Civico Museo del Paesaggio Sonoro, 2007. This
work is licensed under the Creative Commons Attribution 4.0 International License. To
view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/ or send a
letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA</lido:term>
        </lido:rightsType>
        <lido:rightsDate>
            <lido:earliestDate></lido:earliestDate>
            <lido:latestDate></lido:latestDate>
        </lido:rightsDate>
        <lido:rightsHolder>
            <lido:legalBodyName>
                <lido:appellationValue>Civico Museo del Paesaggio Sonoro di Riva
presso Chieri</lido:appellationValue>
            </lido:legalBodyName>

<lido:legalBodyWeblink>http://www.museopaesaggiosonoro.org/</lido:legalBodyWeblink>
        </lido:rightsHolder>
    </lido:recordRights>
    <lido:recordInfoSet>
        <lido:recordInfoLink>link al record sul sito
locale</lido:recordInfoLink>
        <lido:recordMetadataDate>Cristina Ghirardini, aprile
2007</lido:recordMetadataDate>
    </lido:recordInfoSet>
</lido:recordWrap>
<lido:resourceWrap>
    <lido:resourceSet>
        <lido:resourceID lido:pref="preferred"
lido:type="local">0202_SM.tiff</lido:resourceID>
        <lido:resourceRepresentation>
            <lido:linkResource>link alla risorsa</lido:linkResource>
        </lido:resourceRepresentation>
        <lido:resourceType>
            <lido:term>image</lido:term>
        </lido:resourceType>
        <lido:resourceDescription></lido:resourceDescription>
    </lido:resourceSet>
</lido:rightsResource>

```

```
<lido:rightsType>
  <lido:term>Copyright, Civico Museo del Paesaggio Sonoro, 2007. This
work is licensed under the Creative Commons Attribution 4.0 International License. To
view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/ or send a
letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA</lido:term>
</lido:rightsType>
<lido:rightsDate>
  <lido:earliestDate></lido:earliestDate>
  <lido:latestDate></lido:latestDate>
</lido:rightsDate>
<lido:rightsHolder>
  <lido:legalBodyName>
    <lido:appellationValue>Civico Museo del Paesaggio Sonoro di Riva
presso Chieri</lido:appellationValue>
  </lido:legalBodyName>

<lido:legalBodyWeblink>http://www.museopaesaggiosonoro.org/</lido:legalBodyWeblink>
  </lido:rightsHolder>
  <lido:creditLine></lido:creditLine>
</lido:rightsResource>
</lido:resourceSet>
</lido:resourceWrap>
</lido:administrativeMetadata>
</lido:lido>
</lido:lidoWrap>
```

Bibliografia

- Abadal, E. (2014). *Open Access. L'accesso aperto alla letteratura scientifica*. Milano: Ledizioni [anche reperibile online all'indirizzo: <<http://ledibooks.com/openaccess/>>, ultimo accesso: settembre 2016].
- Aliprandi, S. (2005). *Copyleft & opencontent. L'altra faccia del copyright*. Lodi: PrimaOra [anche reperibile online all'indirizzo: <<http://www.copyleft-italia.it/libri/copyleft-opencontent>>, ultimo accesso: settembre 2016].
- Aliprandi, S. (2010). *Apriti standard! Interoperabilità e formati aperti per l'innovazione tecnologica*. Milano: Ledizioni [anche reperibile online all'indirizzo: <<http://www.aliprandi.org/apriti-standard/>>, ultimo accesso: settembre 2016].
- Bailly, R. (s.d.). *EDM Case study: MIMO and EDM*. Europeana [online, reperibile all'indirizzo <http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/Cases_studies/MIMO%20and%20EDM.pdf>, ultimo accesso: settembre 2016].
- Bailly, R. (2011). *Guidelines for harvesting MIMO's database repository*. MIMO [online, reperibile all'indirizzo <<http://www.mimo-db.eu/mimo/doc/D2.3.pdf>>, ultimo accesso: settembre 2016].
- Bailly, R. e Le Meur, T. (2010). *Metadata Mapping and OAI-PMH Implementation Guidelines*. MIMO [online, reperibile all'indirizzo <[http://www.mimo-international.com/documents/MIMO_WP20-%20D2%20%20Guidelines%20v1.0%20\(2010-11-18\).pdf](http://www.mimo-international.com/documents/MIMO_WP20-%20D2%20%20Guidelines%20v1.0%20(2010-11-18).pdf)>, ultimo accesso: settembre 2016].
- Baines, A. (1960). Organology and European folk instruments. *International Folk Music Journal*, 12, 1960, pp. 10-13.
- Bär, F. P. e Pfefferkorn, F. (2011). *Definition Of Scanning Properties And Recommendations For Photographing Musical Instruments, Version 3*. MIMO [online, reperibile all'indirizzo

- <http://www.mimo-international.com/documents/results/MIMO_Deliverable_1.8.pdf>, ultimo accesso: settembre 2016].
- Berners-Lee, T. (2001). *L'architettura del nuovo Web. Dall'inventore della rete il progetto di una comunicazione democratica, interattiva e intercreativa*. Milano: Feltrinelli. (trad. it. di *Weaving the Web: The Past, Present and Future of the World Wide Web by its Inventor*. London: Orion Business Book, 1999).
- Berners-Lee, T., Bizer, C. e Heath T. (2009). Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems*, vol. 5(3), pp. 1-22 [DOI:10.4018/jswis.2009081901].
- Bianchini, C. (2012). Dagli OPAC ai library linked data: come cambiano le risposte ai bisogni degli utenti. *AIB Studi*, vol. 52, n. 3, pp. 303-323 [DOI:10.2426/aibstudi-8597].
- Bianchini, C. (2014). *RDA e la sfida del web semantico*. In: De Castro, F. (a cura di) *Il punto sul Servizio Bibliotecario Nazionale e le sue realizzazioni nel Friuli Venezia Giulia* (pp. 197-206). Trieste: EUT Edizioni Università di Trieste [URI:<http://hdl.handle.net/10077/10578>].
- Blacking, J. (1973). *How musical is man?* Seattle: University of Washington Press.
- Bohannon, J. (2013). Who's Afraid of Peer Review? *Science*, vol. 342, issue 6154, pp. 60-65 [DOI:10.1126/science.342.6154.60].
- Borgo S., Gaio S., Guarino N., Masolo C. e Oltramari, A. (2010). Un'introduzione all'ontologia DOLCE. *AIDAinformazioni*, 1-2/2010, pp. 107-125 [anche reperibile online all'indirizzo: <<https://www.semanticscholar.org/paper/Un-introduzione-All-ontologia-Dolce-Gaio-Borgo/453b8c53d9be03d7b51c79f85275c765d14a9f02/pdf>>, ultimo accesso: settembre 2016].
- Borst, W. (1997). *Construction of Engineering Ontologies*. Tesi di Dottorato, Institute for Telematica and Information Technology, University of Twente [online, reperibile all'indirizzo <<http://eprints.eemcs.utwente.nl/17377/01/t0000004.pdf>>, ultimo accesso: settembre 2016].

- Bonanni, F. (1723). *Gabinetto armonico pieno d'istromenti sonori*. Roma: Nella Stamperia di Giorgio Placho Intagliatore e Gettatore di Caratteri a Piazza di S. Marco.
- Brazzo, L. e Mazzini, S. (2015). *From the Holocaust Victims Names to the Description of the Persecution of the European Jews in Nazi Years: the Linked Data Approach and a New Domain Ontology*. Paper abstract, 2015 Digital Humanities International Conference, Alliance of Digital Humanities Organizations, 29 giugno-3 luglio 2015, Sydney [online, reperibile all'indirizzo <<http://www.regesta.com/wp-content/uploads/2015/03/Abstract.pdf>>, ultimo accesso: settembre 2016].
- Cassella, M. (2010). *Social peer-review e scienze umane, ovvero "della qualità nella Repubblica della scienza"*. *JLIS.it, Italian Journal of Library, Archives and Information Science*, 1/1, pp. 111-132 [DOI:10.4403/jlis-30].
- Coburn, E., Light, R., McKenna, G., Stein, R. e Vitzthum, A. (2010). *LIDO - Lightweight Information Describing Objects Version 1.0*. ICOM-CIDOC for the Data Harvesting and Interchange Working Group [online, reperibile all'indirizzo <<http://www.lido-schema.org/schema/v1.0/lido-v1.0-specification.pdf>>, ultimo accesso: settembre 2016].
- Crofts, N., Doerr, M., Gill, T., Stead, S. e Stiff, M. (2011). *Definition of the CIDOC Conceptual Reference Model Version 5.0.4*. ICOM/CIDOC CRM Special Interest Group [online, reperibile all'indirizzo <http://www.cidoc-crm.org/docs/cidoc_crm_version_5.0.4.pdf>, ultimo accesso: settembre 2016].
- De Robbio, A. (2007a). *Archivi aperti e comunicazione scientifica*. Napoli: ClíoPress [anche reperibile online all'indirizzo: <<http://www.fedoa.unina.it/1093/>>, ultimo accesso: settembre 2016].
- De Robbio, A. (2007b). *Analisi citazionale e indicatori bibliometrici nel modello Open Access*. *Bollettino AIB*, vol. 47, n. 3, pp. 257-289 [anche reperibile online all'indirizzo: <<http://bollettino.aib.it/article/view/5246>>, ultimo accesso: settembre 2016].

- De Robbio, A. (2013a). Chi sorveglia gli editori predatori. *Il BO. Il giornale dell'Università degli Studi di Padova*. [pubblicato online il 20 febbraio 2013, reperibile all'indirizzo <<http://www.unipd.it/ilbo/content/chi-sorveglia-gli-editori-predatori>>, ultimo accesso: settembre 2016].
- De Robbio, A. (2013b). Luci ed ombre sulla "peer review" delle riviste scientifiche, non solo Open Access. *Il BO. Il giornale dell'Università degli Studi di Padova*. [pubblicato online il 30 ottobre 2013, reperibile all'indirizzo <<http://www.unipd.it/ilbo/content/luci-ed-ombre-sulla-peer-review-delle-riviste-scientifiche-non-solo-open-access>>, ultimo accesso: settembre 2016].
- Dempsey, L. (2000). Scientific, Industrial, and Cultural Heritage: a shared approach. *Ariadne*, vol. 5, issue 22 [pubblicato online il 10 gennaio 2000, reperibile all'indirizzo <<http://www.ariadne.ac.uk/issue22/dempsey/>>, ultimo accesso: settembre 2016].
- Gangemi, A., Guarino, N., Masolo, C., Oltramari, A. e Schneider, L. (2002). *Sweetening ontologies with DOLCE*. In: Benjamins, V. R. (Ed.) *Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web* (pp. 166-181). Berlin: Springer-Verlag [DOI:10.1007/3-540-45810-7_18].
- Gay, J. (Ed.) (2002). *Free Software, Free Society: Selected Essays of Richard M. Stallman*. Boston: Free Software Foundation [anche reperibile online all'indirizzo:<<https://www.gnu.org/philosophy/fsfs/rms-essays.pdf>>, ultimo accesso: settembre 2016].
- Ghirardini, C. (2009). Per la definizione di una scheda SM. *Philomusica on-line*, 8/III, pp. 242-264 [DOI:10.6092/1826-9001.8.783].
- Giuliano, S. G. e Proto, S. (a cura di) (2010). *Norme per la catalogazione degli strumenti musicali: scheda SM*. Palermo: CRICD.

- Gruber, T. (1993). A Translation Approach to Portable Ontologies. *KnowledgeAcquisition*, 5(2), pp. 199-220 [DOI:10.1006/knac.1993.1008].
- Guerrini, M. e Possemato, T. (2013). Linked data: a new alphabet for the semantic web. *JLIS.it, Italian Journal of Library, Archives and Information Science*, 4/1, pp. 67-90 [DOI:10.4403/jlis.it-6305].
- Guerrini, M. e Possemato, T. (2015). *Linked data per biblioteche, archivi e musei*. Milano: Editrice Bibliografica.
- Guizzi, F. (2002). *Gli strumenti della musica popolare in Italia*. Lucca: Lim Editrice.
- Haslhofer, B. e Nussbaumer, P. (2007). *Putting the CIDOC CRM into Practice*. University of Wien [online, reperibile all'indirizzo <<http://eprints.cs.univie.ac.at/404/1/covered.pdf>>, ultimo accesso: settembre 2016].
- Hornbostel, E. M. e Sachs, C. (1914). Systematik der Musikinstrumente. Ein Versuch. *Zeitschrift für Ethnologie*, XLVI, pp. 553-590.
- Hubner, S., Neumann, H., Schuster, G., Stuckenschmidt, H., Visser, U., Voge, T. e Wache, H. (2001). *Ontology-Based Integration of Information - A Survey of Existing Approaches*. In: Gómez Pérez, A., Gruninger M., Stuckenschmidt, H. e Uschold, M. (Eds.) *Proceedings of the IJCAI-01 Workshop on Ontologies and Information Sharing* (pp. 108-117). Seattle: IJCAI [DOI:10.1.1.12.8073].
- Isaac, A. (Ed.) (2013). *Europeana Data Model Primer*. Europeana [online, reperibile all'indirizzo <http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf>, ultimo accesso: settembre 2016].
- Kyrillidou, M. (2012). Research Library Trends: A Historical Picture of Services, Resources, and Spending. *Research Library Issues*, n. 280, pp. 20-28 [anche reperibile online all'indirizzo: <<http://publications.arl.org/rli280/20>>, ultimo accesso: settembre 2016].

- La Vena, V. (1996). *Strumenti giocattolo e strumenti da suono a Terranova da Sibari*. Soveria Mannelli: Rubbettino.
- La Vena, V. (a cura di) (2001). *Strumentario minore a Terranova da Sibari e in altre località della Calabria jonica cosentina*. Rossano Calabro: Associazione musicale "Il Cerchio" (2 cd + booklet).
- Le Meur, T. (2010). *Specifications of the Common Data Model for the Description of Musical Instruments*. MIMO [online, reperibile all'indirizzo <[http://www.mimo-international.com/documents/MIMO_Deliverable_%202.1_v1%20%20\(2010-06-30\).pdf](http://www.mimo-international.com/documents/MIMO_Deliverable_%202.1_v1%20%20(2010-06-30).pdf)>, ultimo accesso: settembre 2016].
- Levy, S. (1984). *Hackers. Heroes of the Computer Revolution*. New York: Anchor Press/Doubleday.
- Leydi, R. (2008). *L'altra musica*. Lucca: LIM Editrice.
- Merriam, A. (2007). *Antropologia della musica*. Palermo: Sellerio. (trad. it. di *The anthropology of music*. Evanston: Northwestern University Press, 1964).
- MIMO, (2009). *Description of Work*. MIMO [online, reperibile all'indirizzo <http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/MIMO/Documents/MIMO%20Description%20of%20Work.pdf>, ultimo accesso: settembre 2016].
- MIMO, (2011). *Revision of the Hornbostel-Sachs Classification of Musical Instruments by the MIMO Consortium*. MIMO [online, reperibile all'indirizzo <http://network.icom.museum/fileadmin/user_upload/minisites/cimcim/documents/H-S_20classification_20final_20version_20_282013_29_20without_20editorial_20markings-2.pdf>, ultimo accesso: settembre 2016].
- Møller, A. e Schwartzbach, M. (2007). *Introduzione a XML*. Milano: Paravia Bruno Mondadori. (trad. it. di *An Introduction to XML and Web Technologies 01*. Harlow: Pearson Addison Wesley, 2006).

- Myers, A. (1989). Cataloguing standards for instrument collections. *CIMCIM Newsletter*, XIV, pp. 14-28 [anche reperibile online all'indirizzo: <<http://www.euchmi.ed.ac.uk/itnXIVc.html>>, ultimo accesso: settembre 2016].
- Myers, A., e Karp C. (2005). *Documentation*. In: R. L. Barclay (Ed.) *The Care of Historic Musical Instruments* (cap. 7, s.p.). CIMCIM [pubblicato online, reperibile all'indirizzo <<http://network.icom.museum/cimcim/resources/the-care-of-historic-musical-instruments-full-text/>>, ultimo accesso: settembre 2016].
- Nattiez, J. J. (1975). *Fondements d'une sémiologie de la musique*. Paris: Union Générale des Éditions.
- Oldman, C. e CRM Labs. (2014). *The CIDOC Conceptual Reference Model (CIDOC-CRM): PRIMER*. CIDOC CRM Special Interest Group [online, reperibile all'indirizzo <http://www.cidoc-crm.org/docs/CRMPrimer_v1.1.pdf>, ultimo accesso: settembre 2016].
- Pievatolo, M. C. (2015). Dopo le riviste: il futuro dell'accesso aperto. *Bollettino telematico di filosofia politica*. [pubblicato online il 30 ottobre 2013, reperibile all'indirizzo <<http://btfp.sp.unipi.it/it/2015/11/oltre-la-rivista-laccesso-aperto-futuro/>>, ultimo accesso: settembre 2016].
- Raymond, E. S. (1999). *La cattedrale e il bazaar*. Apogonline [pubblicato online, reperibile all'indirizzo <<http://www.apogonline.com/openpress/cathedral>>, ultimo accesso settembre 2016]. Trad. it. di *The Cathedral and the Bazaar* [pubblicato online, reperibile all'indirizzo <<http://www.catb.org/esr/writings/cathedral-bazaar/>>, ultimo accesso: settembre 2016].
- Rodger, N. (2011). *Final Report*. MIMO [online, reperibile all'indirizzo <<http://www.mimo-international.com/documents/results/D4.6%20Final%20Report.pdf>>, ultimo accesso: settembre 2016].

- Weisser, S. e Quanten, M. (2011). Rethinking musical instrument classification: Towards a modular approach to the Hornbostel-Sachs system. *Yearbook for Traditional Music*, 43, pp. 122-146 [DOI:10.5921/yeartradmusi.43.0122].
- Zani, M. (2006). Granularità: un percorso di analisi. *DigItalia*, 2, 2006, pp. 60-128 [anche reperibile online all'indirizzo: <<http://digitalia.sbn.it/article/view/302>>, ultimo accesso: settembre 2016].