

**ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA  
CAMPUS DI CESENA  
SCUOLA DI INGEGNERIA E ARCHITETTURA**

**CORSO DI LAUREA MAGISTRALE IN  
INGEGNERIA E SCIENZE INFORMATICHE**

**RICERCA, ELABORAZIONE E MAPPING SU STANDARD  
ONTOLOGICI MODERNI**

Tesi in

**SEMANTIC WEB**

Relatore

**ANTONELLA CARBONARO**

Presentata da

**LUCA BEDESCHI**

**SESSIONE III**

**ANNO ACCADEMICO 2015/2016**



## **ABSTRACT**

Questa Tesi, nell'ambito del Semantic Web applicato al campo dei Beni Culturali (BC), si propone lo scopo di definire in un nuovo formato ontologico l'attuale sistema di registrazione e salvataggio delle informazioni che riguardano un Bene Culturale, ad esempio i dati identificativi, bibliografici, di scavo, ecc., attualmente registrati e salvati senza alcuna tecnologia Semantic Web. Nello specifico, questo progetto di Tesi si svilupperà considerando tra le tante schede che descrivono i termini per la catalogazione di un qualsiasi Bene Culturale, la scheda dei Reperti Archeologici (RA). Per fare questo verrà definito un mapping tra l'attuale sistema di registrazione delle informazioni di un Reperto Archeologico, e di conseguenza un nuovo dominio ontologico, in formato standardizzato RDF e OWL, seguendo le direttive sulle informazioni necessarie alla catalogazione dettate dagli organi del settore.

Il mapping della scheda RA, lo standard di contenuto generale italiano per la catalogazione di Reperti Archeologici, in RDF è basato su un sottoinsieme di elementi di catalogo descritti nella versione 3.00 dello standard, rilasciato nell'anno 2014 dall'Istituto Centrale per il Catalogo e la Documentazione (ICCD). Anche se la scheda RA è considerata un punto di riferimento principale per la catalogazione a livello nazionale, è possibile che anche stakeholder internazionali possano trarre beneficio da questo progetto di mapping, che è un insieme di elementi descritti condivisibili per i reperti archeologici.

Lo scopo di questo progetto è infatti quello di garantire l'interoperabilità semantica tra data providers che vogliono pubblicare i propri dati strutturati sui reperti archeologici in RDF come Linked Open Data. Inoltre, questo documento fornisce una guida per mappare elementi

metadata a termini ontologici definiti come well-known vocabularies, per esempio CIDOC Conceptual Reference Model (CRM), e mediante ontologie che sono state sviluppate per descrivere aspetti specifici dello standard della scheda RA. Quindi, ovunque sia necessario, per evitare lack of information, sono stati adottati o creati termini più adatti di altri vocabolari per questo ambito.

È possibile dividere questa tesi in due parti: la prima parte descrive in maniera generale e tecnica gli standard e i modelli del Semantic Web che permettono di pubblicare dati sui Beni Culturali, fornendo una ricerca e uno studio dello stato dell'arte esistente; la seconda parte, traendo le opportune conclusioni dello studio nella prima parte, definisce e descrive una panoramica, a livello tecnico, del mapping concettuale tra lo standard ICCD relativo alla scheda RA e le ontologie che descrivono questi concetti già pubblicate e standardizzate.

Il risultato è una nuova ontologia, *Central Institute for Cataloguing and Documentation Ontology CICDO*, che a sua volta importa diverse ontologie e vocabolari tra cui: *FOAF* (Vocabolario Friend of a Friend), *CITO* (Citation Typing Ontology), *Erlangen CRM* (CIDOC Conceptual Reference Model), *PROV-O*, *FaBiO* (FRBR-aligned Bibliographic Ontology), *HiCO* (Historical Context Ontology), *FRBR* (Functional Requirements for Bibliographic Records), ed altre importate indirettamente.

Nello specifico, CICDO definisce nuove entità e specializza quelle importate per un totale di: quarantasei classi e quarantasette object properties che descrivono le sezioni e le relazioni dei documenti da compilare di un reperto archeologico, e due datatype properties. Il mapping qui presentato è in forma tabellare, gli elementi sono raggruppati, quando serve, in sotto-tabelle, riproducendo parzialmente i paragrafi, campi e sotto-campi dei documenti ICCD di riferimento.

# CONTENUTI

<b>1. BENI CULTURALI NEL SEMANTIC WEB</b>	<b>1</b>
1.1. CARATTERIZZARE I BENI CULTURALI	1
1.2. PORTALI DI INFORMAZIONE PER BENI CULTURALI	3
1.3. SFIDE DEI DATI SUI BENI CULTURALI	5
1.4. PROMESSE DEL SEMANTIC WEB	6
<b>2. MODELLO COLLABORATIVO PER PUBBLICARE BENI CULTURALI</b>	<b>13</b>
2.1. ACCESSO GLOBALE PER CONTENUTO LINKED DATA LOCALE	13
2.1.1. FEDERATED SEARCH	14
2.1.2. DATA WAREHOUSING	15
2.2. PUBBLICAZIONE COLLABORATIVA DI LINKED DATA	16
2.3. BENEFICI PER GLI UTENTI FINALI	18
2.4. BENEFICI PER PUBLISHERS	19
2.5. NUOVE SFIDE	20
<b>3. REQUISITI PER PUBBLICARE LINKED DATA</b>	<b>23</b>
3.1. MODELLO FIVE-STAR PER LINKED DATA	23
3.1.1. PUBBLICARE DATI STRUTTURATI	24
3.1.2. COLLEGARE DATI INTERNAMENTE ED ESTERNAMENTE	27
3.2. REQUISITI PER INTERFACCE ED API	29
3.2.1. LINKED DATA BROWSING	29
3.2.2. ENDPOINT SPARQL	30
3.2.3. FUNZIONALITÀ DI DOWNLOAD	31
3.2.4. HUMAN INTERFACES	31

<b>4. SCHEMI METADATA</b>	<b>33</b>
<b>4.1. TIPI DI METADATI</b>	<b>33</b>
<b>4.2. SCHEMI WEB</b>	<b>36</b>
4.2.1. DUBLIN CORE	36
4.2.2. VRA CORE CATEGORIES	38
<b>4.3. SCHEMI DI CATALOGAZIONE</b>	<b>40</b>
4.3.1. CATEGORIES FOR THE DESCRIPTION OF WORKS OF ART (CDWA)	40
4.3.2. SPECTRUM	41
4.3.3. FORMATI METADATA IN BIBLIOTECHE	42
4.3.4. FORMATI METADATI NEGLI ARCHIVI	42
<b>4.4. CONCEPTUAL HARMONIZATION SCHEMAS</b>	<b>44</b>
4.4.1. APPROCCI ALL'INTEROPERABILITÀ SEMANTICA	44
4.4.2. EUROPEANA SEMANTIC ELEMENTS (ESE)	45
4.4.3. EUROPEANA DATA MODEL (EDM)	45
4.4.4. CIDOC CONCEPTUAL REFERENCE MODEL (CRM)	46
4.4.5. FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS (FRBR)	49
4.4.6. FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS (FRBR)	51
4.4.7. FUNCTIONAL REQUIREMENTS FOR AUTHORITY DATA (FRAD)	52
4.4.8. FUNCTIONAL REQUIREMENTS FOR SUBJECT AUTHORITY DATA (FRSAD)	52
4.4.9. FRBR OBJECT ORIENTED (FRBROO)	53
<b>4.5. HARVESTING SCHEMAS: LIDO</b>	<b>54</b>
<b>4.6. PROTOCOLLI DI HARVESTING E DI RICERCA</b>	<b>57</b>
4.6.1. ENDPOINT SPARQL PER LINKED DATA	57
<b>5. PROGETTO FASE 1: IL CASO DI STUDIO</b>	<b>59</b>
<b>5.1. OBIETTIVI DEL PROGETTO</b>	<b>60</b>
<b>5.2. GOAL DEL PROGETTO</b>	<b>60</b>
<b>5.3. STUDIO SUL DOMINIO</b>	<b>60</b>
<b>5.4. SCHEDA RA</b>	<b>61</b>
5.4.1. IL PARAGRAFO CD	63

## CONTENUTI

<b>6. PROGETTO FASE 2: MAPPING DEL LIVELLO INVENTARIALE</b>	<b>67</b>
<b>6.1. MODELLI STANDARD PER LE ONTOLOGIE</b>	<b>68</b>
<b>6.2. RDF SCHEMA</b>	<b>68</b>
<b>6.3. SIMPLE KNOWLEDGE ORGANIZATION SYSTEM SKOS</b>	<b>69</b>
6.3.1. WEB ONTOLOGY LANGUAGE OWL	70
6.3.2. ONTOLOGIE LUOGO	72
6.3.3. ONTOLOGIE EVENTO	73
<b>6.4. MAPPING DEL LIVELLO INVENTARIALE</b>	<b>74</b>
6.4.1. IL PARAGRAFO CD	75
6.4.2. IL PARAGRAFO OG	76
6.4.3. IL PARAGRAFO LC	77
<b>6.5. TABELLA DEL MAPPING</b>	<b>78</b>
<b>6.6. IMPLEMENTAZIONE DELL'ONTOLOGIA</b>	<b>79</b>
6.6.1. OBJECT PROPERTY VS. DATA PROPERTY	80
<b>6.7. ARCHEOLOGICAL FIND ONTOLOGY (AFO)</b>	<b>81</b>
6.7.1. OSSERVAZIONI	87
<b>7. PROGETTO FASE 3: MAPPING DELLA SCHEDA RA</b>	<b>89</b>
<b>7.1. INTEGRAZIONE DEI MAPPING</b>	<b>89</b>
7.1.1. FRBR-ALIGNED BIBLIOGRAPHIC ONTOLOGY (FABIO)	90
<b>7.2. STESURA DEL MAPPING COMPLETO</b>	<b>91</b>
<b>7.3. IL PARAGRAFO RE</b>	<b>92</b>
7.3.1. IL CAMPO RCG	93
7.3.2. IL CAMPO DSC	95
<b>7.4. IL PARAGRAFO AU</b>	<b>99</b>
7.4.1. IL CAMPO AUT	100
7.4.2. IL CAMPO ATB	100
<b>7.5. IL PARAGRAFO CO</b>	<b>101</b>
7.5.1. IL CAMPO STC	102
<b>7.6. IL PARAGRAFO GP</b>	<b>102</b>

7.6.1. IL CAMPO GPD	103
<b>7.7. IL PARAGRAFO TU</b>	<b>104</b>
7.7.1. IL CAMPO ACQ	104
<b>8. <u>PROGETTO FASE 4: CICDO</u></b>	<b>109</b>
<b>8.1. TABELLA DEL MAPPING</b>	<b>109</b>
<b>8.2. IMPLEMENTAZIONE SU PROTÉGÉ</b>	<b>111</b>
<b>8.3. CICDO E FABIO</b>	<b>123</b>
<b>9. <u>CONCLUSIONI E SVILUPPI FUTURI</u></b>	<b>125</b>
<b><u>BIBLIOGRAFIA</u></b>	<b>127</b>



## RINGRAZIAMENTI

Desidero ringraziare tutti coloro che mi hanno aiutato nella realizzazione della mia Tesi.

Ringrazio anzitutto la professoressa Antonella Carbonaro, Relatore, e Valerio Capoccia, referente dell'azienda: senza il loro supporto e la loro guida sapiente questa Tesi non esisterebbe.

Ringrazio la mia famiglia: nell'ordine mia sorella Sara e suo marito Nicola per avermi sopportato e supportato nei momenti di insofferenza e di ansietà nel periodo di Tesi e non; i miei genitori, Vanna e Giuseppe, per avermi incoraggiato e aver contribuito notevolmente nel conseguimento di questo titolo, oltre alla lettura e interpretazione di quanto scritto in Tesi; mia nipote, Lucia, nata meno di due anni fa, per i momenti, schedulati e non, di pausa e relax dalla Tesi; un pensiero speciale va alla mia nonna, Rosa, venuta a mancare recentemente ma che aspettava sempre notizie riguardanti università e non, e che ricorderò sempre; e, per ultimo ma non per importanza, a mio nonno, Adamo, che a modo suo ha reso più piacevole questo periodo.

Un ringraziamento particolare va ai colleghi ed agli amici che mi hanno incoraggiato e hanno speso tempo in discussioni con me.

Vorrei infine ringraziare le persone a me più care: la mia famiglia ed i miei amici.

Tutte le persone citate in questa pagina hanno svolto un ruolo fondamentale nella stesura della tesi, ma desidero precisare che ogni errore o imprecisione è imputabile soltanto a me.



# CAPITOLO 1

## BENI CULTURALI NEL SEMANTIC WEB

*Beni Culturali* (BC) si riferiscono al patrimonio di oggetti fisici, ambientali, tradizionali, e conoscitivi di una società che ha ereditato dal passato, mantenuto e sviluppato ulteriormente nel presente, e preservato (conservato) a beneficio delle generazioni future. Questo primo capitolo caratterizza la nozione di BC e identifica le sfide specifiche incontrate durante la pubblicazione di contenuti sui BC, specialmente le collezioni di dati, sul Web. Dopo questo, verranno introdotte le tecnologie sul Semantic Web e Linked Data, identificando un approccio per affrontare il problema.

### 1.1. CARATTERIZZARE I BENI CULTURALI

Un BC può essere diviso in tre sotto-aree:

1. **Patrimonio culturale materiale** è costituito da oggetti culturali concreti, come i manufatti, opere d'arte, edifici, e libri;
2. **Patrimonio culturale immateriale** include fenomeni quali le tradizioni, la lingua, le competenze artigianali, folklore, e la conoscenza;
3. **Patrimonio culturale naturale** è costituito da paesaggi culturalmente significativi, la biodiversità e geo-diversità.

Gli attori chiave nel preservare BC sono sistemi di memoria che includono biblioteche, archivi e musei di differenti tipi specializzati in

## 1.1 CARATTERIZZARE I BENI CULTURALI

aree particolari dei BC, come i musei d'arte, musei archeologici, musei e giardini botanici, musei di storia culturale, collezioni mediche, musei scientifici, musei di storia teatrale, musei geologici e mineralogici, e musei zoologici. Anche i sistemi di comunicazione spesso conservano materiali sui BC, in particolare quelli più recenti. Ci sono anche un sacco di materiali sui BC gestiti da associazioni culturali e singoli individui di vario genere. Il patrimonio BC materiali viene memorizzato con metadati allegati, il patrimonio immateriale viene documentato utilizzando descrizioni testuali, fotografie, interviste e video, e ci sono musei di storia naturale ed altri specializzati nella conservazione di tracce e della conoscenza della storia naturale, geologica, e ambientale.

Il Web è diventato un mezzo sempre più importante per la pubblicazione di contenuti sui BC di diverso tipo. Ad esempio, le biblioteche e gli archivi sono online con i propri assortimenti, i musei mostrano le loro collezioni attraverso collection browser, e la documentazione del patrimonio immateriale è disponibile come registrazione audio e video o applicazioni interattive ipertestuali, addirittura come giochi. Ci sono grandi progetti di portali sui BC nazionali e multinazionali attivi nella raccolta e pubblicazione di contenuti da diverse fonti attraverso servizi centralizzati.

Per un utente finale, questi sistemi forniscono un unico punto di accesso alle collezioni eterogenee di massa e un autorevole fonte di informazioni. A differenza delle mostre fisiche tradizionali, i servizi Web sono aperti per tutto il tempo, è possibile accedere senza la presenza fisica in una mostra, il numero sul Web non è limitato dallo spazio fisico disponibile, e le esposizioni possono essere collegate e accessibili in modo flessibile utilizzando strategie diverse, non solo in quelle utilizzate nelle mostre fisiche. Naturalmente, il Web non può sostituire l'esperienza fisica di visitare un museo o una mostra nella realtà, ma fornisce un'alternativa complementare per accedere a collezioni di dati praticamente in qualsiasi momento e da qualsiasi luogo.

## CAPITOLO 1 - BENI CULTURALI NEL SEMANTIC WEB

Per i ricercatori in scienze umane, la disponibilità di grandi quantità di dati BC in formato digitale processabile ha aperto un nuovo paradigma di ricerca denominato Digital Humanities.

### 1.2. PORTALI DI INFORMAZIONE PER BENI CULTURALI

Ci sono diversi tipi di pubblicazioni BC sul Web. In primo luogo, vi è una grande varietà di sistemi well-curated che sono stati hand-crafted per uno scopo specifico con un tema chiuso focalizzato, dataset, e interfacce. Tali sistemi sono spesso implementati utilizzando strumenti quali Adobe Flash con interfacce simili a giochi. Ad esempio, la Spedizione di Lewis e Clark <sup>1</sup>(180e-1806) è documentata sul Web in modo molto dettagliato da diverse applicazioni e sistemi. Tali sistemi possono essere disponibili anche su CD/DVD come applicazioni stand-alone.

D'altra estremità, ci sono servizi di ricerca di collezioni e browser che forniscono accesso a grandi collezioni di database aperti il cui contenuto non è tematicamente focalizzato, i cui percorsi di accesso e interfacce possono risultare mancanti. In cambio, è possibile accedere a database di grandi collezioni provenienti da diverse istituzioni. Ad esempio, è possibile accedere ad una varietà di collezioni BC australiane utilizzando il Collection Australia Network System<sup>2</sup>. Possono essere trovati federated portals simili per la ricerca e la navigazione di collezioni in molti paesi a livello internazionale. Un'applicazione di punta è Europea<sup>3</sup>, basata su di milioni di collezioni di oggetti provenienti da sistemi di memoria in tutta Europa. Ad esempio, in Figura 1 l'utente ha digitato la parola "reperto archeologico" nel campo di ricerca di Europea e il sistema ha trovato varie entità in collezioni associate.

---

<sup>1</sup> <http://lewis-clark.org/>

<sup>2</sup> <http://www.collectionaustralia.net>

<sup>3</sup> <http://www.europeana.eu/>

## 1.2 PORTALI DI INFORMAZIONE PER BENI CULTURALI

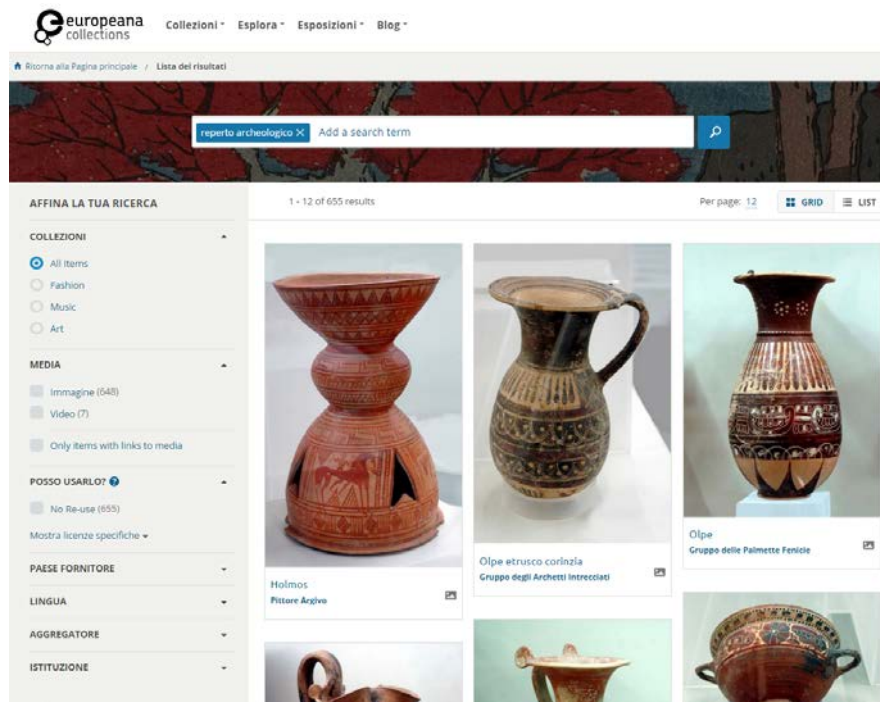


Figura 1: Ricerca nel portale Europeana reperti archeologici provenienti da diverse collezioni Europeana

La ricerca può essere raffinata ulteriormente selezionando i filtri supplementari sulle facets a sinistra, come “collezioni”, “media”, paese fornitore”, “lingua”, ecc. Un altro esempio di portale che raccoglie dati di biblioteche, è WorldCat<sup>4</sup> che contiene metadati di circa 1,5 miliardi di libri, DVD, CD, e articoli in biblioteche associate. Un altro portale internazionale è la World Digital Library, gestito da UNESCO e gli United States Library of Congress, che mette a disposizione, gratuitamente, importanti materie prime multilingue, quali manoscritti, mappe, libri rari, spartiti musicali, registrazioni, film, stampe, fotografie e disegni architettonici.

Quest'ultimo genere di portali sono considerati portali BC-based su grandi dataset di collezioni eterogenee, in cui organizzare i contenuti “a

---

<sup>4</sup> <https://www.worldcat.org/>

## CAPITOLO 1 - BENI CULTURALI NEL SEMANTIC WEB

mano” in un'applicazione specifica, con visualizzazioni e interfacce specifiche dell'applicazione, di solito non è fattibile. Tali portali condivisi facilitano lo scambio di conoscenze per i ricercatori nel campo dei BC, bibliotecari, e archivisti. Per contribuire all'organizzazione di memoria, tali sistemi forniscono l'occasione per entrare in contatto con un pubblico più ampio sul Web con nuove modalità di interazione, e di collaborare con altre organizzazioni. Da un punto di vista sociale, la pubblicazione di BC sul Web stimola il turismo culturale, l'economia creativa, e migliora le relazioni e l'unità tra le parti e le nazioni coinvolte in tali iniziative.

### 1.3. SFIDE DEI DATI SUI BENI CULTURALI

La raccolta di dati sui BC ha molte caratteristiche specifiche, come quelle che seguono:

- **Multi-formato.** I contenuti sono presentati in varie forme, come ad esempio documenti di testo, immagini, tracce audio, video, oggetti da collezione e oggetti di apprendimento;
- **Multi-topical.** I contenuti riguardano vari argomenti, come l'arte, la storia, i manufatti e le tradizioni;
- **Multilingue.** Il contenuto è disponibile in diverse lingue;
- **Multi-culturale.** Il contenuto è legato e interpretato in termini di differenti culture, come religioni o tradizioni nazionali in Occidente e Oriente;
- **Multi-target.** I contenuti sono spesso mirati sia a laici che a esperti, giovani e anziani.

Di conseguenza, un aspetto problematico fondamentale nel trattare con i dati BC è quello di rendere il contenuto reciprocamente *interoperabile*, in modo che possa essere cercato, legato, e presentato in

## 1.4 PROMESSE DEL SEMANTIC WEB

modo armonizzato attraverso i confini dei dataset e dei data silos. Il problema si verifica a livello sintattico, ad esempio, quando si armonizzano diversi set di caratteri, formati di dati, notazioni, e record di collezione utilizzati in diverse collezioni. Ancora più importante, esiste un problema di *interoperabilità semantica*: possono essere interpretati in modo diverso diversi formati di metadati, i dati sono codificati a diversi livelli di precisione, sono diversi i vocabolari e dizionari geografici utilizzati per la descrizione del contenuto, e così via. I Semantic Web standards<sup>5</sup> e le best practices, in particolare quelle sostenute dal World Wide Web Consortium (W3C), forniscono una base condivisa su cui i sistemi Web interoperabili possono essere costruiti in modo ben definito. Le nuove tecnologie non sono ovviamente una panacea per tutti i mali, ma piuttosto uno strumento impostato con cui i problemi difficili possono essere affrontati senza dubbio in modo più efficace rispetto a prima.

### 1.4. PROMESSE DEL SEMANTIC WEB

Le tecnologie del Semantic Web<sup>6</sup> sono un nuovo approccio promettente per affrontare i problemi riguardante la pubblicazione di contenuti BC sul Web. Il termine "semantico" qui si riferisce alla semantica, una disciplina che studia le relazioni tra significati, come parole, frasi, segni e simboli, e ciò che rappresentano, cioè denotano. In semantica applicata alla Computer Science ci si riferisce al significato formale e alla sua interpretazione (dichiarativa o procedurale), che è stata data a strutture sintattiche, come linguaggi di programmazione o strutture di dati simbolici.

Il Semantic Web può essere visto come un nuovo strato di metadati costruito all'interno del Web. Secondo la definizione tradizionale, i

---

<sup>5</sup> <https://www.w3.org/>

<sup>6</sup> <https://www.w3.org/standards/semanticweb/>



## CAPITOLO 1 - BENI CULTURALI NEL SEMANTIC WEB

metadati sono dati sui dati. Ad esempio, un record di metadati di un libro (dato) può specificare il suo titolo, autore, soggetto, e anno di pubblicazione. Tuttavia, il termine "metadato" è usato più ampiamente nel contesto del Semantic Web come sinonimo di macchina per dati processabili o interpretabili. L'idea chiave è che le strutture di metadati sintattiche rendano i contenuti Web "comprensibili" per le macchine, in base alle specifiche semantiche condivise fondate sulla logica formale. In questo modo è possibile creare servizi Web più interoperabili e intelligenti.

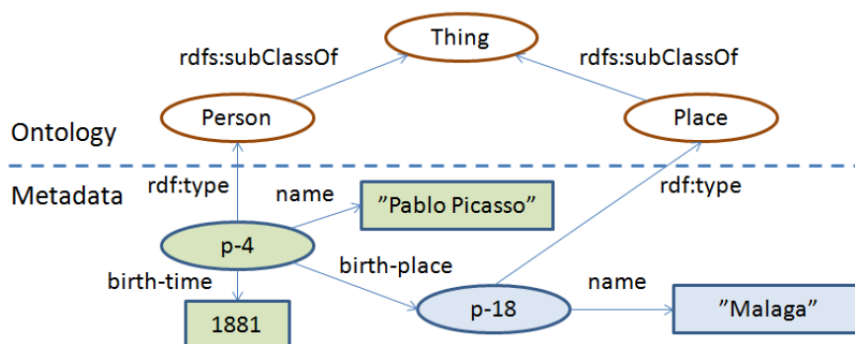


Figura 2: Il modello RDF è un grafo etichettato diretto.

La metodologia per rappresentare metadati e concetti ontologici sul Web è basata su un semplice modello di dati: un grafo etichettato diretto, cioè una rete semantica. Ad esempio, la [Figura 2](#) illustra un grafo RDF dicendo a un livello di metadati che l'identità p-4 è un individuo della classe Persona (indicato con l'arco `rdf:type`) con il nome di "Pablo Picasso" nato nel 1881 in un'istanza p-18 della classe Place il cui nome è "Malaga". Nel grafo RDF, le classi come luoghi e persone sono rappresentate come sottoclassi (arco `rdfs:subClassOf`) della classe Thing a livello ontologico, mentre gli individui delle classi sono considerati metadati. Sia metadati che le ontologie vengono rappresentate in modo uniforme nello stesso grafo. Nella figura, le identità che possono avere proprietà, cioè, possono avere archi uscenti, vengono raffigurate come

## 1.4 PROMESSE DEL SEMANTIC WEB

ovali, mentre valori letterali atomici letterali senza ulteriori proprietà (nell'esempio stringhe e numeri) come contenitori rettangolari.

La figura illustra che in realtà ci sono diversi livelli di descrizioni necessarie sul Semantic Web:

1. **Mondo reale.** Al livello più basso, c'è il mondo reale, cioè il dominio del discorso, come ad esempio persone, oggetti e luoghi;
2. **Livello dei dati.** Poi c'è il livello dei dati, dal momento che gli elementi del mondo reale devono essere rappresentati come dati di qualche tipo in un computer. Ad esempio, le immagini e i documenti sono dati, come anche un URI che fa riferimento ad una persona;
3. **Livello metadati.** Dopo i dati, abbiamo i metadati relativi ai dati, ad esempio, record in un database di collezioni riguardanti immagini, persone o artefatti;
4. **Livello ontologico.** Successivamente, il livello ontologico definisce le classi generiche e le proprietà utilizzate per descrivere un dominio, vale a dire, i vocabolari in cui sono rappresentati termini metadati. Fanno parte di questo livello sia gli schemi di metadati utilizzati nella catalogazione che vocabolari di subject headings (cioè parole standardizzate assegnate ad un concetto);
5. **Livello di meta-ontologia.** Infine, ci sono i principi di modellazione tra domini generali di ontologie che sono indipendenti per ognuna. Ad esempio, le nozioni di relazione subclass-of e di classe sono generiche e non limitate ad un particolare dominio. Tali principi generici sono specificati dagli standard del Semantic Web, come RDF(S) e OWL, e facilitano l'interoperabilità tra domini di contenuti.

Su scala Web globale, il Semantic Web forma un Giant Global Graph (GGG) delle risorse di dati collegati. Il GGG può essere utilizzato e sfogliato in modo analogo al WWW, ma mentre i collegamenti WWW

## CAPITOLO 1 - BENI CULTURALI NEL SEMANTIC WEB

associano pagine Web per uso umano, i link GGG associano concetti base e risorse dati. Ad esempio, il GGG può dire che le anatre sono uccelli, e che Donald è un'istanza di un'anatra (e quindi un uccello), mentre le pagine WWW relative possono costituire un libro di fumetti su Donald Duck.

Un'idea chiave di linked data è che le varie parti del GGG possano provenire da diverse data sources. Ad esempio, nella [Figura 2](#) i metadati relativi alle persone, come Pablo Picasso, possono provenire da un authority database, le informazioni sui luoghi, come Malaga, possono essere forniti da una organizzazione per il rilevamento topografico, e l'ontologia classe può essere basata su un keyword thesaurus esistente in uso in una libreria. La differenza tra data sources viene illustrata nella figura con diversi colori/densità.

Sulla base di rappresentazioni RDF-based armonizzate di dati, possono essere costruite con meno sforzo più applicazioni Web "intelligenti". Dal punto di vista delle applicazioni tecniche, le tecnologie Semantic Web hanno molte caratteristiche promettenti:

- **Descrizioni del contenuto più precise.** La tecnologia si basa su un univoco e globale Universal Resource Identifiers (URI), che consente di fare riferimento a significati più accurati rispetto all'utilizzo di espressioni letterali. Ad esempio, i nomi di persona e luogo possono essere disambiguati: ci sono un sacco di "John Smith" in giro per il mondo, "Parigi" si trova in Francia, in Texas, e in molti altri luoghi, ed i nomi possono avere diverse traslitterazioni in diversi sistemi linguistici. Nelle biblioteche, la nozione di Shakespeare interpreta "Amleto" può riferirsi alla storia astratta, alla sua manifestazione come un testo o un video della rappresentazione, a diverse traduzioni, a varianti della storia, a diverse edizioni e, infine, a libri o DVD sugli scaffali di una biblioteca. La modellazione di tali distinzioni semantiche può essere fatta utilizzando standard BC "ontology-based";

## 1.4 PROMESSE DEL SEMANTIC WEB

- **Interoperabilità.** Le tecnologie Semantic Web forniscono un nuovo approccio per la creazione di linked data interoperabili;
- **Un modello di dato semplice per l'aggregazione.** Due grafi (interoperabili) RDF possono essere uniti insieme, tecnicamente in modo banale, semplicemente eseguendone l'unione (cioè i corrispondenti set di triple);
- **L'aggregazione di dati dai linked data.** Grazie alla combinazione di data sources in modo interoperabile, i dati da una sorgente possono essere arricchiti con linked data aggiuntivi da un'altra fonte. Una notevole iniziativa internazionale verso questo obiettivo è Linked Data<sup>7</sup>, in cui dataset aperti come Wikipedia/DBpedia<sup>8</sup> relativi a conoscenza comune, GeoNames<sup>9</sup> per milioni di luogo, o il progetto Gutenberg<sup>10</sup> per oltre 40.000 ebook gratis vengono descritti in termini di standard Semantic Web e interconnessi tra loro;
- **Servizi del Semantic Web.** Linked data semantici vengono pubblicati non solo come insiemi di dati passivi, ma come servizi operativi che possono essere utilizzati come legacy e altre applicazioni BC tramite Application Programming Interface (API) aperte e generiche. Utilizzando questi servizi condivisi, i programmatori possono riutilizzare il lavoro fatto da altri. Questa idea può essere messa in parallelo con le Mappe di Google e Yahoo! che forniscono un servizio su base globale per le applicazioni tramite le API di facile utilizzo per lo sviluppo di mash-up.

Publicare BC sul Web non è solo una sfida tecnica ma esistono problemi di affidabilità sui contenuti, diritti d'autore, e concessione di licenze. Molti contenuti sui BC sono protetti da copyright, e ci sono

---

<sup>7</sup> <http://linkeddata.org/>

<sup>8</sup> <http://wiki.dbpedia.org/>

<sup>9</sup> <http://www.geonames.org/>

<sup>10</sup> <http://www.gutenberg.org/>

## CAPITOLO 1 - BENI CULTURALI NEL SEMANTIC WEB

anche altri motivi per cui le organizzazioni non possono pubblicare i propri dati apertamente, ad esempio, questioni di privacy personale. Tuttavia, sulla base delle idee di Linked Open Data, il mondo WWW sta chiaramente prendendo passi verso la pubblicazione di open data. L'idea è che il contenuto BC debba essere condiviso al massimo. Di solito vengono sono un prodotto che parte da finanziamenti pubblici e in un certo senso già pagati dal pubblico. La disponibilità di questi open data favorisce anche l'interoperabilità e crea una base su cui le applicazioni commerciali possono essere costruite più facilmente.



## **CAPITOLO 2**

### **MODELLO COLLABORATIVO PER PUBBLICARE BENI CULTURALI**

Questa sezione presenta un “modello prototipico” e un’architettura per la pubblicazione di collezioni BC e altri materiali in modo collaborativo sul Semantic Web come Linked Data. Vengono inoltre analizzati sia i benefici per i pubblicatori e gli utenti finali fruitori di questi dati che le prospettive e le sfide discusse.

#### **2.1. ACCESSO GLOBALE PER CONTENUTO LINKED DATA LOCALE**

Di solito, i contenuti di memoria delle organizzazioni vengono mantenute in silos di dati diversi: i database vengono distribuiti in luoghi differenti e utilizzano diversi sistemi e schemi. Ogni organizzazione pubblica in genere pubblica il proprio contenuto utilizzando homepage custom e collection browser. Questi confini delle organizzazioni creano gravi ostacoli all’information retrieval per l’utente finale, che deve sapere quale contenuto può essere trovato in quale collezione, deve imparare ad usare diversi sistemi di ricerca, e deve più volte interrogare e raccogliere dati con interfacce web differenti. Piuttosto, l’utente finale ha bisogno di un servizio unico globale per l’accesso ai dati situati in diversi database, ad esempio dei musei, invece di avere un insieme di sistemi specifici relativi alle organizzazioni locali. Per facilitare questo, i contenuti locali devono essere aggregati e combinati in qualche modo, e successivamente

## 2.1 ACCESSO GLOBALE PER CONTENUTO LINKED DATA LOCALE

utilizzati tramite un'interfaccia utente armonizzata, come un portale Web.

Con il termine *data integration* si fa riferimento alla combinazione di dati provenienti da diverse sorgenti e fornirgli una vista unificata. Data integration può avvenire in diverse fasi dell'information retrieval. Esistono due tipi di strategie principali:

1. **Integrare i dati in modo dinamico durante l'elaborazione di query;**
2. **Integrare i dati in anticipo in una fase separata di pre-processing.**

### 2.1.1. FEDERATED SEARCH

Prendendo in considerazione il caso precedente, una query di ricerca può essere inviata al database locale distribuito, e il risultato può essere combinato con una lista di risultati globali. Questo approccio viene chiamato *federated search*, anche conosciuto come *multi-search* o *meta-search*. La federated search è costituita dai seguenti passaggi:

1. **Trasformare** la query per ogni database service e trasmetterlo;
2. **Raccogliere** i risultati dei servizi ed unirli in un unico result set;
3. **Presentare** il result set in un formato conciso con il minimo di duplicazione;
4. **Fornire** all'utente finale mezzi per il controllo del result set unito, ad esempio, classificandolo in diversi modi.

Un notevole vantaggio della federated search è che richiede solamente loose coupling dei sistemi di collezione dei database coinvolti: le organizzazioni partecipanti devono solo essere d'accordo sul protocollo di integrazione, ed i database system possono continuare indipendentemente come prima. Grazie alla sua semplicità, federated



## CAPITOLO 2 - MODELLO COLLABORATIVO PER PUBBLICARE BENI CULTURALI

search è stata applicata ampiamente portali BC (ed altri), e ci sono molti sistemi di raccolta culturali nazionali sul Web basati su di esso, come ad esempio i portali come Collection Australia Network<sup>1</sup> e Artefacts Canada<sup>2</sup>. L'approccio è particolarmente utile quando la ricerca di oggetti di tipo simile è rappresentata tramite formato di metadati simile, come i libri nelle biblioteche.

Federated search può essere effettuata utilizzando due strategie. Nella strategia *Local as View* (LAV) i database locali forniscono il sistema di ricerca con le proprie viste locali, su cui la federated query deve essere trasformata. Nella strategia *Global as View* (GAV), viene mediata una visione globale condivisa ai servizi locali a cui deve rispondere. Vengono detti *wrapper* i moduli software che permettono la trasformazione delle query necessarie.

Una limitazione fondamentale della federated search è che processando la query indipendentemente in ogni database locale, le dipendenze globali, vale a dire associazioni tra oggetti attraverso diverse collezioni, sono difficili da trovare. Dal momento che trovare collegamenti semantici tra collection item è uno dei maggiori obiettivi dei linked data, un portale BC semantico per linked data non può essere basato sul paradigma della federated search senza sacrificare gran parte del potenziale del Semantic Web.

### 2.1.2. DATA WAREHOUSING

L'approccio alternativo alla federated search è quello di consolidare prima le collezioni locali in un repository globale, un data warehouse, e quindi ricercare sul database globale. Possono essere utilizzati modelli concettuali reciprocamente condivisi ed ontologie per l'arricchimento dei contenuti e l'interoperabilità delle collezioni. Per mostrare le

---

<sup>1</sup> <https://aiccm.org.au/collections-australia-network-0>

<sup>2</sup> <http://www.rcip-chin.gc.ca/artefacts/index-eng.jsp>

## 2.2 PUBBLICAZIONE COLLABORATIVA DI LINKED DATA

associazioni all'utente finale, gli elementi di raccolta possono essere rappresentati come pagine Web collegate tra di loro attraverso associazioni semantiche. Una sfida importante in questo approccio è che risulta necessario un processo di creazione di contenuti separato per consolidare il repository globale su database locali, e il suo mantenimento al evolversi dei database partecipanti. È necessaria una collaborazione maggiore e un coordinamento tra i content providers di tale portale.

### 2.2. PUBBLICAZIONE COLLABORATIVA DI LINKED DATA

I Linked Data sui BC sono eterogenei e allo stesso tempo interconnessi. Per esempio, il contenuto potrebbe riguardare la biografia narrativa di una persona, il suo lavoro creato, posti d'interesse in cui è vissuto, articoli di Wikipedia o romanzi su di lui o scritti di lei, i rapporti sociali con i suoi parenti o altre persone, e gli eventi storici in cui la persona è stata coinvolta. Tali contenuti multi-prospective rischiano di essere creati in modo indipendente l'uno dall'altro, utilizzando diversi schemi di metadati, vocabolari, accordi di catalogazione e lingue. Il collegamento di dati eterogenei in un ambiente così distribuito è impegnativo per utenti finali di dati culturali, nonché alle organizzazioni e alle comunità che producono questo tipo di contenuti.

Fortunatamente, le idee del Semantic Web e dei linked data possono essere applicate per affrontare problemi di interoperabilità dei dati e allo stesso tempo per la creazione di contenuti distribuiti, come illustrato nella [Figura 3](#). Nella rappresentazione, il sistema di pubblicazione è illustrato da un cerchio. Nel mezzo si trova una rete semantica condivisa, un'infrastruttura ontologica, e consiste in ontologie condivise allineate e modelli di metadati utilizzati per descrivere contenuti del sistema. Se i providers di contenuti esterni al cerchio forniscono il sistema con metadati relativi ai BC utilizzando gli stessi

## CAPITOLO 2 - MODELLO COLLABORATIVO PER PUBBLICARE BENI CULTURALI

concetti, i dati verranno automaticamente collegati gli uni agli altri e formeranno un GGG.



*Figura 3: Un modello per pubblicare linked BC, basato su un'infrastruttura ontologica condivisa al centro*

Per esempio, se i metadati di un dipinto creato da Picasso provenisse da un museo d'arte, potrebbe essere arricchito (collegato) con le biografie di Wikipedia ed altre sorgenti, alcune sue fotografie, informazioni sulle sue mogli, i libri in una libreria che descrive le sue opere d'arte, mostre aperte delle sue opere nei musei, e così via. In questo modo, i metadati del dipinto vengono semanticamente arricchiti, se il collegamento viene stabilito in modo corretto. Allo stesso tempo, anche i contenuti di organizzazioni presenti nel portale che riguardano Picasso vengono arricchite dagli stessi metadati della nuova opera inserita nel sistema. Si tratta chiaramente di una situazione vantaggiosa per tutti i partecipanti del sistema, in cui la collaborazione ripaga.

## 2.3 BENEFICI PER GLI UTENTI FINALI

I potenziali benefici sono particolarmente elevati in situazioni in cui il contenuto BC distribuito è relativo a un argomento in luoghi diversi, e ottenere una visione olistica è più impegnativo. Ad esempio, in Europa esistono un numero ingente di Paesi in cui le linee di confine sono cambiate nel corso della storia, e il contenuto BC è stato distribuito a differenti collezioni nazionali. Allo stesso modo, gran parte del contenuto BC delle ex colonie, come i tesori dell'antico Egitto, sono stati trasportati in musei occidentali, e non sono più disponibili nelle sedi originali. Utilizzando un sistema come quella in [Figura 3](#), collezioni BC potrebbero almeno essere virtualmente unite sul Semantic Web.

### 2.3. BENEFICI PER GLI UTENTI FINALI

Un portale applicativo semantico, seguendo il modello illustrato precedentemente, è utile dal punto di vista degli utenti finali in diversi modi:

- **Vista globale, eterogenea e distribuita dei contenuti.** Il contenuto di diversi content providers è accessibile attraverso un servizio unico e una repository omogenea. Ci si dovrà abituare solo ad un'unica interfaccia utente;
- **Aggregazione automatica dei contenuti.** Soddisfare la necessità di informazione di un utente richiede spesso aggregazione di contenuti da diversi information providers. Ad esempio quando, nella ricerca di dati su un artista, le informazioni rilevanti possono essere fornite da collezioni museali, biblioteche, archivi, authority records, ontologie ed altre fonti;
- **Ricerca semantica.** Nei portali tradizionali, di solito la ricerca è basata su *free text search* (e.g., Google), query a database, e/o gerarchie di classificazione. Un contenuto semantico rende possibile fornire all'utente finale strutture più "intelligenti" basate su concetti

## CAPITOLO 2 - MODELLO COLLABORATIVO PER PUBBLICARE BENI CULTURALI

e strutture ontologiche, come la ricerca semantica, l'auto-completamento e la faceted search;

- **Navigazione semantica e raccomandazioni.** Un contenuto semantico facilita anche la navigazione e la raccomandazione semantica per ricevere ulteriori informazioni. In questo caso, possono essere esposte all'utente finale associazioni semantiche tra oggetti di ricerca come collegamenti di raccomandazione, possibilmente con spiegazioni esplicite sull'associazione;
- **Altri servizi intelligenti.** Possono essere creati anche altri servizi sulla base dei "machine interpretable content", come la scoperta di conoscenza e associazioni, personalizzazione, e le visualizzazioni semantiche basate su, ad esempio, mappe storiche e contemporanee e/o timelines.

### 2.4. BENEFICI PER PUBLISHERS

Questo tipo di tecnologie sono molto interessanti ed utili anche dal punto di vista di "publisher" di contenuto:

- **Creazione di contenuti distribuiti.** Un portale di contenuti viene generalmente creato in modo centralizzato utilizzando un sistema di gestione dei contenuti (CMS). Questo approccio è costoso e non possibile se viene creato in modo distribuito da publishers indipendenti, ad esempio, da diversi musei e altre memory organizations. Le tecnologie semantiche possono essere utilizzate per la raccolta e l'aggregazione di contenuti distribuiti omogenei in portali globali di contenuti;
- **Manutenzione dei link automatica.** Il problema di mantenere i link up-to-date è il costo dal punto di vista di manutenzione. In un portale semantico, i link possono essere creati e mantenuti automaticamente in base ai metadati e alle ontologie;

## 2.5 NUOVE SFIDE

- **Canale comune di pubblicazione dei contenuti.** Nel dominio culturale di solito i publishers condividono l'obiettivo comune di promuovere la conoscenza culturale al pubblico e tra professionisti. Un portale semantico è in grado di fornire alle organizzazioni partecipanti un canale di pubblicazione condiviso e allo stesso tempo economico;
- **Arricchimento di altri contenuti semantici.** L'interconnessione del contenuto tra le organizzazioni che collaborano arricchisce i contenuti di tutti "for free";
- **Riutilizzo di contenuti aggregati.** I contenuti aggregati in un portale semantico possono essere riutilizzati in diverse applicazioni e sistemi cross-portal.

## 2.5. NUOVE SFIDE

La produzione e l'utilizzo di dati semantici pone nuove sfide. Ottenere interoperabilità richiede, in pratica, l'uso più disciplinato di standard, modelli di metadati armonizzati, vocabolari condivisi, e best-practices condivise. La sfida principale spesso più organizzativa che tecnica: la modifica, ad esempio, di pratiche di catalogazione non è semplice, e se vengono apportate modifiche, bisogna decidere cosa farne dei metadati legacy già catalogati. Una difficoltà pratica è che i sistemi di gestione di contenuti in uso non supportano la creazione dei dati Semantic Web. Inoltre è necessario, quando si aggregano contenuti tra confini organizzativi, più collaborazione possibile e armonizzazione del dato.

Anche lo sviluppo di applicazioni più "intelligenti" stabilisce nuove sfide: tali sistemi sono in genere più complessi, richiedono competenze specifiche da parte dei programmatori, nuovi strumenti, e sono in genere computazionalmente complessi e non necessariamente scalano facilmente. L'arricchimento dei dati tramite linked data è promettente, ma in pratica i dataset disponibili hanno problemi di qualità:

## **CAPITOLO 2 - MODELLO COLLABORATIVO PER PUBBLICARE BENI CULTURALI**

ad esempio, DBpedia viene prodotto automaticamente senza supervisione umana. Il problema è che gli identificatori URI utilizzati per i concetti (ad esempio, persone e luoghi) in dataset differenti sono tipicamente diversi, e le mappature dei dati non sono complete o contengono errori. I vocabolari utilizzati possono contenere loops, etichette mancanti, violare vincoli di norme semantiche, e così via.

Tuttavia, è chiaro che queste sfide devono essere affrontate in un modo o nell'altro quando si integrano dati di collezioni a livello semantico. Le tecnologie del Semantic Web forniscono un approccio standard ed un set di strumenti che sono stati applicati con successo al compito, quindi da sperimentare prima di reinventare soluzioni simili.





## CAPITOLO 3

# REQUISITI PER PUBBLICARE LINKED DATA

Questo capitolo per prima cosa indica i requisiti tecnici generali per la pubblicazione di contenuti BC come Linked (Open) Data, sulla base di cinque livelli di qualità. Successivamente, vengono presentate API standard per repository Linked Data. Infine, i problemi di qualità dei linked data discussi.

### 3.1. MODELLO FIVE-STAR PER LINKED DATA

I dati possono essere pubblicati sul Web a livelli incrementali che riguardano openness e linkage, caratterizzato dal sistema five-star presentato in

★	<b>Struttura dati.</b> I dati sono disponibili come dati strutturati.
★★	<b>Licensing.</b> I dati sono disponibili sul Web (in qualsiasi formato) sotto una licenza open.
★★★	<b>Formato.</b> Vengono utilizzato formati aperti non proprietari, ad esempio CSV al posto di Excel.
★★★★	<b>Identifiers.</b> Vengono utilizzati URI per identificare le cose.
★★★★★	<b>Data Linking.</b> I dati sono collegati internamente ed esternamente ad altri per fornire un contesto.

Tabella 1: Sistema Five-Star per valutare Linked Data

## 3.1 MODELLO FIVE-STAR PER LINKED DATA

L'obiettivo generale è quello di cercare di guadagnare il maggiore numero di stelle possibile, una per ogni punto elencato in Tabella 1. Di seguito, vengono presentate le linee guida per pubblicare linked data, partendo dal requisito di dato strutturato.

### 3.1.1. PUBBLICARE DATI STRUTTURATI

La pubblicazione di dati in formati strutturati significa che i dati non possano essere interpretati solo da esseri umani, ma anche interpretati ed utilizzati da elaboratori (ad esempio, computer). La pubblicazione di un foglio di Excel in formato CSV è molto più riutilizzabile che la pubblicazione di una scansione di un'immagine di un tavolo, anche se l'immagine potrebbe essere utile per i lettori umani.

La prima stella è guadagnata utilizzando tutti i formati strutturati, ma per Linked Data la scelta ottimale è quella di utilizzare il Resource Description Framework RDF<sup>1</sup>. RDF ha molte sintassi, e la scelta dipende dal caso d'uso in mente. Le sintassi testuali RDF vengono chiamate serializzazioni perché il testo rappresenta i grafi multi-dimensionali come una serie di caratteri in un file.

1. **RDF/XML**<sup>2</sup> è il modo standard originale di rappresentazione di grafi RDF. È basato su XML e può essere analizzato e convalidato facilmente dalla maggior parte degli strumenti di RDF. Ad esempio, i metadati sulla risorsa Pablo Picasso *p-4* in Figura 2 può essere rappresentato in forma RDF/XML come:

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf=http://www.w3.org/1999/02/22-rdf-syntax-ns#
  xmlns:ex="http://example.org">
  <rdf:Description rdf:about="http://example.org/p-4">
    <rdf:type rdf:resource="http://example.org/Person"/>
    <ex:name>Pablo Picasso</ex:name>
```

---

<sup>1</sup> <https://www.w3.org/RDF/>

<sup>2</sup> <http://www.w3.org/TR/REC-rdf-syntax/>

## CAPITOLO 3 - REQUISITI PER PUBBLICARE LINKED DATA

```
<ex:birth-place rdf:resource="http://example.org/p-18"/>
<ex:birth-time>1881</ex:birth-time>
  </rdf:Description>
</rdf:RDF>
```

La sintassi RDF/XML è prolissa da scrivere e di difficile lettura per gli utenti. Pertanto, sono state sviluppate sintassi più semplici come N-triples, N3 e Turtle;

2. **N-Triple** è un modo semplice di serializzazione RDF come un insieme di triple. Si tratta di una versione semplificata di Notation 3 (N3)<sup>3</sup>, una notazione human-friendly ed estensione per grafi RDF. Il formato N-Triple elenca una tripla <oggetto, predicato, oggetto> in una riga. È utile quando si creano dataset che possono essere caricati in modo efficiente in un triplestore senza bisogno di caricare tutti i dati in memoria. Ad esempio, i metadati sulla *p-4* in [Figura 2](#) in formato N-Triple diventa:

```
@prefix ex: <http://example.org/>.
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
ex:p-4 rdf:type ex:Person . ex:p-4 ex:name "Pablo Picasso".
ex:p-4 ex:birth-place ex:p-18 . ex:p-4 ex:birth-time 1881.
```

3. **Terse RDF Triple Turtle**<sup>4</sup> è probabilmente il formato più semplice da utilizzare dal punto di vista umano, ed è diventato molto popolare. È compatibile con N-Triple, ma è conciso da scrivere. È inoltre compatibile con query pattern syntax utilizzato in SPARQL<sup>5</sup>, il linguaggio di query per il Semantic Web.

Per esempio, i metadati riguardanti la risorsa *p-4* della [Figura 2](#) in Turtle possono essere rappresentati senza ripetizioni del soggetto:

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
@prefix ex: <http://example.org/>.
```

---

<sup>3</sup> <https://www.w3.org/TeamSubmission/n3/>

<sup>4</sup> <https://www.w3.org/TR/turtle/>

<sup>5</sup> [https://www.w3.org/standards/techs/sparql#w3c\\_all](https://www.w3.org/standards/techs/sparql#w3c_all)

### 3.1 MODELLO FIVE-STAR PER LINKED DATA

```
ex:p-4 rdf:type ex:Person ; ex:name "Pablo Picasso";  
ex:birth-place ex:p-18 ; ex:birth-time 1881.
```

4. **RDFa**<sup>6</sup> è uno schema di mark-up per l'incorporamento di descrizioni RDF in HTML e HTML5. RDFa fornisce un meccanismo per la pubblicazione di metadati su pagine Web, da cui possono essere estratti e raccolti da motori di ricerca. Ad esempio, nel seguente paragrafo HTML, la risorsa *p-4* viene definita utilizzando il testo del documento:

```
... <p about="http://example.org/p-4">  
In his paintings <span property="ex:name">Pablo Picasso</span>,  
born in <span property="ex:birth-time">1881</span>, expressed ...  
</p> ...
```

Il meccanismo GRDDL per Gleaning Resource Descriptions from Dialects of Languages<sup>7</sup> può essere utilizzato per estrarre RDF da una pagina Web.

RDFa è la versione Semantic Web dell'idea di 1) *microformati*<sup>8</sup> e 2) *microdata*. In tutti e tre gli approcci, sviluppati da diverse comunità, l'idea di fondo è la stessa: incorporare dati strutturati per le macchine da utilizzare in pagine HTML human-readable sul Web. Facendo embedding, la quantità di markup da scrivere può essere ridotta, perché non c'è bisogno di scrivere separatamente HTML e descrizioni strutturate di metadati.

Le principali differenze tra gli approcci descritti sono: i Microformati. I Microformati sono basati sull'utilizzo di tag HTML tradizionali, non specificatamente destinati allo scopo di rappresentare i metadati, in modo tale che le pagine Web possano essere renderizzate dai browser senza problemi per l'utilizzo umani. I Microdata, d'altra parte, utilizzano tag specifici di HTML5 per rappresentare i dati strutturati, e

---

<sup>6</sup> <http://www.w3.org/TR/xhtml1-rdfa-primer/>

<sup>7</sup> <https://www.w3.org/TR/grddl-primer/>

<sup>8</sup> <http://microformats.org/>

## CAPITOLO 3 - REQUISITI PER PUBBLICARE LINKED DATA

forniscono un insieme di vocabolari condivisi che i webmasters possono utilizzare per “marcare” le loro pagine in modo che possano essere comprese dai motori di ricerca.

Una specifica più importante che tratta l’incorporamento dei metadata in XHTML per motori di ricerca è Schema.org<sup>9</sup>, supportato dai principali motori di ricerca come Google, Bing, Yahoo, Yandex. La specifica è Semantic Web ed include vocabolari condivisi ontology-like per annotare i contenuti. Una delle principali differenze tra Schema.org, sviluppato da industrie, e RDFa, sviluppato da W3C, è che RDFa si basa sull’utilizzo di URI, il che rende più compatibile con Linked Data. Tuttavia, è possibile mappare descrizione Schema.org in rappresentazioni corrispondenti RDF. I vocabolari Schema.org sono disponibili in formati RDF diversi, JSON, e CSV. JSON (JavaScript Object Notation) è un formato lightweight data-interchange che è diventato molto popolare sul Web e Linked Data.

### 3.1.2. COLLEGARE DATI INTERNAMENTE ED ESTERNAMENTE

In una tripla RDF <S, P, O>, il soggetto S e il predicato P devono essere risorse identificate da un URI, mentre l’oggetto O può essere sia una risorsa o un literal (stringa, numero, data, ecc.). Un valore literal, ad esempio, il nome o età di una persona, non può avere altre proprietà proprie. Solo se l’oggetto O è un URI, allora viene stabilito un collegamento tra due risorse S ed O.

Un collegamento può collegare due risorse all’interno di un dataset, o due risorse da dataset separati. La potenza dell’approccio Linked Data si basa sulla creazione di collegamenti ricchi di dati di entrambi i tipi. È facile creare dataset RDF con solo proprietà literal, ma un dataset del

---

<sup>9</sup> <http://schema.org/>

### 3.1 MODELLO FIVE-STAR PER LINKED DATA

genere non è realmente Linked Data ma solamente una variante sintattica dei dati originali in RDF. La proporzione tra il collegamento di triple di dati in un dataset e la proporzione di collegamenti interni ed esterni può essere utilizzata come misura di *linkedness* di un dataset.

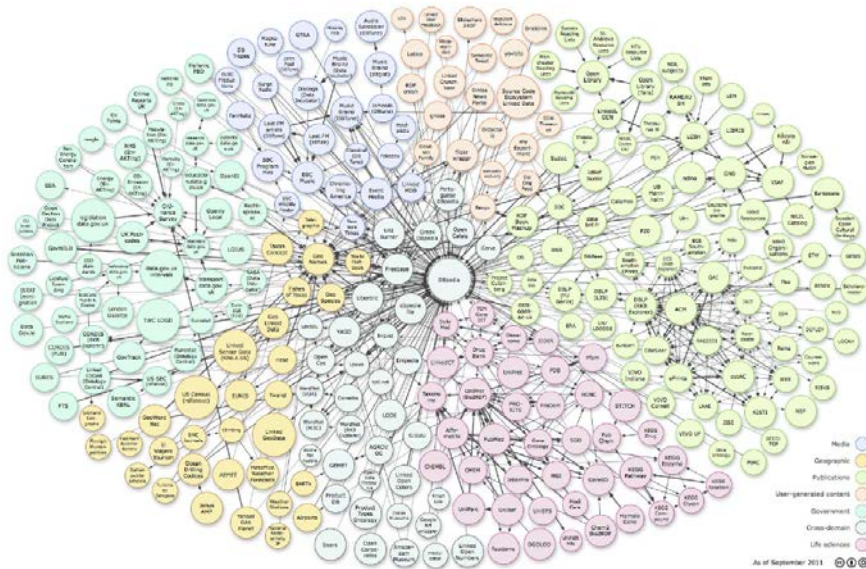


Figura 4: Il Linked Open Data Cloud consiste in dataset (rappresentati in bolle) in differenti domini (individuati dal colore), e mapping tra risorse simili nei dataset (illustrati con gli archi)

Molti dataset utilizzano URI indipendenti coniate con politiche proprie, e un concetto spesso ha diverse copie di sé stesso in diversi dataset con URI differenti. Per esempio, stessi luoghi e persone spesso appaiono in collezioni differenti con identificatori diversi. Al fine di collegare queste risorse correttamente, vengono generalmente creati gruppi di link esterni, i mapping dataset. Come in [Figura 4](#) viene illustrato il Linked Open Data Cloud<sup>10</sup>, dove ogni bolla è un dataset. Un arco tra due dataset A e B rappresenta un mapping M tra di loro ed è rappresentato come set di data link (un set di triple).

---

<sup>10</sup> <http://lod-cloud.net/>

## CAPITOLO 3 - REQUISITI PER PUBBLICARE LINKED DATA

Tipicamente vengono utilizzate per il mapping delle triple “owl:sameAs”  $\langle R_A, owl : sameAs, R_B \rangle$ , dove  $R_A$  è una risorsa in A ed  $R_B$  è una risorsa in B. Molti triplestore possono trattare automaticamente le risorse  $R_A$  e  $R_B$  come se fossero uguali, nascondendo dettagli di mapping all’utente. Possono essere utilizzati anche altri tipi di predicati in dataset mapping, come “rdfs:subClassOf” o “rdfs:subPropertyOf”, stabilendo gerarchie di classi e proprietà. Nel vocabolario standard SKOS<sup>11</sup>, possono essere creati mapping “imprecisi” utilizzando proprietà di mapping come “skos:closeMatch”.

### 3.2. REQUISITI PER INTERFACCE ED API

Un linked dataset RDF è pubblicato rendendolo disponibile tramite interfacce. Tipicamente, vengono fornite le seguenti:

1. **Linked Data Browsing.** Interfaccia browser Linked Data basata su deferenza di URI. Questa interfaccia rende i dataset consultabili da vari tabulatori e browser LD creati per il Web;
2. **SPARQL endpoint.** Pubblicato per interrogare i dati in modo standard per, ad esempio, applicazioni mash-up;
3. **Funzionalità di download.** Possibilità di scaricare i dati come un RDF data dump;
4. **Human interfaces.** Interfacce human end-user per la ricerca e la navigazione dei dati.

#### 3.2.1. LINKED DATA BROWSING

Nel Linked Data Browsing ogni risorsa viene renderizzata mostrando il literal e valori delle risorse come proprietà. Le risorse vengono viste come link HTML; una risorsa è renderizzata cliccando è seguendo il

---

<sup>11</sup> <http://www.w3.org/2004/02/skos/>

## 3.2 REQUISITI PER INTERFACCE ED API

browsing. Sono stati creati diversi browser LD, come Tabulator<sup>12</sup>, Sig.ma<sup>13</sup>, OpenLink Data Explorer<sup>14</sup>, Ontology-browser<sup>15</sup> e Falcon Explorer<sup>16</sup>.

### 3.2.2. ENDPOINT SPARQL

SPARQL Protocol and RDF Query Language (SPARQL)<sup>17</sup> è il linguaggio di query standard per i dati RDF. Possiede alcune somiglianze con SQL per database relazionali. Ci sono quattro forme di query disponibili:

1. **Query SELECT** si basa su un graph pattern, rappresentato utilizzando Turtle e variabili di query che iniziano con “?” per le risorse. Il modello è abbinato col grafo RDF sottostante, che risulta associando diverse variabili che corrispondono al pattern. I bindings sono i risultati delle query e vengono restituiti in forma tabellare. Per esempio, la query seguente restituisce una tabella di due colonne che lista i nomi e corrispondenti materiali di tutte le istanze della classe Artifact.

```
PREFIX ex: <http://museum.org/artifact/>
SELECT ?name ?material WHERE {
  ?artifact a ex:Artifact . ?artifact ex:name ?name .
  ?artifact ex:material ?material . }
```

2. **Query CONSTRUCT** viene utilizzato per estrarre informazioni da endpoint SPARQL, e trasforma i risultati in RDF;
3. **Query ASK** verifica solo se un determinato graph pattern corrisponde al grafo RDF sottostante, e restituisce vero o falso;

---

<sup>12</sup> <https://www.w3.org/2005/ajar/tab>

<sup>13</sup> <http://sig.ma/>

<sup>14</sup> <http://linkeddata.uriburner.com/ode/>

<sup>15</sup> <http://code.google.com/p/ontology-browser/>

<sup>16</sup> <http://ws.nju.edu.cn/explorer/>

<sup>17</sup> <https://www.w3.org/TR/rdf-sparql-query/>



## CAPITOLO 3 - REQUISITI PER PUBBLICARE LINKED DATA

4. **Query DESCRIBE** viene utilizzato per leggere le descrizioni RDF di un determinato URI.

Un endpoint SPARQL fornisce agli utenti un servizio dinamico per information retrieval tramite il protocollo http, senza la necessità di copiare e trasferire il dataset. Sono disponibili una varietà di software endpoint SPARQL, inclusi open source e sistemi commerciali, come ad esempio 4Store<sup>18</sup>, AllegroGraph<sup>19</sup>, Jena<sup>20</sup>, Sesame<sup>21</sup> e Virtuoso<sup>22</sup>.

È presente una lista degli endpoint SPARQL online nelle W3C wiki pages<sup>23</sup>.

### 3.2.3. FUNZIONALITÀ DI DOWNLOAD

La funzionalità di download è utile per, ad esempio, utenti esterni disposti ad analizzare, modificare o arricchire il dataset, o utilizzarlo in modalità locale e altre applicazioni.

### 3.2.4. HUMAN INTERFACES

Infine, possono essere create varie interfacce applicative del dataset per l'utente, come ontology browser nei sistemi bibliotecari ontologici. Sono stati creati diversi motori di ricerca per dati RDF, come ad esempio Falcons<sup>24</sup>, Sindice<sup>25</sup>, Swoogle<sup>26</sup>, SWSE<sup>27</sup> e Watson<sup>28</sup>.

---

<sup>18</sup> <http://4store.org/>

<sup>19</sup> <http://franz.com/agraph/allegrograph/>

<sup>20</sup> <http://jena.apache.org/>

<sup>21</sup> <http://www.openrdf.org/>

<sup>22</sup> <https://virtuoso.openlinksw.com/>

<sup>23</sup> <http://www.w3.org/wiki/SparqlEndpoints>

<sup>24</sup> <http://ws.nju.edu.cn/falcons/objectsearch/index.jsp>

<sup>25</sup> <http://www.sindice.com/>

<sup>26</sup> <http://swoogle.umbc.edu/>

<sup>27</sup> <http://saimuseiri-mode.com/>

<sup>28</sup> <http://watson.kmi.open.ac.uk/WatsonWUI/>



## CAPITOLO 4

### SCHEMI METADATA

Questo capitolo riguarda i metadati, la base del Semantic Web e dei Linked Data. Nella prima sezione vengono classificati diversi tipi di metadati. Successivamente, vengono presentati schemi di metadati utilizzati per scopi diversi a rappresentare contenuti BC.

#### 4.1. TIPI DI METADATI

Metadato significa letteralmente “dati sui dati”. L’American Library Association definisce nel 1999 i metadati nel seguente modo:

*Metadata is structured, encoded data that describe characteristics of informationbearing entities to aid in the identification, discovery, assessment, and management of the described entities.*

Nel contesto del Web, la nozione di metadati ha ricevuto una connotazione più ampia:

*Metadata is machine understandable information for the Web.*

In quest’ultima connotazione, le nozioni di strutture di metadati e simboli utilizzati per la *knowledge representation* (KR), diventano strettamente correlate. KR è un’area del campo di ricerca dell’intelligenza artificiale, che rappresenta, e inferenza conoscenza in termini di strutture di simboli. I metadati sul Semantic Web vengono

## 4.1 TIPI DI METADATI

utilizzati non solo per descrivere oggetti da collezione, ma per modellare il mondo reale sottostante e anche la nostra conoscenza di esso.

I metadati sono stati tradizionalmente creati da professionisti dell'informazione durante la classificazione e l'indicizzazione di oggetti di collezioni durante i processi di catalogazione. Al giorno d'oggi, vengono prodotti sempre più metadati da figure non professioniste, pubblicando il loro contenuto come immagini e video online. Inoltre, gran parte dei metadati sul Web sono stati prodotti automaticamente utilizzando tecniche e strumenti di *data mining e information retrieval*.

I metadati relativi a collection items possono essere divisi in diversi tipi:

1. **Metadata amministrativi**, come l'acquisizione o location information, viene utilizzato per la gestione delle collezioni e information resources;
2. **Metadata descrittivo**, ad esempio, il tipo di materiale di un artefatto, viene utilizzato per identificare e descrivere i contenuti delle collezioni e le relative information resources;
3. **Metadata di conservazione**, ad esempio, le informazioni sulla condizione fisica di un collection item, sono legate alla gestione della conservazione della collezione;
4. **Metadata tecnico**, come l'hardware o la documentazione software, che descrive come si comporta una funzionalità del sistema;
5. **Metadata utilizzo**, come rights metadata e log di ricerca, è collegato al tipo e al livello di utilizzo delle collezioni e information resources.

I metadati possono essere divisi tramite indicizzazione e visualizzazione delle informazioni:

## CAPITOLO 4 - SCHEMI METADATA

1. **Indexing metadata** viene utilizzato dai motori di ricerca e sistemi di raccomandazione per la ricerca di informazioni memorizzate;
2. **Displaying metadata** viene utilizzato per specificare, ad esempio, l'ordine particolare in cui i dati devono essere mostrati in un'interfaccia dell'applicazione.

Quando si utilizzano schemi XML, l'informazione sull'ordine degli elementi è fissata dallo schema e può essere utilizzato per motivi di visualizzazione. Tuttavia, quando si utilizzano dati RDF, l'ordine delle proprietà di una risorsa non è specificato. Ci sono modi per specificare questo ordine dell'informazione in RDF: 1) RDF contiene alcuni costrutti per rappresentare l'ordinazione, come le liste. 2) È possibile utilizzare l'ordinamento della serializzazione XML di RDF in maniera personalizzata. 3) È possibile creare un metaschema personalizzato in RDF che specifica in modo esplicito l'ordine delle proprietà rispetto alle risorse.

I contenuti culturali nelle collezioni dei musei, biblioteche e altre repository di contenuti sono descritte utilizzando schemi di metadati, detti anche *annotation schemas*. Uno schema di metadati è un modello che specifica il formato in cui deve essere rappresentato il metadato.

Gli schemi di metadati sono stati sviluppati per scopi diversi:

1. **Web schemas** per descrivere i documenti e gli oggetti sul web;
2. **Cataloging schemas** descrivono un insieme di elementi obbligatori e facoltativi, cioè, le proprietà con la quale i metadati per un elemento devono essere descritti, e i vincoli per l'inserimento di valori negli elementi. I valori degli elementi metadata vengono compilati quando avviene la catalogazione di un item. Per esempio, metadati di catalogazione descrittivi possono indicare il titolo, l'autore(i), e l'argomento(i) di un libro in una biblioteca;

3. **Harmonization schemas** sono modelli ontologici sui quali vengono trasformati modelli di metadati eterogenei per interoperabilità semantica. Questi schemi sono destinati ad essere utilizzati per la catalogazione, ma possono essere anche fonte di ispirazione per lo sviluppo di tali schemi;
4. **Harvesting schemas** sono utilizzati per la raccolta di metadati eterogenei provenienti da schemi diversi, con l'obiettivo di creare un portale applicativo aggregato.

### 4.2. SCHEMI WEB

Lo scopo degli schemi Web è quello di fornire elementi metadati armonizzati per descrivere documenti e altri oggetti sul Web. Il modello più notevole è quello del Dublin Core e le sue applicazioni. Gli elementi del DC e i suoi derivati sono spesso utilizzati anche nella catalogazione degli schemi, e la distinzione tra schema Web e schema di catalogazione può essere vaga.

#### 4.2.1. DUBLIN CORE

Dublin Core (DC) è un modello di metadati e framework ampiamente utilizzato nelle biblioteche e organizzazioni.

Il cuore del DC è il DC Metadata Element Set<sup>1</sup> (DCES), che contiene 15 elementi standardizzati elencati nella Tabella 2.

title	creator	subject	description	publisher
contributor	date	type	format	identifier
source	language	relation	coverage	rights

Tabella 2: DC Metadata Element Set

---

<sup>1</sup> <http://dublincore.org/documents/dces/>

## CAPITOLO 4 - SCHEMI METADATA

Tutti gli elementi DC hanno cardinalità (0, n), cioè sono opzionali e possono avere più valori (ad esempio, un libro può avere diversi creatori). Gli elementi sono rappresentati in ordine title, creator, subject, ecc., come mostrato in Tabella 2.

I 15 elementi di base possono essere estesi in modo interoperabile utilizzando il principio “dumb-down”. Ciò significa che un nuovo elemento “qualificato” può essere introdotto nel vocabolario con un significato più specifico di un elemento esistente. In questo modo è possibile interoperare descrizioni più specifiche con quelle meno specifiche. Ad esempio, l’elemento “manufacturer” sarebbe più specifico di “creator”. Quindi, durante la ricerca per creators potrebbe essere mostrato come risultato. Il vocabolario può anche essere esteso introducendo nuovi elementi. Il risultato all’estensione e all’aggiunta di elementi nel Core è chiamato *DC application*.

L’attuale DC recommendation è quella di utilizzare una versione estesa degli elementi Core chiamata DCMI (Metadata) Terms<sup>2</sup>. È diventato uno dei vocabolari più popolari in uso, e le sue specifiche sono compatibili con quelle del Semantic Web e Linked Data. DCMI Terms contiene 55 elementi definiti nel seguente namespace: <http://dublincore.org/documents/dcmi-terms/>.

Le specifiche del DCMI Terms include la *encoding schemes* che vengono utilizzati per armonizzare le modalità con cui sono espressi i valori degli elementi. In primo luogo, ci sono *vocabulary encoding schemes* che sono riferimenti ai vocabolari, come la Library of Congress Subjects (LCSH) o Medical Subject Heading Headings (MeSH). In secondo luogo, vi è anche menzionata una lista di *syntax encoding schemes*. Vale a dire i riferimenti ai metodi standard di codifica di date, lingue, codici di paesi, ecc.

---

<sup>2</sup> <http://dublincore.org/documents/dcmi-terms/>

Ci sono anche un insieme di classi definite e utilizzate come range constraints per alcuni elementi. Per esempio, il valore del creatore si presume essere un'istanza della classe Agent identificata dallo URI: <http://purl.org/dc/terms/Agent>.

Infine, la specifica contiene un generico DCMI *Type Vocabulary* di classi, come *Collection*, *Dataset*, *Event*, e *Image*, le quali istanze possono essere annotate utilizzando il metadata schema.

Oltre a queste specifiche tecniche, la DC community ha pubblicato varia linee guida per l'utilizzo del framework DC, ad esempio, su come creare profili applicativi.

### 4.2.2. VRA CORE CATEGORIES

DC viene utilizzato come base per schemi di metadati culturali più dettagliati, come per il Visual Resource Association's (VRA) Core Categories<sup>3</sup>. L'element set nel Core fornisce gli elementi per la descrizione di 1) lavori di cultura visiva, 2) immagini che li documentano, 3) collezioni di oggetti. È possibile anche indicare le relazioni tra opere e le loro immagini, quindi le collezioni possono contenere sia opere o immagini. La maggior parte degli elementi VRA sono definiti come *refinements* (subproperties) dei DC elements.

Un esempio di un'istanza di metadato VRA è dato nella notazione Turtle sottostante:

```
rijks:artefactSK-C-K vra:type vra:Work;  
vra:title "The Night Watch"; vra:date "1642" ;  
vra:creator ulan:500011051; # Rembrandt  
vra:subject iconclass:45F31; # Call to arms  
vra:culture tgn:7006952; # Amsterdam  
vra:material aat:30015050. # Oil paint
```

---

<sup>3</sup> <http://www.vraweb.org/>



## CAPITOLO 4 - SCHEMI METADATA



Figura 5: "The Night Watch" by Rembrandt van Rijn

Il record rappresenta il dipinto di Rembrandt "The Night Watch" nella collezione presente nel museo Rijks ad Amsterdam (Figura 5). Lo schema ha proprietà come *vra:types* (il tipo dell'opera d'arte come riferimento per il vocabolario VRA), *vra:title* (titolo literal dell'opera d'arte), *vra:creator*, *vra:subject*, *vra:culture*, *vra:material*. Gli Element values con namespace sono referenziati all'ontologia di dominio sottostante: ULAN<sup>4</sup> è un'ontologia per actors, Inconclass<sup>5</sup> per descrizioni iconografiche, TGN<sup>6</sup> per luoghi storici, e AAT<sup>7</sup> per concetti d'arte, architetture e cultura.

VRA è stato approvato come extension schema per oggetti METS<sup>8</sup> che contengono immagini di risorse del patrimonio culturale. Nell'esempio precedente, VRA è stato utilizzato come schema Web ma il sistema è utilizzato anche come base per sistemi di catalogazione.

---

<sup>4</sup> <http://www.getty.edu/research/tools/vocabularies/ulan/about.html>

<sup>5</sup> <http://www.iconclass.org/>

<sup>6</sup> <http://www.getty.edu/research/tools/vocabularies/tgn/about.html>

<sup>7</sup> <http://www.getty.edu/research/tools/vocabularies/aat/about.html>

<sup>8</sup> <http://www.loc.gov/standards/mets/>

### 4.3. SCHEMI DI CATALOGAZIONE

Gli oggetti da collezioni sui BC di diversi tipi hanno diverse caratteristiche. La loro descrizione richiede spesso diversi modelli di metadati. Ad esempio, un dipinto può descrivere eventi che si sono verificati nella storia o nella fantasia, mentre un manufatto, come un tavolo, potrebbe essere caratterizzato dal tipo e materiale di costruzione. Questa sezione presenta alcuni modelli di metadati sviluppati ed utilizzati per la catalogazione nelle biblioteche, musei e archivi.

#### 4.3.1. CATEGORIES FOR THE DESCRIPTION OF WORKS OF ART (CDWA)

Categories for the Description of Works of Art (CDWA) (1) è un modello metadata mirato per la catalogazione delle opere d'arte. Il sistema CDWA è più esteso rispetto a VRA ed include 532 categorie, vale a dire elementi e subcategories. Comunque, è possibile trovare una versione Lite disponibile di CDWA che contiene un minimal set di *core categories* selezionate per identificare univocamente e senza ambiguità un'opera d'arte o architettonica. Viene riportata una sintesi delle categorie principali del CDWA divisi in gruppi per dare un'idea generale del sistema:

1. OBJECT, ARCHITECTURE, OR GROUP: Catalog Level;  
Object/Work Type; Classification Term; Title or Name;  
Measurements Description; Materials and Techniques Description;  
Creator Description; Creator Identity; Creator Role; Creation Date;  
Subject Matter; Current Location Repository Name/Geographic  
Location; Current Repository Numbers;
2. RELATED TEXTUAL REFERENCES AUTHORITY: Brief  
Citation; Full Citation;

## CAPITOLO 4 - SCHEMI METADATA

3. CREATOR IDENTIFICATION AUTHORITY: Name; Source;  
Display Biography; Birth Date; Death Date;  
Nationality/Culture/Race; Life Roles;
4. PLACE/LOCATION AUTHORITY: Place Name; Source; Place Type;  
Broader Context;
5. GENERIC CONCEPT AUTHORITY: Term; Source; Broader  
Context; Scope Note; Source;
6. SUBJECT AUTHORITY: Subject Name; Source; Broader Context.

Regole ed esempi per l'utilizzo del core subset del CDWA categories e del VRA Core Categories sono documentate in (2).

### 4.3.2. SPECTRUM

Standard ProcEdures for CollecTions Recording Used in Museums (SPECTRUM)<sup>9</sup> è uno che schema di metadati XML-based per la catalogazione di collezioni per musei. È uno standard di gestione di collezioni open e disponibile gratuitamente. Originariamente UK-based, ora è utilizzato in circa 7.000 musei in 40 paesi, e include anche traduzioni e localizzazioni.

Questo standard comprende lo schema SPECTRUM, lo Standard wiki, e le SPECTRUM terminologies. Comprende linee guida procedurali per la catalogazione e allo stesso tempo fornisce un livello di metadati per la raccolta della documentazione museale. Come nel caso del CDWA, lo schema è vasto, ed è presente una versione semplificata chiamata SPECTRUM Essentials creata per la documentazione “light weight” per musei più piccoli.

---

<sup>9</sup> <http://www.collectionstrust.org.uk/spectrum>

## 4.3 SCHEMI DI CATALOGAZIONE

### 4.3.3. FORMATI METADATA IN BIBLIOTECHE

Le biblioteche hanno una lunga tradizione nel creare metadata records sulle pubblicazioni. Uno standard ampiamente utilizzato è MARC (Machine-Readable Cataloging)<sup>10</sup>, sviluppato dal Library of Congress negli anni '60. È un modello sofisticato per l'uso nelle biblioteche, quindi risulta difficile da apprendere ed utilizzare dal punto di vista umano. Ad esempio, il sistema si basa su tag a tre cifre difficili da ricordare e codici subfield. Questo modello è stato quindi modernizzato e trasformato in MARC-XML<sup>11</sup> ed è presente una versione più human-friendly disponibile, Metadata Object Description Standard (MODS)<sup>12</sup>. Il Metadata Authority Description Schema (MADS)<sup>13</sup> è uno schema XML correlato per fornire metadata relativi ad actors, events, terms (topics, geographics, genres, ecc.).

In molti casi la documentazione di un oggetto contiene descrizioni su di esso, come ad esempio i metadati di un documento, le immagini su di esso, recensioni connesse, ecc. la Metadata Encoding and Transmission Standard (METS)<sup>14</sup> è uno standard “wrapper” per la rappresentazione di tali raccolte di descrizioni dei metadati. È possibile utilizzare DC, MODS e VRA per metadata di collection items.

### 4.3.4. FORMATI METADATI NEGLI ARCHIVI

Gli archivi sono il terzo principale tipo di organizzazione di memoria oltre a musei e biblioteche. Archivi, biblioteche e musei differiscono l'uno dall'altro per quello che ricordano e per chi fruisce del servizio.

---

<sup>10</sup> <http://www.loc.gov/marc/>

<sup>11</sup> <http://www.loc.gov/standards/marcxml/>

<sup>12</sup> <http://www.loc.gov/standards/mods/>

<sup>13</sup> <http://www.loc.gov/standards/mads/>

<sup>14</sup> <http://www.loc.gov/standards/mets/>

## CAPITOLO 4 - SCHEMI METADATA

- Un archivio conserva il materiale di base su cui si basa la nostra comprensione storica; i materiali sono prove legali e storiche degli eventi. La memoria legale e storica richiede un alto grado di fiducia degli utenti nell'autenticità e integrità degli atti e documenti. I materiali in archivi e biblioteche sono record unici creati da organi sociali, individui e famiglie. I documenti generati da una persona giuridica, individuo o famiglia è chiamato *collection* o *fond*:

*A fond is a whole of the documents, regardless of form or medium, organically created and/or accumulated and used by a particular person, family, or corporate body in the course of that creator's activities and functions.*

- Le biblioteche raccolgono libri singoli pubblicati e i seriali, o insiemi delimitati di singoli elementi. Gli elementi in collezione non sono unici, nel senso che esistono più copie di una pubblicazione. Qualsiasi copia generalmente soddisferà il bisogno del cliente di informazione ugualmente;
- I musei differiscono molto gli uni dagli altri in termini di materie che conservano. Anche all'interno di un singolo museo, le collezioni possono essere molto eterogenee tra loro, ad esempio, l'arte, manufatti, libri, fotografie, ecc. gli elementi di una collezione sono tipicamente unici, come i dipinti, ma possono esistere anche più copie, come nelle biblioteche.

Musei, biblioteche e archivi forniscono servizi diversi, anche se sovrapponibili. I musei e le biblioteche generalmente offrono un servizio alle comunità pubbliche, educative e accademiche. Molti archivi a loro volta, offrono servizi alla legge, funzionando come memoria istituzionale degli organi sociali specifici. Le agenzie governative, istituzioni pubbliche, e le imprese hanno requisiti legali relativi al mantenimento dei registri. Queste differenze di materiali di raccolta e comunità che utilizzano questi servizi rappresentano le differenze tra le pratiche di

## 4.4 CONCEPTUAL HARMONIZATION SCHEMAS

catalogazione di metadati tra musei, biblioteche e archivi. Di conseguenza, sono stati sviluppati diversi formati standard di metadati utilizzati negli archivi.

### 4.4. CONCEPTUAL HARMONIZATION SCHEMAS

La definizione di schema affronta il problema di *interoperabilità sintattica* e *semantica* delle descrizioni degli oggetti contenuti. Ovviamente, i problemi di interoperabilità possono essere affrontati in modo efficace utilizzando un unico schema. Tuttavia, sono necessari e vengono utilizzati diversi schemi per diversi tipi di oggetti in applicazioni che si occupano di contenuti cross-domain. Questa prima sezione discute approcci per schema integration e presenta approcci basati su ontologie per armonizzare i modelli metadata.

#### 4.4.1. APPROCCI ALL'INTEROPERABILITÀ SEMANTICA

L'interoperabilità sintattica può essere ottenuta mediante l'armonizzazione delle forme strutturali che rappresentano i dati (ad esempio, utilizzando elementi DC o uno schema XML), e fissando le convenzioni di codifica per i valori dell'elemento (ad esempio, vocabolari controllati, formato della data, coordinate, ecc.). L'interoperabilità semantica è ottenuta tramite convenzioni condivise per interpretare le rappresentazioni sintattiche, per esempio, che la proprietà `dc:subject` descrive l'oggetto di un documento come un insieme di keyword tratta da un'ontologia.

Fare diversi schemi di metadati semanticamente interoperabili richiede due attività secondarie. In primo luogo, l'interoperabilità semantica dei valori degli elementi deve essere affrontata utilizzando vocabolari e ontologie (in comune). In secondo luogo, deve essere facilitata l'interoperabilità tra elementi di diverso schema.

## CAPITOLO 4 - SCHEMI METADATA

Ci sono due approcci principali per la preparazione di schemi di metadati interoperabili: un approccio semplice per mappare elementi di schema con altri è utilizzare il principio di dump-down. Un altro approccio è quello di creare un modello di ontologia concettuale nel quale schemi differenti possono essere trasformati e armonizzati. Europeana Semantic Elements (ESE) si presenta come un esempio al primo approccio. Di seguito, alcuni modelli che usano il secondo approccio sono: Europeana Data Model (EDM), CIDOC CRM proveniente dal dominio dei musei e FRBR dal dominio bibliotecario.

### 4.4.2. EUROPEANA SEMANTIC ELEMENTS (ESE)

Europeana Semantic Elements (ESE)<sup>15</sup> è un esempio di un'applicazione tipica DC. ESE è costituito da 15 elementi originali DC raffinati con una selezione di 21 elementi aggiuntivi da DCMI Terms. Inoltre, vengono introdotti 14 nuovi elementi nel namespace Europeana, come *country* e *dataProvider*.

ESE è utilizzato massicciamente per l'iniziativa europea di creare un portale pan-European sui BC. La specifica è fornita come schema XML per la convalida automatica dei dati ed è la normativa per i content providers di Europeana. Con armonizzazione dei dati si intende che i dati provenienti da diversi sistemi di raccolta vengano trasformati in un singolo schema. ESE è un'estensione dello schema DCMI Terms e accetta tutti i DC terms, compresi quelli che non possono essere importati nel portale.

### 4.4.3. EUROPEANA DATA MODEL (EDM)

A dispetto del suo nome, l'uso di ESE in Europeana non è particolarmente "semantico" da un punto di vista Linked Data. Per risolvere questo problema, è stato proposto il Europeana Data Model

---

<sup>15</sup> <http://pro.europeana.eu/page/ese-documentation>

## 4.4 CONCEPTUAL HARMONIZATION SCHEMAS

(EDM)<sup>16</sup> basato su RDF ed utilizzato nella pubblicazione di linked open data dei contenuti di Europeana<sup>17</sup> (3). I dataset contenuti in questa pubblicazione contengono 2,4 milioni di testi, immagini, video e suoni raccolti da Europeana, e sono liberamente disponibili in formato RDF e XML. ESE è un sottoinsieme di EDM e può essere mappato direttamente su di essa.

EDM è un framework Semantic Web-based per la rappresentazione cross-domain di collection metadata nei musei, biblioteche e archivi. Il modello facilita descrizioni più ricche dei contenuti di ESE, e collegamento tra i dati basato su risorse condivise. EDM fa una distinzione semantica tra le descrizioni intellettuali o tecniche di oggetti che vengono raccolti da content providers, l'oggetto delle descrizioni di cui si parla, e rappresentazioni digitali dell'oggetto. Per esempio, il dipinto "Mona Lisa" è al Louvre, ma ci sono numerose copie, fotografie, testi, disegni, e statue che la raffigurano in vari musei europei. I diversi punti di vista dello stesso oggetto possono essere rappresentati utilizzando uno speciale meccanismo di proxy. Si prevede che il meccanismo di proxy venga sostituito in futuro da grafi etichettati dopo che il modello verrà standardizzato.

A differenza di ESE, EDM non è uno schema fisso che determina il modo di rappresentare i dati, ma piuttosto un quadro concettuale o ontologico in cui possono essere collegati e interoperati modelli più specifici.

### 4.4.4. CIDOC CONCEPTUAL REFERENCE MODEL (CRM)

Molti schemi elencati, come EDM, non sono destinati ai fini catalogativi, ma sono utilizzati per scopi di armonizzazione e integrazione dei dati. Un

---

<sup>16</sup> <http://pro.europeana.eu/web/guest/edm-documentation>

<sup>17</sup> <http://labs.europeana.eu/api/linked-open-data-introduction>



## CAPITOLO 4 - SCHEMI METADATA

esempio di quanto detto è il CIDOC Conceptual Reference Model (CIDOC CRM) (4) (cf. (5)).

L'idea è che vengano utilizzati diversi modelli di metadati di organizzazioni di memoria per i diversi tipi di oggetti da collezione per essere trasformati in una scheda conforme a un modello ontologico semantico più fondamentale. Il modello serve come guida per la creazione di forma, una lingua per l'analisi, e formato di trasporto dati.

Dublin Core può essere caratterizzato come *centric document* o *centric object*: i suoi elementi vengono utilizzati per descrivere documenti ed oggetti. Al contrario, CIDOC CRM è *event centric* incentrato sul concetto più fondamentale di evento legato alla creazione, l'utilizzo e la manutenzione di documento di raccolta e oggetti. Gli eventi possono essere considerati una sorta di "colla semantica" che mette in relazione aspetti diversi della conoscenza (objects, actors, times, places, names, ecc.) in una collezione BC con le altre in modo armonizzato. Sono stati creati vari mapping e strumenti di mapping per trasformare modelli diversi di metadati in CIDOC CRM, tra cui Dublin Core, LIDO<sup>18</sup>, FRBR<sup>19</sup> e EDM<sup>20</sup>, e disponibili<sup>21</sup> in un elenco completo con la documentazione.

CIDOC CRM "fornisce definizioni e una struttura formale per descrivere i concetti e le relazioni implicite ed esplicite utilizzate nella documentazione del patrimonio culturale"<sup>22</sup>. Questo include non solo la rappresentazione degli oggetti, ma anche le loro informazioni di provenienza potenzialmente complesse.

Le specifiche standard (6) includono 98 entità di classe (numerata da E1-E98) e un set di 181 proprietà (P1-P181) che collegano entità e

---

<sup>18</sup> <http://www.lido-schema.org/>

<sup>19</sup> <http://www.ifla.org/VII/s13/frbr/>

<sup>20</sup> <http://pro.europeana.eu/share-your-data/data-guidelines/edm-documentation>

<sup>21</sup> <http://www.cidoc-crm.org/node/209>

<sup>22</sup> <http://www.cidoc-crm.org/>

#### 4.4 CONCEPTUAL HARMONIZATION SCHEMAS

proprietà le une con le altre. Le classi fondamentali del modello sono presentate in [Figura 6](#). Inoltre, le proprietà sono organizzate in una gerarchia subproperty, come in RDFS.

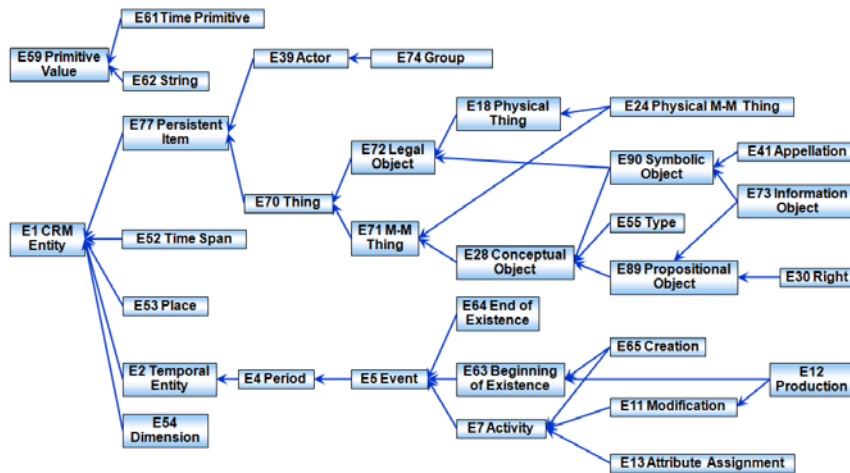


Figura 6: Gerarchia Core dei componenti del CIDOC CRM. Gli archi denotano relazioni subclass-of.

La trasformazione di collezioni di metadata in questo modello prevede la creazione di un'istanza di classe nel modello. Ad esempio, il data record di un artefatto può essere un'istanza basata sulla classe *E22 Man-Made Object*. Le relazioni istanza classe sono espresse utilizzando la proprietà *P2 has type of*. All'istanza dell'oggetto è dato un identificatore e nome chiamato *appellations (E41 Appellation)*, informazioni sull'acquisizione e di proprietà, informazioni sulla posizione come riferimenti a luoghi in un'ontologia, informazioni storiche sugli eventi correlati all'oggetto, descrizioni della materia, e le descrizioni fisiche dell'oggetto (ad esempio, *P43 has dimension*). I valori di annotazione delle proprietà sono in genere i riferimenti alle risorse. Ad esempio, i valori (range) di *P52 has current owner* sono actors (*E39 Actor*). Durante la trasformazione, le istanze di tali valori devono essere create in base ai collection metadata e/o acquisiti da ontologie condivise.

## CAPITOLO 4 - SCHEMI METADATA

Gran parte dei contenuti sono descritti in termini di eventi (*E5 Event*) e attività di vario tipo, come ad esempio *E10 Transfer of Custody*, *E11 Modification*, *E65 Creation* o *E87 Curation Activity*. Questi sono caratterizzati dal tempo (ad esempio, *P10 falls within* e *P4 has time-span*), luogo (ad esempio, *P7 took place at*), e attori coinvolti (ad esempio, *P11 had participant*).

CIDOC CRM è in uso in diverse organizzazioni ed è mantenuto online da una comunità sotto cura del ICOM CIDOC, l'International Committee for Documentation<sup>23</sup> nel dominio museale.

### 4.4.5. FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS (FRBR)

DC ha origine dal dominio applicativo delle biblioteche e si concentra sulla rappresentazione di metadati relativi ai documenti, cioè, raccolte di articoli nel patrimonio bibliotecario. Le raccolte di articoli possono essere collegate tra loro in diversi modi sulla base di processi reali di creazione intellettuale.

Ad esempio, uno dei lavori di Omero l'Odissea è stata realizzata in varie espressioni, cioè in versioni greche e tradotta in altre lingue. Altre pubblicazioni sono basate sull'Odissea, come ad esempio, gli adattamenti cinematografici, che a loro volta possono essere disponibili in diverse versioni. Ognuna di queste "espressioni" possono essere "manifestate", cioè, disseminate in vari formati fisici, come ad esempio in copertina rigida o in formato tascabile, o VHS e DVD. Gli articoli fisici posseduti dalle biblioteche sono esemplari di queste manifestazioni. Al fine di fornire i bibliotecari e i clienti nella ricerca del materiale BC relativo a opere creative e alle loro pubblicazioni, sono necessari modelli

---

<sup>23</sup><http://icom.museum/the-committees/international-committees/international-committee/international-committee-for-documentation/>

#### 4.4 CONCEPTUAL HARMONIZATION SCHEMAS

concettuali più versatili del DC. Per prendere queste relazioni in considerazione, l'International Federation of Library Associations and Institutions (IFLA), l'organismo nazionale leader nel settore delle biblioteche e dei servizi di informazione, ha sviluppato e standardizzato il "Functional Requirements for Bibliographic Records (FRBR)"<sup>24</sup>, un framework e una famiglia di standard per rappresentare metadati concettuali nel dominio bibliotecario. (7) Il termine "funzionale" in queste specifiche significa che l'obiettivo del lavoro concettuale sottostante e le sue implementazioni è quello di supportare quattro attività o funzioni base delle biblioteche, eseguite da clienti o bibliotecari.

1. **Find.** Trovare un'entità o set di entità corrispondenti ai criteri indicati;
2. **Identify.** Identificare un'entità (confermare che l'entità trovata corrisponde all'entità ricercata);
3. **Select.** Selezionare un'entità che è appropriata alle esigenze degli utenti;
4. **Obtain.** Ottenere l'accesso all'entità descritta.

La famiglia di standard FRBR contiene tre modelli concettuali entità-relazione, elencate nella Tabella 3, ciascuna delle quali contempla un aspetto dei dati registrati nei record bibliografici e proprietari.

Nome	Focus	Data
FRBR	Functional Requirements for Bibliographic Records	1998
FRAD	Functional Requirements for Authority Data	2009
FRSAD	Functional Requirements 4 Subject Authority Data	2010

Tabella 3: Famiglia FRBR del modello metadata

---

<sup>24</sup><http://www.ifla.org/publications/functional-requirements-for-bibliographic-records>

## CAPITOLO 4 - SCHEMI METADATA

Di seguito, l'introduzione a questi modelli.

### **4.4.6. FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS (FRBR)**

A differenza di un centric DC, FRBR fa una distinzione semantica tra entità di lavoro, espressioni, manifestazioni, e item riguardanti record bibliografici. Un lavoro è una creazione intellettuale o artistica distinta in senso astratto, ad esempio, una storia come quella di William Shakespeare "A Midsummer Night's dream". Un lavoro è realizzato attraverso diverse espressioni, cioè, realizzazioni intellettuali o artistiche del lavoro. Per esempio, lo spettacolo "A Midsummer Night's Dream" potrebbe essere espresso in un'edizione critica, una versione censurata o tradotto. Ogni espressione può essere realizzata in diverse manifestazioni, cioè, in forme fisiche di espressione, come ad esempio una particolare pubblicazione di una determinata traduzione. Infine, le manifestazioni sono esemplificate da oggetti, singoli esemplari di una manifestazione, come i libri negli scaffali di una biblioteca o copie del film in un archivio audiovisivo. Un elemento può esemplificare una sola manifestazione. Questo modello è illustrato in [Figura 7](#), dove le frecce con testa singola indicano le relazioni uno-a-uno, mentre le frecce a doppia testa indicano le relazioni uno-a-molti. In breve: con FRBR è possibile rappresentare con precisione i diversi tipi di elementi relativi a una creazione intellettuale o artistica come "A Midsummer Night's Dream".

## 4.4 CONCEPTUAL HARMONIZATION SCHEMAS

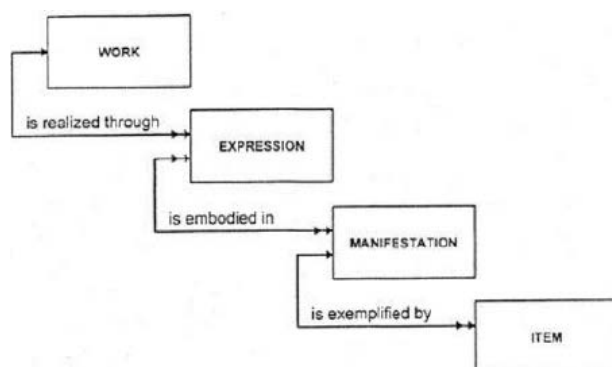


Figura 7: Entità e relazioni primarie in FRBR. Grafico da (7)

### 4.4.7. FUNCTIONAL REQUIREMENTS FOR AUTHORITY DATA (FRAD)

Functional Requirements for Authority Data (FRAD) (8) precedentemente noto come Functional Requirements for Authority Record (FRAF), estende il modello FRBR con 1) l'aggiunta di attributi per actors e relazioni tra di essi, 2) nuove relazioni tra gli actors e la loro appellations, 3) la gestione da parte delle biblioteche delle informazioni relative agli actors.

### 4.4.8. FUNCTIONAL REQUIREMENTS FOR SUBJECT AUTHORITY DATA (FRSAD)

Il terzo modello concettuale della famiglia FRBR è Functional Requirements for Subject Authority Data (FRSAD)<sup>25</sup> (9). FRSAD è un modello per le relazioni tra lavori e subjects. FRSAD estende il modello FRBR con nuove entità e relazioni nello stesso modo di FRAD. Le relazioni vengono presentate utilizzando le entità nella Tabella 4 che sono già state dichiarate in FRBR.

Entità	Significato
--------	-------------

<sup>25</sup> <http://www.ifla.org/node/5849>

## CAPITOLO 4 - SCHEMI METADATA

CONCEPT	Nozione o idea astratta
OBJECT	Oggetto materiale
EVENT	Azione o occorrenza
PLACE	Luogo

Tabella 4: Entità FR SAD

L'idea è che queste entità, così come le entità primarie di FRBR e FR SAD, vengono utilizzate come descrizioni soggetto chiamate *themas* (entità THEMA) che sono identificate da denominazioni (entità NOMEN), vale a dire, da qualsiasi segno o sequenza di segni (alfanumerico, simbolico, sonoro, ecc.) di un thema conosciuto, riferite o indirizzate alle relazioni tra entità nel modello FRBR.

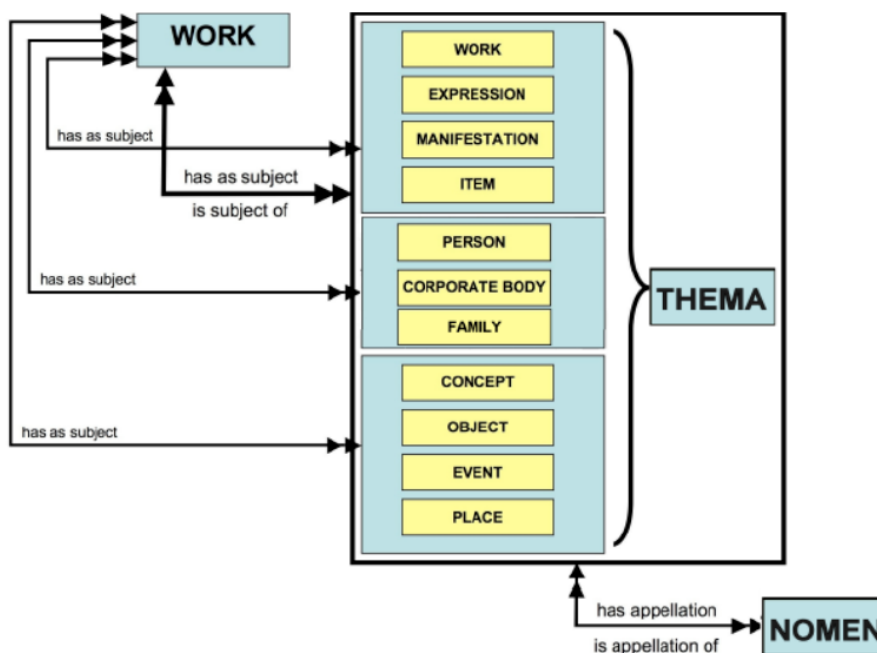


Figura 8: Relazioni FR SAD a FRBR e FRAD. Grafico da (9)

### 4.4.9. FRBR OBJECT ORIENTED (FRBROO)

I lavori di normalizzazione sul FRBR nei domini bibliotecari e su CIDOC CRM nei domini museali condividono lo stesso obiettivo: rendere l'informazione più reperibile armonizzando i formati eterogenei in uso.

## 4.5 HARVESTING SCHEMAS: LIDO

Inoltre, entrambe i lavori si basano su modelli concettuali semantici entità-relazione simili. Il passo successivo è quello di applicare lo stesso approccio di armonizzazione su FRBR e CIDOC CRM. A questo scopo, è stato costituito un gruppo di lavoro nel 2003. Il modello risultante, FRBR object oriented (FRBRoo) (10), è un'estensione del CIDOC CRM, dove le entità FRBR vengono aggiunte nella struttura entità gerarchica del CIDOC CRM (cfr. [Figura 6](#)) e le relazioni e gli attributi utilizzati nei modelli sono allineati. FRBRoo aggiunge in FRBR gli aspetti di event centric di CIDOC CRM, e una serie di altri parametri. L'obiettivo di FRBRoo è quello di creare un modello di un'ontologia per armonizzare i metadati bibliografici da biblioteche con le collezioni di metadati dai musei. È pianificato un allineamento tra EDM e FRBRoo.

## 4.5. HARVESTING SCHEMAS: LIDO

I format metadati vengono anche utilizzati per scopi di harvesting (raccolta), fornendo un modello comune su cui diversi publishers possono trasformare i propri database. Questa sezione presenta un recente sviluppo verso questo argomento: LightWeight Information Describing Objects (LIDO)<sup>26</sup>. LIDO è destinato all'utilizzo, ad esempio, nelle versioni del portale Europeana linked data-based.

LIDO è un harvesting schema XML ed è destinato a fornire metadati per servizi online aggregati, cioè i portali: non è inteso per essere usato come base per un sistema di gestione di raccolta o per sostenere l'attività di acquisizione o "prestito". Come formato di aggregazione, il modello può essere utilizzato per descrivere diversi tipi di oggetti, tra cui artistici, architeturali, di storia culturale, tecnologica, e naturale. I metadati possono provenire da diversi sistemi di gestione di collezioni utilizzando modelli di metadati in diverse lingue. A questo scopo, LIDO combina

---

<sup>26</sup> <http://www.lido-schema.org/>



## CAPITOLO 4 - SCHEMI METADATA

caratteristiche CDWA Lite-based, lo schema German *museumdat*<sup>27</sup>, e il sistema è allineato con lo standard di gestione delle collezioni britannico SPECTRUM. Il sistema è compatibile anche con lo standard CIDOC CRM. Di seguito, viene presentata la struttura di LIDO sulla base della sua documentazione standard (11).

Un record LIDO può includere 14 gruppi di elementi descrittivi e amministrativi di metadati, tre dei quali sono obbligatori (cfr. Tabella 5). L'elemento Subject Set è usato per descrivere l'oggetto utilizzando le categorie della Tabella 6. Da un punto di vista semantico, un elemento particolarmente versatile è Event Set, che descrive l'oggetto in termini di istanze di evento che vengono descritte utilizzando le categorie di elementi evento della Tabella 7. Gli eventi si verificano nel tempo, in luoghi diversi, e sono associati ad actors e altre risorse. In questo modo è possibile esprimere relazioni semantiche indirette dettagliate tra gli elementi raccolti e altre risorse nello spazio e tempo. I vantaggi della modellazione event centric vengono discussi a fine capitolo.

<b>Tipi Metadata</b>	<b>Elementi</b>
Object Classifications	Oggetto/Work type (obbligatorio) Classification
Identification	Title/Name (obbligatorio) Inscriptions Repository/Location State/Edition Object Description Measurements
Events	Event Set
Relations	Subject Set Related Works
Administrative	Rights Record (obbligatorio) Resource

Tabella 5: Elementi metadata LIDO

---

<sup>27</sup> <http://www.museumdat.org/>

## 4.5 HARVESTING SCHEMAS: LIDO

<b>Category</b>	<b>Significato</b>
Extent Subject	Quando ci sono più subjects, termine che indica la parte dell'object/work alla quale i subject fanno riferimento.
Subject Concept	Fornisce riferimenti a concetti legati al tema dell'object/work.
Subject Actor	Persona, gruppo o istituzione rappresentata da un object/work, o di quello che si tratta, forniti come elementi di visualizzazione e di indice.
Subject Date	Specifica di tempo rappresentante di o da un object/work, o di quello che si tratta, fornita come elemento di visualizzazione e di indice.
Subject Place	Luogo raffigurato in o da un object/work, o di quello che si tratta, fornito come elemento di visualizzazione e di indice.
Subject Event	Evento raffigurato in o da un object/work, o di quello che si tratta, fornito come elemento di visualizzazione e di indice.
Subject Object	Oggetto, ad esempio, un edificio o un'opera d'arte rappresentate da un object/work, o di quello che si tratta, fornito come elemento di visualizzazione e di indice.

Tabella 6: Elementi Subject Set LIDO

<b>Category</b>	<b>Significato</b>
Event ID	Event identifier.
Event Type	The nature of the event associated with an object/work.
Role in Event	The role played within this event by the described entity.
Event Name	An appellation for the event, e.g., a title, identifying phrase, or name given to it.
Event Actor	Wrapper for display and index elements for an actor with role information (participating or being present in the event).
Culture	Cultural context.
Event Date	Date specification of the event.
Period Name	A period in which the event happened.
Event Place	Place specification of the event.
Event Method	The method by which the event is carried out.
Materials/Technique	Indicates the substances or materials used within the event (e.g., the creation

## CAPITOLO 4 - SCHEMI METADATA

	of an object/work), as well as any implements, production or manufacturing techniques, processes, or methods incorporated.
Thing Present	References another object that was present at this same event.
Event Related	Display and index elements for the event related to the event being recorded.
Event Description	Wrapper for a description of the event, including description identifier, descriptive note of the event, and its sources.

Tabella 7: Elementi Event Description LIDO

### 4.6. PROTOCOLLI DI HARVESTING E DI RICERCA

Nel modello di creazione di contenuto distribuito, le memory organizations 1) forniscono API nel servizio centralizzato per la ricerca di contenuti (federated search) o 2) lasciano che sia il portale a fare harvesting di metadati pubblicati ad un certo indirizzo Web. Questa sezione presenta un approccio per la ricerca di linked data: endpoint SPARQL.

#### 4.6.1. ENDPOINT SPARQL PER LINKED DATA

Il metodo standard per fare federated queries a repository distribuite linked data è quello di utilizzare endpoint SPARQL, sulla base del protocollo SPARQL per RDF (SPROT<sup>28</sup>), un linguaggio di query e protocollo per RDF. Un endpoint SPARQL consente alle macchine ed a utenti finali di eseguire query SPARQL ad un repository RDF in modo semplice tramite HTTP. I risultati di questa query vengono poi restituiti in uno o più formati machine-processable, tipicamente XML, RDF o JSON.

---

<sup>28</sup> <https://www.w3.org/TR/rdf-sparql-protocol/>

## 4.6 PROTOCOLLI DI HARVESTING E DI RICERCA

Per esempio, la seguente query SPARQL può essere utilizzata per ricercare un libro ed il suo autore, dove vengono restituiti tutti quelli con specifiche triple nella clausola where:

```
PREFIX dc: <http://purl.org/dc/elements/1.1/> SELECT ?book ?author
WHERE {
  ?book dc:creator ?creator.
  ?creator :name ?author.
  ?book dc:time 2012. }
```

In questo caso, vengono restituiti gli URI di libri con i nomi degli autori come literal. Nel modello di metadati sottostante, ogni risorsa libro ha una proprietà *dc:creator* con un autore come valore, e ogni risorsa autore ha una proprietà *:name* con un literal indicante il suo nome.

La formulazione delle query e la presentazione human-readable dei risultati sono tipicamente creati da un software agent. Tuttavia, gli endpoint SPARQL vengono utilizzati anche non da macchine per ispezionare dataset e filtrare subset di dati. In questo caso, viene descritto come è possibile utilizzare facilmente un browser Web e un indirizzo URL in una query SPARQL.

Un endpoint SPARQL può essere utilizzato non solo per la ricerca di risultati tramite federate search, ma anche per fare harvesting di linked data. Tuttavia, utilizzando l'endpoint per harvesting di grandi dataset tramite query, rallenta il servizio per gli utenti. Un consiglio è quello di scaricare i file su cui effettuare poi harvesting.

## CAPITOLO 5

### PROGETTO FASE 1: IL CASO DI STUDIO

In questo capitolo vengono descritti nei minimi dettagli il caso di studio del progetto svolto nell'Azienda 3D Informatica<sup>1</sup>, vengono definiti gli obiettivi e i goal di questo progetto e viene fatto uno studio del dominio e dello stato attuale, nello specifico quello italiano. L'Azienda in qui è stato commissionato il lavoro realizza progetti di comunicazione e di integrazione tecnologica della propria piattaforma documentale; progetta, realizza e fornisce tecnologie ed applicazioni software per:

- la gestione, il trattamento e la valorizzazione dei Beni Culturali;
- la dematerializzazione della c.d. Amministrazione Digitale e la conservazione a norma (in primis per i Poli archivistici di conservazione);
- servizi ASP (Application Service Provider) / SaaS (Software as a Service) e ne fornisce la relativa gestione e manutenzione in una logica BPM (Business Process Management).

Per quanto riguarda questo progetto, l'Azienda fornisce strumenti ad enti catalogatori appunto per la catalogazione di qualsiasi bene culturale (foto, reperto archeologico, monete, ecc.). Attualmente, l'azienda propone come tecnologia base di salvataggio di tutte le informazioni relative ad un bene culturale l'XML. Una volta salvato, i dati sul bene

---

<sup>1</sup> <http://www.3di.it/>

## 5.1 OBIETTIVI DEL PROGETTO

culturale, proprio per il formato scelto, non permette nessuna delle innovazioni e rinnovamenti elencati precedentemente. Senza entrare troppo nello specifico di questo applicativo, è chiaro che l'utilizzo di una tecnologia più moderna rispetto a XML, ad esempio RDF/OWL, permetterebbe di arricchire ulteriormente le informazioni sui beni culturali salvate, come spiegato nei capitoli precedenti.

### 5.1. OBIETTIVI DEL PROGETTO

Un obiettivo di questo progetto non esplicitato, però derivabile, è quello di mantenere standardizzazione e consistenza pari o uguale a quella già presente grazie alle tecnologie utilizzate attualmente. Inoltre, l'utilizzo delle best practise del settore, integrato alle ontologie che saranno prima valutate e poi importate per l'utilizzo, permetterà di definire un'ontologia corretta e che segua una linea coerente con le altre sviluppate e pubblicate.

### 5.2. GOAL DEL PROGETTO

Il goal di questo progetto è quindi quello di creare, dopo un'attenta analisi delle ontologie standardizzate già pubblicate sul Web, un'ontologia custom (preferibilmente in RDF/OWL) che sia dipendente da ontologie standardizzate pubblicate ma che abbia un suo scopo, date le limitazioni e i requisiti di compilazione dettati dagli organi istituzionali italiani.

### 5.3. STUDIO SUL DOMINIO

Come accennato precedentemente, un bene culturale deve seguire ed essere compilato secondo delle regole fornite dall'Istituto Centrale per il Catalogo e la Documentazione (ICCD)<sup>2</sup> italiano.

---

<sup>2</sup> <http://www.iccd.beniculturali.it/>

## CAPITOLO 5 - PROGETTO FASE 1: IL CASO DI STUDIO

*"La Repubblica promuove lo sviluppo della cultura e della ricerca scientifica e tecnica. Tutela il paesaggio e il patrimonio storico e artistico della Nazione"*

Questo ente, seguendo l'Articolo 9 della Costituzione, lo interpreta e realizza attraverso l'emissione di norme tecnico-amministrative, in particolare per quanto riguarda l'organizzazione sistematica di conoscenze e informazioni sul patrimonio culturale d'Italia, quindi il conseguimento degli obiettivi di conservazione culturale e di sviluppo.

L'Istituto Centrale per il Catalogo e la Documentazione coordina la ricerca per la definizione degli standard di catalogazione per le diverse tipologie di beni culturali che afferiscono agli ambiti di tutela del MiBACT (archeologico; architettonico-paesaggistico; storico-artistico ed etnoantropologico).

Gli standard catalografici sono costituiti dalle normative, da specifici strumenti terminologici e da un insieme di indirizzi di metodo, definiti per attuare la catalogazione secondo criteri omogenei e condivisi a livello nazionale, in modo funzionale alla gestione informatizzata. Il catalogo del bene culturale viene definito come bene culturale, al pari dei beni che in esso sono schedati.

### **5.4. SCHEDA RA**

Questo istituto fornisce delle direttive sulle informazioni necessarie alla catalogazione di un qualsiasi bene culturale. Nel caso di questo progetto e della tesi, la scheda presa in considerazione è quella relativa ai Reperti Archeologici (RA)<sup>3</sup>. A questo link<sup>4</sup> è possibile scaricare, in formato xls, un file in cui sono definite tutte le informazioni che un ente catalogatore

---

<sup>3</sup><http://www.iccd.beniculturali.it/index.php?it/473/standard-catalografici/Standard/4>

<sup>4</sup> <http://www.iccd.beniculturali.it/getFile.php?id=4530>

## 5.4 SCHEDA RA

dovrà compilare riguardanti al reperto archeologico che sta archiviando. Viene presentata un'immagine di un campo della scheda RA per chiarire di cosa si sta parlando.

Acronimo				Definizione	LUN.	RIP.	OBB.	VOC.	VIS.
CD				CODICI			*		
	TSK			Tipo Scheda	4		*	C	1
	LIR			Livello ricerca	5		*	C	1
	NCT			CODICE UNIVOCO			*		
		NCTR		Codice regione	2		*	C	1
		NCTN		Numero catalogo generale	8		*	numero assegnato da ICCD	1
		NCTS		Suffisso numero catalogo generale	2				1
	ESC			Ente schedatore	25		*	codice assegnato da ICCD	1
	ECP			Ente competente	25		*	codice assegnato da ICCD	1
	EPR			Ente proponente	25				1

Figura 9: Campo Codici Scheda RA

È importante distinguere alcuni paragrafi da altri. Nello specifico, sono presenti dei paragrafi obbligatori e non obbligatori indicati con un asterisco. In più, esistono paragrafi in comune a tutte le schede che sono obbligatori. In questo caso, questi paragrafi compongono un livello speciale che viene denominato “Livello Inventariale”. Il livello inventariale è il livello minimo obbligatorio di compilazione che deve essere eseguito appena il bene viene catalogato.

I paragrafi che compongono il livello inventariale sono:

- **paragrafo CD:** Codici;
- **paragrafo OG:** Oggetto;
- **paragrafo LC:** Localizzazione Geografico-Amministrativa;
- **paragrafo DT:** Cronologia;
- **paragrafo MT:** Dati Tecnici;



## CAPITOLO 5 - PROGETTO FASE 1: IL CASO DI STUDIO

- **paragrafo TU:** Condizione Giuridica e Vincoli;
- **paragrafo DO:** Fonti e Documenti di Riferimento;
- **paragrafo AD:** Accesso ai Dati;
- **paragrafo CM:** Compilazione.

A loro volta, i paragrafi del livello inventariale (ma anche quelli non del livello inventariale), sono divisi in campi. In questo modo, il paragrafo CD, può essere scomposto nei suoi campi che forniscono informazioni aggiuntive singole. A loro volta, i campi possono essere divisi in sottocampi che descriveremo successivamente.

### 5.4.1. IL PARAGRAFO CD

Il paragrafo CD registra i dati che, nell'ambito del processo della catalogazione, identificano univocamente a livello nazionale il bene associato alla scheda che lo descrive.

Le definizioni e descrizioni dei sotto-campi del campo CD sono:

- **campo \*TSK:** (Tipo di Scheda), indica la sigla che contraddistingue il modello di scheda previsto per il bene catalogato. L'indicazione del tipo di scheda è necessaria per le procedure di trasferimento delle informazioni catalografiche fra banche-dati e per l'immediata individuazione della tipologia del bene. La compilazione del campo è obbligatoria;

Esempi:

- RA [Reperti Archeologici].
- **campo \*LIR:** (Livello ricerca), indica la sigla che individua il livello di indagine effettuato nel processo catalografico. La compilazione del campo è obbligatoria;

Esempi:

- I [Inventario];
  - P [Precatalogo];
  - C [Catalogo].
- **campo \*NCT<sup>5</sup>:** (Codice univoco), il codice univoco identifica a livello nazionale il bene e la scheda che lo descrive. Si genera dalla concatenazione dei valori registrati nei due sottocampi *Codice Regione* (NCTR) e *Numero catalogo generale* (NCTN), che viene assegnato dall'ICCD, e dall'eventuale *Suffisso numero catalogo generale* (NCTS). La compilazione del campo è obbligatoria;
  - **sottocampo \*NCTR:** (Codice regione), indica il numero di codice ISTAT che individua la Regione in cui ha sede l'Ente competente sul bene catalogato. La compilazione del sottocampo è obbligatoria;

Esempio:

- 12 [Lazio].
- **sottocampo \*NCTN:** (Numero catalogo generale), indica il numero, composto da otto cifre, assegnato dall'ICCD a ciascun bene catalogato secondo l'ordine progressivo relativo ad una determinata Regione. La compilazione del sottocampo è obbligatoria;

Esempi:

- 00005438;

---

<sup>5</sup> In riferimento ad uno stesso bene culturale possono essere redatte nel corso del tempo più schede, per attività di aggiornamento o di approfondimento del livello di catalogazione: **il codice univoco** (cfr. campo NCT) **resta sempre lo stesso in quanto si tratta del medesimo bene.**

## CAPITOLO 5 - PROGETTO FASE 1: IL CASO DI STUDIO

- 000034567.
- **Campo \*ESC:** (Ente Schedatore), indica il codice che identifica l'Ente responsabile della redazione della scheda di catalogo. *Tale codice viene assegnato dall'ICCD* ed annotato nel registro generale dei soggetti accreditati per effettuare campagne di catalogazione. La compilazione del campo è obbligatoria;

Esempi

- R08: Regione Emilia-Romagna;
- RM0175: Biblioteca della Camera dei Deputati.
- **campo \*ECP:** (Ente Competente), indica il codice che identifica l'Ente sotto la cui competenza ricade la tutela o la delega alla tutela del bene catalogato. *Tale codice viene assegnato dall'ICCD* ed annotato nel registro generale dei soggetti accreditati per effettuare campagne di catalogazione. La compilazione del campo è obbligatoria.

Esempi:

- S25: Soprintendenza per i Beni Culturali Archeologici della Lombardia;
- S45: Soprintendenza per i Beni Archeologici Dell'Etruria Meridionale.

Un esempio di compilazione del campo CD in formato XML relativo alla scheda RA è il seguente:

## 5.4 SCHEDA RA

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<scheda_RA>
  <!-- CD
  ID univoco del bene
  -->
  <CD>
    <TSK>RA</TSK> <!-- Tipo di scheda -->
    <LIR>P</LIR> <!-- Livello di ricerca -->
    <NCT>
      <NCTR>05</NCTR> <!-- Codice di regione -->
    </NCT>
    <ESC>R05</ESC> <!-- Ente schedatore -->
    <ECP>S32</ECP> <!-- Ente competente -->
  </CD>
```

Figura 10: Esempio di scheda RA e paragrafo CD in XML

Nella fase successiva, si effettuerà uno studio delle ontologie esistenti, pubblicate e standardizzate, che riescano quantomeno a mappare al meglio paragrafi, campi e sottocampi mostrati precedentemente.

## CAPITOLO 6

### PROGETTO FASE 2: MAPPING DEL LIVELLO INVENTARIALE

In questo capitolo, viene effettuata una ricerca sulle ontologie esistenti, pubblicate e standardizzate, che permettono di esprimere le informazioni di un'entità, in questo caso un bene culturale, in modo completo e corretto, una volta che la sua scheda è stata compilata e archiviata. Viene quindi fatta una valutazione ed una scelta tra le ontologie recuperate, fornendo motivi e giustificazioni, per quanto riguarda la completezza della scheda RA.

Successivamente verrà proposto un mapping iniziale su Excel, entità ICCD/entità ontologica, sul livello inventariale. L'Azienda ha voluto assegnare a ciascuno una scheda, nel mio caso quella dei Reperti Archeologici (RA), e visto che il livello inventariale è comune a tutte le schede, è stato deciso che il mapping del suddetto livello fosse suddiviso al numero dei tesisti. Perciò, verrà mostrato come sono stati mappati i paragrafi **CD**, **OG** e **LC**.

Il capitolo si conclude con la definizione dell'ontologia su Protégé, in cui vengono presentate le motivazioni delle scelte prese, vengono descritte nello specifico le entità mappate tramite lo strumento e vengono proposte schermate ed esempi per chiarezza.

In questo paragrafo, viene esposta la principale e più importante parte di questa fase: vale a dire la ricerca di ontologie su cui poter mappare le schede ICCD. L'obiettivo del mapping deve essere quello di, oltre a

## 6.1 MODELLI STANDARD PER LE ONTOLOGIE

mantenere la standardizzazione che già è presente, riuscire ad arricchire ulteriormente i dati archiviati su di un bene culturale, acquisendo una correttezza e visione più ampia rispetto ad una locale e chiusa.

Proprio per questo motivo, le ontologie che andremo a descrivere e valutare nei paragrafi successivi, per essere importate in quella custom del progetto, dovranno essere quantomeno standardizzate e pubblicate da enti affidabili.

### 6.1. MODELLI STANDARD PER LE ONTOLOGIE

Nel Semantic Web e Linked Data i modelli standard per rappresentare ontologie sono RDF Schema, Simple Knowledge Organization System SKOS, e Web Ontology Language OWL. Di seguito, viene presentata una breve panoramica di questi standard.

### 6.2. RDF SCHEMA

RDFS introduce la modellazione object oriented in RDF. L'idea è quella di descrivere un dominio del discorso in termini di classi, istanze individuali appartenenti a classi, e le proprietà che descrivono classi e istanze. Una classe, come *Painting*, rappresenta un set di istanze (ad esempio, Mona Lisa) che condividono le proprietà della classe. Queste proprietà sono specificate quando viene creata un'istanza, vale a dire, viene istanziata la classe. Ad esempio, se *Painting* ha proprietà *dc:title* e *dc:creator*, allora la stringa "Mona Lisa" può essere settata come titolo e un URI che referencia Leonardo Da Vinci per il creatore quando si istanzia l'oggetto Mona Lisa. La relazione istanza-classe è rappresentata da un arco proprietà *rdf:type* dall'istanza (URI) alla sua classe (URI).

Le classi sono organizzate in gerarchie di sottoclasse utilizzando la proprietà *rdfs:subClassOf*. Un'istanza viene considerata automaticamente facente parte non solo alla classe a cui punta l'arco *rdf:type*, ma anche alla sua superclasse, ereditando la sua definizione.

## CAPITOLO 6 - PROGETTO FASE 2: MAPPING DEL LIVELLO INVENTARIALE

Inoltre, le proprietà vengono organizzate in una gerarchia basata sulla proprietà *rdfs:subPropertyOf*. Le proprietà in RDFS sono risorse e possono quindi avere delle loro proprietà.

RDFS introduce in RDF l'idea di vincoli di proprietà: 1) Un constraint range descrive che il valore di una proprietà (ad esempio, *dc:creator*) deve sempre essere un'istanza di una classe specifica (ad esempio, *Person*). 2) Un domain constraint descrive nello stesso modo che la property subject deve essere un'istanza di una determinata classe, per esempio, che la proprietà *dc:creator* può essere utilizzata solo per istanze della classe *Work*.

Un'ontologia RDFS consiste quindi di 1) una gerarchia di classi, 2) una gerarchia di proprietà, e 3) un insieme di vincoli di proprietà. Un dataset RDF può essere creato istanziando le classi con valori specifici di proprietà.

### 6.3. SIMPLE KNOWLEDGE ORGANIZATION SYSTEM SKOS

L'obiettivo principale dello standard SKOS è quello di fornire un formato ontologico leggero in RDF per la rappresentazione di vocabolari, come legacy thesauri e classificazioni all'utilizzo. Da un punto di vista filosofico, il focus di modellazione è piuttosto incentrato sulla descrizione dei termini e delle strutture dei thesauri per presentare una concettualizzazione del mondo reale, anche se i thesauri vengono utilizzati per rappresentare concetti del mondo reale.

Il concetto chiave in SKOS è la classe *skos:Concept*. Un vocabolario SKOS viene costruito come rete semantica che collega le istanze di questa classe che rappresenta i termini di un thesaurus o una classificazione, o collezione di concetti (*skos:Collection*, *skos:OrderedCollection*). Il vocabolario include proprietà predefinite per equivalenza terminologica (ad esempio, *skos:prefLabel* e *skos:altLabel*

## 6.3 SIMPLE KNOWLEDGE ORGANIZATION SYSTEM SKOS

per preferred e alternative concept labeling), relazioni per rappresentare gerarchie thesaurus, come ad esempio *skos:narrower* e *skos:broader*, e relazione per concetti associativi, come *skos:related*. Ogni vocabolario SKOS è un'istanza della classe *skos:ConceptScheme*.

La specifica SKOS include un set di condizioni di integrità che possono essere utilizzate per validare schemi di concetti, e un set di regole logiche che possono essere utilizzate per arricchire vocabolari di relazioni tramite reasoning. Il modello è compatibile con RDF(S), standard OWL per rappresentare ontologie, e può essere esteso utilizzando i loro costrutti.

SKOS è “semplice” ma abbastanza versatile per catturare gran parte della semantica dei thesauri esistenti. È ampiamente utilizzato in Linked Data, concentrandosi sull'utilizzo di ontologie light-weight i dati RDF basati su risorse dati esistenti.

### 6.3.1. WEB ONTOLOGY LANGUAGE OWL

OWL è stato creato per affrontare le limitazioni semantiche di RDFS che seguono:

- I vincoli di proprietà non possono essere specificati class-wise in RDFS, ad esempio, il range di *dc:creator* per i libri (ad esempio, *Author*) dovrebbe essere diverso da quello di una sinfonia (ad esempio, *Composer*);
- La cardinalità delle proprietà non può essere specificata, per esempio, che una *Person* ha due genitori;
- Le proprietà base semantiche di relazioni sono spesso necessarie nella modellazione di ontologie e reasoning, ma mancano in RDFS. Per esempio, ci sono le proprietà inverse (ad esempio, moglie-di vs. marito-di), proprietà transitive (ad esempio, antenati di una persona sono antenati dei suoi figli), relazione funzionali (ad esempio, una



## CAPITOLO 6 - PROGETTO FASE 2: MAPPING DEL LIVELLO INVENTARIALE

persona ha un padre e una madre univoci), e proprietà funzionali inverse (ad esempio, il numero del documento di identità determina univocamente il suo titolare);

- Non è possibile definire classi con operazioni basic set in RDFS. Per esempio, non è possibile dire che una classe *Person* è unione tra *Man* e *Woman*.

OWL estende RDFS con nuovi costrutti per la modellare tali situazione ontologiche. Aumentando la potenza espressiva, possono essere create definizioni più accurate e più reasoning performed, ma al prezzo di efficienza computazionale. Per bilanciare questo trade-off tra espressività ed efficienza, sono state standardizzate tre versioni di OWL nel 2014<sup>1</sup>: OWL Lite, OWL DL, e OWL Full, in ordine crescente di potenza espressiva. Nel 2009, la specifica OWL successiva, chiamata OWL 2<sup>2</sup>, è stata standardizzata con tre profili linguistici per i diversi casi d'uso:

- OWL 2 EL per ontologie con un gran numero di classi/proprietà;
- OWL 2 QL per ontologie con un gran numero di dati istanza;
- OWL 2 RL per applicazioni che necessitano un reasoning scalabile e una decente potenza espressiva allo stesso tempo.

OWL2 impiega anche una “sintassi Manchester” alternativa per specificare ontologie in modo sintetico, l’annotazione human friendly. OWL2 è inoltre compatibile con la specifica originale OWL.

Una sfida di OWL dal punto di vista pratico è la complessità concettuale. Costrutti ontologici più espressivi sono più difficili da padroneggiare dagli utenti e non possono sempre essere creati

---

<sup>1</sup> <https://www.w3.org/2001/sw/wiki/OWL>

<sup>2</sup> <http://www.w3.org/TR/owl2-overview/>

## 6.3 SIMPLE KNOWLEDGE ORGANIZATION SYSTEM SKOS

automaticamente da data sources esistenti. Un'altra difficoltà in molti scenari è che molti concetti reali complessi sono difficili da modellare in termini di formulazioni logiche precise. Ad esempio, i concetti possono essere vaghi per natura, o la nostra conoscenza può essere incerta o incompleta, portando a problemi di non monotonia nel ragionamento. Di conseguenza, la modellazione ontologica complessa non è di solito utilizzata in Linked Data.

### 6.3.2. ONTOLOGIE LUOGO

Le Ontologie geo-spaziali definiscono classi e individui per rappresentare le regioni geografiche, le loro proprietà e relazioni reciproche (12) (13). L'interoperabilità in termini di luoghi geografici può essere promossa attraverso la condivisione delle risorse relative ai luoghi in diverse collezioni e domini applicativi (14) (15). Esistono molti dataset e repository disponibili per luoghi contemporanei forniti dalle organizzazioni e consorzi nazionali.

In molti casi, questi dati sono disponibili con licenze aperte. Un'ontologia particolarmente importante che si riferisce a luoghi in Linked Data è GeoNames<sup>3</sup>. Appartiene alla LOD cloud ed è stato collegato con diversi dataset, come DBpedia. Il dizionario geografico di GeoNames copre tutti i paesi e contiene più di otto milioni di descrizioni di luogo. I dati di GeoNames vengono utilizzati (harvested) da decine di altri database ed è presente anche un'interfaccia interattiva Web 2.0 per l'aggiunta di luoghi dalla comunità di sviluppatori. Il dataset è disponibile per il download gratuito.

Trattandosi di contenuto storico geografico si aggiunge una dimensione temporale e la nozione di cambiamento di sistemi informativi territoriali. Ad esempio, un riferimento a "Germania" o "the U.S." può riferirsi a diverse regioni, a seconda del periodo di riferimento. Ad

---

<sup>3</sup> <http://www.geonames.org/>

## CAPITOLO 6 - PROGETTO FASE 2: MAPPING DEL LIVELLO INVENTARIALE

esempio, la Germania del 1943, 1968, e quella attuale si estendono su aree molto diverse, e la questione di cosa fosse la Germania nel 18° e 19° secolo può essere vaga.

Per una rappresentazione più accurata e machine interpretable dei luoghi storici, è necessaria una nozione spazio-temporale del periodo e luogo necessario. Correlare queste regioni o luoghi ontologicamente con le altre è utile in information retrieval, perché l'utente finale potrebbe non utilizzare gli stessi nomi dei luoghi in query di ricerca che vengono utilizzate per le annotazioni, ma legate solamente al nome del luogo. Più in generale, sono necessarie relazioni ontologiche, topologiche e altre tra luoghi storici al fine di collegare i contenuti semanticamente correlati tra loro nelle applicazioni.

### 6.3.3. ONTOLOGIE EVENTO

Un'ontologia evento è utile per i seguenti casi d'uso (16) (17).

- **Uno schema metadata.** L'ontologia definisce un modello di metadata che possono essere utilizzati per rappresentare gli eventi storici in domini BC;
- **Un dominio ontologico/dizionario geografico.** I singoli eventi, conformi allo schema di metadata, possono essere utilizzati come un dizionario geografico per indicizzare contenuto storico culturale, come una fotografia scattata a un particolare evento;
- **Una data repository.** L'ontologia con il suo contenuto può essere utilizzata come data source di storia propria, ed essere collegato con dataset correlati. Possono essere poi collegati ulteriori dati e descrizioni riguardanti gli eventi.

## 6.4 MAPPING DEL LIVELLO INVENTARIALE

Sono stati proposti diversi modelli semantici per rappresentare eventi e le loro relazioni, come *Ontology Event*<sup>4</sup>, *LODE ontology*<sup>5</sup>, e *SEM* (18). Il *CIDOC CRM* è basato su eventi, associando al suo livello più alto oggetti fisici, oggetti concettuali, attori, luoghi e intervalli di tempo.

Esistono anche ricerche sullo sviluppo di ontologie storiche e markup languages, la visualizzazione di eventi su linee temporali e mappe. Il progetto *Papyrus* è stato sviluppato per produrre ontologie il cui scopo è quello di colmare il divario tra collezioni BC e i loro attributi storici in archivi, basato su *CIDOC CRM* (19).

### 6.4. MAPPING DEL LIVELLO INVENTARIALE

Riprendendo quanto detto nel capitolo in cui veniva spiegato il livello inventariale, i paragrafi da mappare per questo paragrafo sono i campi:

- **CD:** Codici;
- **OG:** Oggetto;
- **LC:** Localizzazione Geografico-Amministrativa.

Dopo un'accurata analisi delle ontologie standardizzate e pubblicate esistenti fatte nei paragrafi precedenti, è stato scelto di importare come prima ontologia di dominio *CIDOC-CRM*.

È stata inoltre necessaria un'attenta analisi di tutta la documentazione di questa ontologia. Nello specifico, viste le molteplici versioni e gli aggiornamenti dell'ontologia, la versione in uso è la *CIDOC-CRM 6.2.1*. Questa versione comprende 94 entità (classi) e 168 proprietà che collegano le entità definite.

---

<sup>4</sup> <http://motools.sourceforge.net/event/event.html>

<sup>5</sup> <http://linkedevents.org/ontology/>

## CAPITOLO 6 - PROGETTO FASE 2: MAPPING DEL LIVELLO INVENTARIALE

Ogni nome di entità in CIDOC-CRM è definito con la lettera *E* seguito da un numero progressivo. Per esempio, come detto precedentemente, l'entità che descrive tutti gli oggetti fisici appositamente creati da un'attività umana, è definita come *E22 Man-Made Object*, dove *E* sta per *Entity*.

In questa ontologia CIDOC-CRM, un bene culturale, e nello specifico un reperto archeologico, viene mappato esattamente con l'entità *E22 Man-Made Object*. Da questo punto in avanti, tutti paragrafi che verranno mappati, i campi e i sottocampi, e le proprietà faranno riferimento ad un'entità *E22 Man-Made Object*.

Per quanto riguarda invece i nomi delle proprietà, analogamente vengono definiti dalla lettera *P* seguita da un numero progressivo. Per esempio, la proprietà che descrive la denominazione o identificazione di un qualsiasi oggetto reale al mondo tramite un nome o un qualsiasi altro identificatore, è definita come *P1 is identified by (identifies)*, dove *P* sta per *Property*.

### 6.4.1. IL PARAGRAFO CD

Il primo paragrafo ad essere mappato è stato il paragrafo CD, che descrive e identifica in modo univoco tutti gli identificatori del bene culturale associato. L'entità che riesce a descrivere in maniera più corretta e significativa dell'ontologia CIDOC-CRM un codice identificativo è la *E42 Identifier*. Viene riportato di seguito lo scope di questa classe:

*“Questa classe comprende stringhe o codici assegnati a istanze di E1 CRM Entity che portano all'identificazione univoca e permanente all'interno di un contesto di una o più organizzazioni. Questi codici sono spesso conosciuti come numeri inventariali, codice di registrazione, ecc. e sono tipicamente composti da sequenze alfanumeriche”.*

## 6.4 MAPPING DEL LIVELLO INVENTARIALE

In CIDOC-CRM, questa entità *E42 Identifier* è collegata al reperto archeologico, descritto dall'entità *E22 Man-Made Object*, con la proprietà *P46 has preferred identifier (is preferred identifier of)*. Di seguito viene riportato lo scope di questa proprietà:

*“Questa proprietà salva l'identificatore E42 Identifier che viene usato per identificare un'istanza di E1 CRM Entity. È possibile assegnare più di un identificatore per un'oggetto nel tempo. L'uso di questa proprietà richiederà quindi un meccanismo esterno per validità temporale alla rispettiva istanza CRM”*. In questo caso, l'istanza di *E1 CRM Entity* è padre della gerarchia di classi che porta al reperto archeologico *E22 Man-Made Object*.

Il paragrafo CD, come spiegato precedentemente, è costituito da altri identificatori, e non altro. In questo caso quindi, la totalità del paragrafo CD viene mappata nella nuova ontologia custom tramite l'entità CIDOC-CRM *E42 Identifier*.

### 6.4.2. IL PARAGRAFO OG

Il secondo paragrafo ad essere mappato è il paragrafo OG, che descrive e contiene le informazioni necessarie per l'individuazione del reperto che si sta catalogando. Nell'ontologia CIDOC-CRM, l'entità che riesce a descrivere in maniera più corretta e significativa questo concetto è la *E55 Type*. Viene riportato di seguito lo scope di questa classe:

*“Questa classe comprende concetti denotati da termini proveniente da thesauri e vocabolari controllati utilizzati per caratterizzare e classificare istanze di classi CRM”*.

In CIDOC-CRM, questa entità *E55 Type* è collegata al reperto archeologico, descritto dall'entità *E22 Man-Made Object*, con la proprietà *P2 has type*. Di seguito viene riportato lo scope di questa proprietà:

## CAPITOLO 6 - PROGETTO FASE 2: MAPPING DEL LIVELLO INVENTARIALE

*“Questa proprietà permette di sotto-tipizzare entità CRM, come una forma di specializzazione, attraverso l’uso di una gerarchia di termini, o thesauri”.*

Anche in questo caso come in quello precedente, la totalità del paragrafo OG è rappresentata da entità che possono essere mappate tutte tramite la classe CIDOC-CRM *E55 Type*.

### 6.4.3. IL PARAGRAFO LC

Il terzo ed ultimo paragrafo di questo primo mapping a livello inventariale ad essere mappato è il paragrafo LC, che registra i dati relativi alla localizzazione geografico-amministrativa del reperto archeologico nel momento in cui viene compilata la scheda. L’entità che riesce a descrivere in maniera più corretta e significativa dell’ontologia CIDOC-CRM un codice identificativo è la ***E53 Place***. Viene riportato di seguito lo scope di questa classe:

*“Questa classe comprende estensioni nello spazio, in particolare sulla superficie terrestre, nel puro senso fisico: indipendentemente da fenomeni temporali o altro. Le istanze di E53 Place sono tipicamente determinate dalla posizione di oggetti “immobili” come edifici, città, montagne, fiumi”.*

In CIDOC-CRM, questa entità *E53 Place* è collegata al reperto archeologico, descritto dall’entità *E22 Man-Made Object*, con la proprietà ***P55 has current location***. Di seguito viene riportato lo scope di questa proprietà:

*“Questa proprietà salva la posizione di un E19 Physical Object durante il tempo di validità del record o database contenente la dichiarazione che utilizza questa proprietà”.*

## 6.5 TABELLA DEL MAPPING

In questo caso, l'entità CIDOC-CRM *E19 Physical Object* è padre della gerarchia di classi che porta al reperto archeologico identificato come *E22 Man-Made Object*.

Anche in questo caso come in quello precedente, la totalità del paragrafo OG è rappresentata da entità che possono essere mappate tutte tramite la classe CIDOC-CRM *E55 Type*.

### 6.5. TABELLA DEL MAPPING

Di seguito, viene presentata una tabella che rappresenta un primo mapping del livello inventariale. Vengono descritte unicamente le scelte fatte che riguardano il match dei paragrafi ICCD con le entità dell'ontologia scelta CIDOC-CRM.

In questa fase non viene detto, né la tecnologia, né le scelte implementative che verranno studiate e descritte successivamente.

I paragrafi vengono indicati con un colore di sfondo verde, i campi sono indicati con un colore di sfondo grigio, e l'obbligatorietà di un campo viene indicata con un asterisco.

CD – Codici *		
TSK *	Tipo scheda	E42 Identifier
LIR *	Livello ricerca	E42 Identifier
NCT – Codice Univoco *		
NCTR *	Codice regione	E42 Identifier
NCTN *	Numero catalogo generale	E42 Identifier
ESC *	Ente schedatore	E42 Identifier
ECP *	Ente proponente	E42 Identifier



## CAPITOLO 6 - PROGETTO FASE 2: MAPPING DEL LIVELLO INVENTARIALE

OG – Descrizione di Oggetto e Soggetto *		
OGT – Oggetto *		
OGTD *	Definizione	E55 Type
LC – Localizzazione Geografico-Amministrativa *		
PVC – Posizione corrente *		
PVCS *	Stato	E53 Place
PVCR *	Regione	E53 Place
PVCP *	Provincia	E53 Place
PVCC *	Comune	E53 Place

Tabella 8: Mapping del Livello Inventariale

### 6.6. IMPLEMENTAZIONE DELL'ONTOLOGIA

Il passo successivo sarà quello di implementare il mapping utilizzando un tool apposito, Protégé.

Protégé è un editor di ontologie open source gratuito, un sistema di knowledge management, e fornisce un'interfaccia grafica per definire ontologie. Comprende anche classificatori deduttivi per la convalidazione dei modelli, per fare in modo che siano coerenti, è capace di dedurre nuove informazioni sulla base dell'analisi dell'ontologia. Come Eclipse, Protégé è un framework per il quale vari altri progetti suggeriscono plugins da poter installare. La versione che è stata scelta di utilizzare è l'ultima, vale a dire la 5.1.0.

È stato deciso di mantenere un certo grado di distacco dalle ontologie importate. Per questo motivo, dove è necessario l'utilizzo di entità da ontologie importate, vengono definite delle sottoclassi per esprimere un concetto nella nuova ontologia custom. Così facendo, vengono mantenute

## 6.6 IMPLEMENTAZIONE DELL'ONTOLOGIA

sia le proprietà che le caratteristiche di quella entità senza bisogno di “toccare” e modificare l'ontologia importata.

### 6.6.1. OBJECT PROPERTY VS. DATA PROPERTY

Prima di continuare e descrivere il lavoro fatto con Protégé, è necessario descrivere e spiegare due concetti: le *Object Property* e le *Datatype Property*. Questi due elementi descrivono quale tipo di *valore* una tripla con proprietà debba avere.

Con Object Property si correlano individui con altri individui.

Invece, con Datatype Property si correlano individui a dei valori literal (ad esempio, stringhe, numeri, datatype, ecc.).

Un esempio di utilizzo potrebbe essere qualcosa del tipo una proprietà *hasAge* tipicamente sarà una Datatype Property, visto che l'età è un numero, invece la proprietà *hasMother* sarà una Object Property, perché una madre è un altro individuo.

Come nell'esempio precedente, anche in questo progetto è stato necessario determinare quali triple sarebbero dovuto diventare Datatype Property e Object Property una volta implementate in Protégé, modificando leggermente il mapping fatto precedentemente. Di seguito, vengono indicate in [Tabella 9](#) le Object Property e le Datatype Property relative alla [Tabella 8](#).

CD – Codici *	
TSK *	^^xsd:string - Datatype Property
LIR *	^^xsd:string - Datatype Property
NCT – Codice Univoco *	
NCTR *	Object Property

## CAPITOLO 6 - PROGETTO FASE 2: MAPPING DEL LIVELLO INVENTARIALE

NCTN *	Object Property
ESC *	Object Property
ECP *	Object Property
OG – Descrizione di Oggetto e Soggetto *	
OGT – Oggetto *	
OGTD *	Object Property
LC – Localizzazione Geografico-Amministrativa *	
PVC – Posizione corrente *	
PVCS *	Object Property
PVCR *	Object Property
PVCP *	Object Property
PVCC *	Object Property

Tabella 9: Mapping Object Property vs. Datatype Property

### 6.7. ARCHEOLOGICAL FIND ONTOLOGY (AFO)

Archeological Find Ontology (AFO) è l'ontologia creata di questa prima parte seguendo lo standard dettato dall'ICCD e importando solamente l'ontologia pubblicata CIDOC-CRM.

L'ontologia è stata creata, come detto precedentemente, utilizzando l'applicativo Protégé, e per tutta la durata dell'implementazione è stato eseguito un reasoner denominato Hermitt (versione 1.3.8.413). In questo modo, è possibile avere un feedback istantaneo, o quasi, sulla corretta modellazione e implementazione dell'ontologia.

## 6.7 ARCHEOLOGICAL FIND ONTOLOGY (AFO)

Evitando di mostrare troppe schermate di Protégé, viene data una definizione delle classi in formato tabellare, con esempi per capirne meglio il funzionamento. La sintassi utilizzata è di Turtle.

CD – Codici *		<p>AFO#Iccd_CD  rdfs:subClassOf <b>crm:E42_Identifier</b>  (Class)</p> <p>La classe CD indica tutti i dati che identificano un Bene Culturale.</p> <pre>[rdf:type owl:Restriction owl:onProperty : identifies; owl:maxQualifiedCardinality "1"^^xsd:nonNegativeInteger; owl:onClass :Bene_Culturale ];</pre>
TSK *	<p><b>Tipo scheda</b></p> <p>Termine da un vocabolario controllato per fornire informazioni riguardo il tipo della voce.</p>	<p>AFO#Iccd_TSK  rdfs:subClassOf :Iccd_CD;  <b>(Datatype Property)</b></p> <p>La classe TSK, sottoclasse di crm:E42 Identifier, è una classe che indica la sigla che contraddistingue il modello di scheda previsto.</p> <p>Esempio di utilizzo:  :’RA’  a Iccd_TSK, crm:E42_Identifier</p> <p>Proprietà:  • AFO#hasTSK  rdfs:domain : Iccd_TSK;  rdfs:range : xsd:string;</p> <p>Esempio di utilizzo:  afentry FAO#hasTSK ‘RA’;</p>
LIR *	<p><b>Livello ricerca</b></p> <p>Termine da dizionario controllato che fornisce informazioni sull’accuratezza della catalogazione.</p>	<p>AFO#Iccd_LIR  rdfs:subClassOf :Iccd_CD;  <b>(Datatype Property)</b></p> <p>Esempio di utilizzo:  :’I’  a Iccd_LIR, crm:E42_Identifier</p> <p>Proprietà:  • AFO#hasLir</p>

**CAPITOLO 6 - PROGETTO FASE 2: MAPPING DEL LIVELLO  
INVENTARIALE**

		<p>rdfs:domain : Iccd_LIR; rdfs:range : xsd:string;</p> <p>Esempio di utilizzo: afentry FAO#hasLir 'I';</p>
<b>NCT – Codice Univoco *</b>		<p>AFO#Iccd_NCT rdfs:subClassOf :Iccd_CD; (Class)</p>
<b>NCTR *</b>	<p><b>Codice regione</b></p> <p>Codice che identifica la regione dove l'istituto competente risiede: questo codice, combinato con altri, definisce un identificatore nazione univoco per l'entry corrente.</p>	<p>AFO#Iccd_NCTR rdfs:subClassOf :Iccd_NCT; (Class)</p> <p>Esempio di utilizzo: :nctr001 a Iccd_NCTR, crm:E42_Identifier</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• AFO#has_identifier rdfs:subPropertyOf crm:P48; rdfs:domain : Bene_Culturale; rdfs:range : Iccd_NCTR;</li> <li>• AFO:is_identified_by rdfs:subPropertyOf crm:P48i; rdfs:domain : Iccd_NCTR; rdfs:range : Bene_Culturale.</li> </ul> <p>Esempio di utilizzo: afentry FAO#has_identifier nctr001;</p>
<b>NCTN *</b>	<p><b>Numero catalogo generale</b></p> <p>Un codice che identifica l'entry, fornito dall'istituzione competente, ICCD, che può essere combinato con altri codici per definire ulteriormente un identificatore unico per l'entry corrente.</p>	<p>AFO#Iccd_NCTN rdfs:subClassOf :Iccd_NCT; (Class)</p> <p>Esempio di utilizzo: :nctn001 a Iccd_NCTN, crm:E42_Identifier</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• AFO#has_identifier rdfs:subPropertyOf crm:P48; rdfs:domain : Bene_Culturale; rdfs:range : Iccd_NCTN;</li> <li>• AFO:is_identified_by rdfs:subPropertyOf crm:P48i; rdfs:domain : Iccd_NCTN; rdfs:range : Bene_Culturale.</li> </ul> <p>Esempio di utilizzo: afentry FAO#has_identifier nctn001;</p>

## 6.7 ARCHEOLOGICAL FIND ONTOLOGY (AFO)

ESC *	<p>Ente schedatore</p> <p>Codice che identifica l'istituzione catalogatrice del reperto archeologico, fornito da istituzione competente, ICCD.</p>	<p>AFO#Iccd_ESC rdfs:subClassOf :Iccd_CD; (Class)</p> <p>Esempio di utilizzo: :esc001 a Iccd_ESC, crm:E42_Identifier</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• AFO#has_identifier rdfs:subPropertyOf crm:P48; rdfs:domain : Bene_Culturale; rdfs:range : Iccd_ESC;</li> <li>• AFO:is_identified_by rdfs:subPropertyOf crm:P48i; rdfs:domain : Iccd_ESC; rdfs:range : Bene_Culturale.</li> </ul> <p>Esempio di utilizzo: afentry FAO#has_identifier esc001;</p>
ECP *	<p>Ente competente</p> <p>Codice che identifica l'istituto competente che supervisiona il processo di catalogazione.</p>	<p>AFO#Iccd_ECP rdfs:subClassOf :Iccd_CD; (Class)</p> <p>Esempio di utilizzo: :ecp001 a Iccd_ECP, crm:E42_Identifier</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• AFO#has_identifier rdfs:subPropertyOf crm:P48; rdfs:domain : Bene_Culturale; rdfs:range : Iccd_ECP;</li> <li>• AFO:is_identified_by rdfs:subPropertyOf crm:P48i; rdfs:domain : Iccd_ECP; rdfs:range : Bene_Culturale.</li> </ul> <p>Esempio di utilizzo: afentry FAO#has_identifier ecp001;</p>
OG – Descrizione di Oggetto e Soggetto *		<p>AFO#Iccd_OG rdfs:subClassOf <b>crm:E55_Type</b> (Class)</p> <p>La classe OG contiene le informazioni necessarie per l'individuazione del reperto archeologico.</p>
OGTD *	Definizione	<p>AFO#Iccd_OGTD rdfs:subClassOf :Iccd_OG;</p>

**CAPITOLO 6 - PROGETTO FASE 2: MAPPING DEL LIVELLO INVENTARIALE**

	Termini che identifica la tipologia del reperto archeologico. Potrebbe riferirsi a un thesauro locale e/o ad uno stabilito, come il <i>Art and architecture Thesaurus</i> .	<p>(Class)</p> <p>Esempio di utilizzo: :ogtd001 a Iccd_OGTD, crm:E55_Type</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• AFO#has_type rdfs:subPropertyOf crm:P2; rdfs:domain : Bene_Culturale; rdfs:range : Iccd_OGTD;</li> <li>• AFO:is_type_of rdfs:subPropertyOf crm:P2i; rdfs:domain : Iccd_OGTD; rdfs:range : Bene_Culturale.</li> </ul> <p>Esempio di utilizzo: afentry FAO#has_type ogtd001;</p>
	<b>LC – Localizzazione Geografico-Amministrativa *</b>	<p>AFO#Iccd_LC rdfs:subClassOf <b>crm:E53_Place</b> (Class)</p>
	<b>PVC – Posizione corrente *</b>	<p>AFO#Iccd_PVC rdfs:subClassOf AFO#Iccd_LC (Class)</p> <p>Questa classe rappresenta la posizione corrente dove il bene è collegato.</p>
<b>PVCS *</b>	<p><b>Stato</b></p> <p>Il nome dello stato dove il reperto archeologico è collocato.</p>	<p>AFO#Iccd_PVCS rdfs:subClassOf :Iccd_PVC; <b>(Object Property)</b></p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• AFO#has_current_location rdfs:subPropertyOf crm:P55; rdfs:domain : Bene_Culturale; rdfs:range : Iccd_LC;</li> <li>• AFO#fall_within rdfs:subPropertyOf crm:P89; rdfs:domain : Iccd_LC; rdfs:range : Iccd_LC.</li> </ul> <p>Esempio di utilizzo: afentry a crm:E22_Man-Made_Object; AFO#has_current_location :gallery001. gallery001 a Iccd_PVC; AFO#falls_within gallery-palace. gallery-palace a Iccd_PVC; AFO#falls_within :alfonsine.</p>

## 6.7 ARCHEOLOGICAL FIND ONTOLOGY (AFO)

		<p>alfonsine a Iccd_PVC;  AFO#falls_within :ravenna.  ravenna a Iccd_PVC;  AFO#falls_within :emilia-romagna.  emilia-romagna a Iccd_PVC;  AFO#falls_within :italy;  italy a Iccd_PVC;  AFO#has_type :country.  : country a crm:E55_Type.</p>
PVCR *	<p><b>Regione</b></p> <p>Il nome della regione dove il reperto archeologico è collocato.</p>	<p>AFO#Iccd_PVCR  rdfs:subClassOf :Iccd_PVC;  <b>(Object Property)</b></p> <p>Esempio di utilizzo:  :emilia-romagna a AFO#Iccd_PVC;  crm:P2_has_type AFO#Iccd_PVCR;  AFO#falls_within :italy.  :AFO#Iccd_PVCR a crm:E55_Type.</p>
PVCP *	<p><b>Provincia</b></p> <p>Il nome della provincia dove il reperto archeologico è collocato.</p>	<p>AFO#Iccd_PVCP  rdfs:subClassOf :Iccd_PVC;  <b>(Object Property)</b></p> <p>Esempio di utilizzo:  :ravenna a AFO#Iccd_PVC;  crm:P2_has_type AFO#Iccd_PVCP;  AFO#falls_within :emilia-romagna.  :AFO#Iccd_PVCP a crm:E55_Type.</p>
PVCC *	<p><b>Comune</b></p> <p>Il nome del comune dove il reperto archeologico è collocato.</p>	<p>AFO#Iccd_PVCC  rdfs:subClassOf :Iccd_PVC;  <b>(Object Property)</b></p> <p>Esempio di utilizzo:  :alfonsine a AFO#Iccd_PVC;  crm:P2_has_type AFO#Iccd_PVCC;  AFO#falls_within :ravenna.  :AFO#Iccd_PVCC a crm:E55_Type.</p>

Tabella 10: Livello Inventariale in mapping RDF



### 6.7.1. OSSERVAZIONI

Ci sono alcune osservazioni da fare sul mapping creato precedentemente, sia riguardanti il mapping singolo, che riguardanti la fase successiva di integrazione del lavoro con gli altri due colleghi e di completamento del mapping della scheda RA.

1. Il mapping è relativo: cioè significa che, anche se si è cercati di rimanere il più generali possibile, alcuni aspetti sono stati mappati riguardanti un reperto archeologico. Nel caso di una integrazione, verranno rivisti alcuni aspetti. Questo non vuole dire che il mapping fatto sia sbagliato, solo che dovranno essere effettuati alcuni accorgimenti e allineamenti tra i diversi lavori;
2. Come è possibile constatare, è stato descritto un mapping utilizzando solamente come ontologia importata CIDOC-CRM. Infatti, visto che questo modello è basato sugli eventi, è meno complicato spiegare, sia nel mapping attuale che in quello che verrà, la logica adottando un approccio CRM. Quindi è stato deciso di non importare altre ontologie in questa fase, comunque non necessarie, visto il potere espressivo di CRM;
3. Il paragrafo Localizzazione Geografico-Amministrativa (LC), come è possibile vedere, è stato mappato tramite due Object Property. La prima Object Property *has current location* identifica la posizione del reperto archeologico in questione che viene catalogato. Questa posizione viene mappata con il campo AFO#Iccd\_PVC.

La seconda proprietà, *falls within*, con dominio e range entrambi *E53 Place*, identifica che un luogo ricade all'interno dell'estensione di un altro luogo. Viene utilizzata solamente per indirizzare un contenimento spaziale, e non implica alcuna relazione o fenomeno che occupano questi luoghi;

## 6.7 ARCHEOLOGICAL FIND ONTOLOGY (AFO)

4. Nonostante i consistenti aggiornamenti, la scheda RA è ancora alla versione 3.00. ICCD sta completando il proprio modello e la versione 4.00 di standard alla catalogazione, rendendolo ancora più razionale. D'altra parte, anche CIDOC-CRM è in continua evoluzione. Quindi, i concetti comuni ed elementi cross-sezioni mappati precedentemente, renderanno il mapping di tutti gli altri schemi ICCD più facili e veloci da eseguire anche una volta aggiornati.

## **CAPITOLO 7**

### **PROGETTO FASE 3: MAPPING DELLA SCHEDA RA**

In questo capitolo viene presentata una rapida descrizione dell'integrazione tra i singoli lavori che hanno portato alla definizione di un'ontologia formale per la copertura del livello inventariale. Dopo questa breve parentesi, verrà presentato il mapping completo della scheda RA, in maniera simile a come esplicitato nel capitolo precedente per il mapping dei tre paragrafi.

#### **7.1. INTEGRAZIONE DEI MAPPING**

Prima di parlare della parte vera e propria, specifica alla scheda RA, del mapping, senza risultare troppo discorsivi, si descriverà il lavoro di integrazione. In questo paragrafo è stato scelto di presentare unicamente le variazioni che sono state necessarie per poter integrare le tre ontologie del livello inventariale create per questo progetto. Ognuno dei colleghi ha mappato tre paragrafi obbligatori della propria scheda (nel mio caso i paragrafi CD, OG ed LC). L'ontologia creata fornirà una base su cui completare il mapping per le schede del progetto (nel mio caso la scheda RA).

### 7.1.1. FRBR-ALIGNED BIBLIOGRAPHIC ONTOLOGY (FABIO)

Per aumentare la potenza espressiva e semantica del lavoro singolo fatto è necessario introdurre una nuova ontologia che dovrà essere integrata correttamente, e che porterà ad una variazione della modellazione dell'ontologia del livello inventariale vista nel capitolo precedente, vale a dire l'ontologia FRBR-aligned Bibliographic Ontology (FaBiO).

FaBiO è un'ontologia per la registrazione e la pubblicazione di documenti bibliografici sul Semantic Web e fa parte di SPAR, una suite di *Semantic Publishing and Referencing Ontologies*.

Il motivo per il quale si è scelto di importare questa ontologia è per la potenza espressiva che riesce a dare agli elementi che vengono mappati con essa. L'ontologia FaBiO definisce quattro tipi di Entità principali:

- **Work:** sottoclasse di FRBR work, ristretta ai lavori che sono pubblicati o che potrebbero essere pubblicati, e che contengono o sono riferiti a documenti bibliografici, o entità usate per definire bibliografie. Per esempio, un'entità mappabile come Work, potrebbe essere l'idea che ha avuto Dante Alighieri per la stesura della Divina Commedia;
- **Expression:** sottoclasse di FRBR expression, ristretta all'espressione di un fabio:Work. In questo caso, con espressione viene intesa, per esempio, una prima bozza, o appunti, o qualsiasi altro materiale che sia un'espressione dell'idea Work, se esiste;
- **Manifestation:** sottoclasse di FRBR manifestation, ristretta a manifestazioni di una fabio:Expression. fabio:Manifestation si applica specificatamente a manifestazione di espressioni in formato elettronico (digitale) ma anche fisico. Un esempio, sempre riferito a quello precedente, potrebbe essere il manoscritto originale della Divina Commedia;

## CAPITOLO 7 - PROGETTO FASE 3: MAPPING DELLA SCHEDA RA

- **Item:** sottoclasse di FRBR item, ristretto a esemplari di una fabio:Manifestation. In questo caso, un esempio di fabio:Manifestation potrebbero essere le varie versioni e copie della Divina Commedia negli anni.

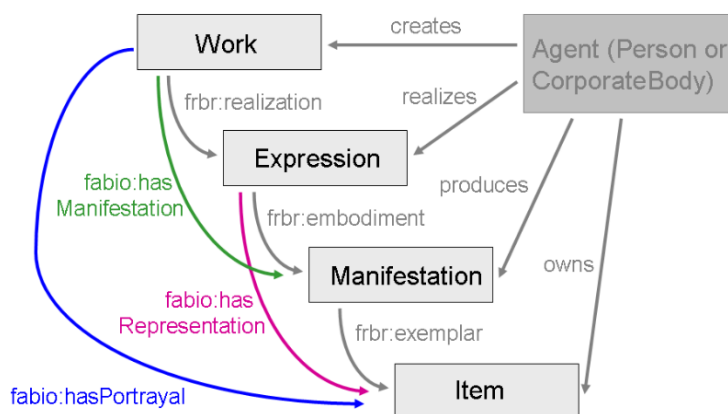


Figura 11: Rappresentazione di FaBiO - FRBR-aligned Bibliographic Ontology

Ognuno di questi elementi è collegato all'altro tramite proprietà, come mostrato nella [Figura 11](#). Non è obbligatorio l'utilizzo di ogni classe e proprietà, infatti, nella stesura del mapping della scheda RA verranno utilizzati solo alcuni dei costrutti di questa ontologia. Senza anticipare nulla, alcune entità già definite come elementi di CIDOC-CRM dovranno quindi essere ampliate e modellate seguendo la logica di questa ontologia dato che porta sia maggiore espressività che maggiore semantica. Nei capitoli successivi, relativi al mapping vero e proprio dell'ontologia della scheda RA, verranno mostrati anche questi accorgimenti fatti durante l'integrazione sulle entità già create.

### 7.2. STESURA DEL MAPPING COMPLETO

Per questa fase del progetto, verranno mappati i paragrafi rimanenti della scheda RA più significativi per un bene culturale, nello specifico un reperto archeologico. Perciò, i paragrafi scelti sono:

## 7.3 IL PARAGRAFO RE

- RE: Modalità di Reperimento;
- AU: Definizione Culturale;
- CO: Conservazione;
- GP: Georeferenziazione tramite punto;
- TU: Condizione Giuridica.

Come fatto nella fase precedente del livello inventariale, è stata eseguita un'analisi però relativa all'ontologia CIDOC-CRM per verificare la copertura anche di questi paragrafi. Come risultato, è stato possibile mappare i paragrafi in CIDOC-CRM, con alcune variazioni spiegate successivamente.

Ricordo che, nell'ontologia CIDOC-CRM, un bene culturale, e nello specifico un reperto archeologico, viene mappato con l'entità *E22 Man-Made Object*.

### 7.3. IL PARAGRAFO RE

Il primo paragrafo di questa fase ad essere mappato è stato il paragrafo RE, che raccoglie informazioni relative alle indagini che hanno consentito il reperimento del bene.

L'evento ritrovamento è molto importante per l'archeologia, e rappresenta una pietra angolare nella ricostruzione della storia dell'oggetto. Dal punto di vista CIDOC-CRM, il ritrovamento di un oggetto è un tipo di acquisizione (*E8 Acquisition*) che può verificarsi durante (*P117 occurs during*) un esame o scavo (*E7 Activity*) cambiando la proprietà dell'oggetto (*P24 changed ownership through*), che viene acquisita dall'ente che esegue la scoperta.

## CAPITOLO 7 - PROGETTO FASE 3: MAPPING DELLA SCHEDA RA

Quindi, seguendo la logica di CIDOC-CRM, un ritrovamento si riconduce ad un'acquisizione **E8 Acquisition**. Viene riportato di seguito lo scope di questa classe:

*“Questa classe comprende i trasferimenti di proprietà legale da uno o più istanze di E39 Actor ad una o più E39 Actor”.*

Questa acquisizione avviene durante un'**E7 Activity**. Viene riportato di seguito lo scope di questa classe:

*“Questa classe comprende le azioni intenzionalmente svolte da un E39 Actor che si traducono in cambiamenti di stato nei sistemi culturali, sociali o fisici documentati”.*

In CIDOC-CRM, questa entità **E7 Activity** è collegata al reperto archeologico, descritto dall'entità **E22 Man-Made Object**, con la proprietà **P24 changed ownership through**. Di seguito viene riportato lo scope di questa proprietà:

*“Questa proprietà identifica il E18 Physical Thing o Things coinvolte in un'E8 Acquisition”. In questo caso, è necessario un collegamento ulteriore per poter rappresentare l'evento ritrovamento, e per fare questo viene utilizzata la proprietà **P117 occurs during**. Di seguito viene riportato lo scope di questa proprietà:*

*“Questa proprietà consente che l'intero lasso di tempo di un evento temporale E2 Temporal Entity sia situato all'interno di un lasso temporale di un'altra entità che inizia prima e finisce dopo l'entità temporale compresa”.*

### 7.3.1. IL CAMPO RCG

Il primo campo da mappare del paragrafo RE è il campo *Ricognizioni* (RCG). Questo campo contiene informazioni relative alla ricognizione che ha consentito il reperimento del bene. È stato deciso di non definire

### 7.3 IL PARAGRAFO RE

questo campo, in quanto non porta informazioni neppure come nodo vuoto. Si passa quindi alla mappatura dei suoi sottocampi.

- **Sottocampo RCGA:** indica il nome del responsabile scientifico della ricognizione nella forma ‘cognome,nome’. Viene mappato con l’entità CIDOC-CRM *E39 Actor*. Di seguito viene fornito lo scope di questa entità: *“Questa classe comprende persone, individui o gruppi, che hanno il potenziale per eseguire azioni intenzionali per le quali qualcuno potrebbe essere responsabile”*.

La proprietà che collega un *E39 Actor* all’entità *E7 Activity* è la proprietà *PI4 carried out by*. Di seguito viene mostrato lo scope di questa proprietà: *“Questa proprietà descrive i partecipanti attivi di un E39 Actor in un E7 Activity”*;

- **Sottocampo RCGE:** indica il motivo che ha determinato la ricognizione. Viene mappato con l’entità CIDOC-CRM *E17 Type Assignment*. Di seguito viene fornito lo scope di questa entità: *“Questa classe comprende le azioni di classificazione di item di qualsiasi genere, i quali includono oggetti, campioni, azioni e concetti”*.

Esempio: archeologia preventiva, fotointerpretazione, ecc.

La proprietà che collega un *E17 Type Assignment* all’entità *E7 Activity* è la proprietà *PI7 was motivated by*. Di seguito viene mostrato lo scope di questa proprietà: *“Questa proprietà descrive uno o più item considerati la ragione che porta al completamento di una E7 Activity”*;

- **Sottocampo RCGM:** indica il metodo di ricognizione adottato. Viene mappato con l’entità CIDOC-CRM *E55 Type*. Di seguito viene fornito lo scope di questa entità: *“Questa classe comprende concetti*



## CAPITOLO 7 - PROGETTO FASE 3: MAPPING DELLA SCHEDA RA

*denotati da termini provenienti da thesauri e vocabolari controllati utilizzati per caratterizzare e classificare istanze di classi CRM”.*

Esempio: occasionale, per campioni, ecc.

La proprietà che collega un *E55 Type* all’entità *E7 Activity* è la proprietà ***P32 used general technique***. Di seguito viene mostrato lo scope di questa proprietà: *“Questa proprietà identifica la tecnica o metodo che è stato utilizzato in un’attività”*;

- **Sottocampo RCGD**: indica l’epoca in cui è stata effettuata la ricognizione archeologica. Viene mappato con l’entità CIDOC-CRM ***E52 Time-Span***. Di seguito viene fornito lo scope di questa entità: *“Questa classe comprende estensioni temporali astratte, nel senso della fisica Galileiana, avendo un inizio, una fine e una durata”*.

Esempio: 1998/11/06, primavera del 2012, ecc.

La proprietà che collega un *E52 Time-Span* all’entità *E7 Activity* è la proprietà ***P4 has time-span***. Di seguito viene mostrato lo scope di questa proprietà: *“Questa proprietà descrive il confinamento temporale di un’istanza di un E2 Temporal Entity”*. Nel mio caso, *E52 Time-Span* sottoclasse di *E2 Temporal Entity*.

### 7.3.2. IL CAMPO DSC

Il secondo campo del paragrafo *RE* è relativo ai *Dati di Scavo* (DSC). Questo campo contiene informazioni relative all’intervento di scavo archeologico che ha consentito il reperimento del bene. Come nel caso precedente del campo *RCG*, è stato deciso di non definire questo campo. Si passa quindi alla mappatura dei suoi sottocampi.

- **Sottocampo NCUN**: indica il codice che individua lo scavo in modo univoco a livello nazionale, e viene assegnato dall’ICCD. Viene mappato con l’entità CIDOC-CRM ***E42 Identifier***. Di seguito viene

### 7.3 IL PARAGRAFO RE

fornito lo scope di questa entità: *“Questa classe comprende stringhe o codici assegnati a istanze di E1 CRM Entity che portano all’identificazione univoca e permanente all’interno di un contesto di una o più organizzazioni”*.

La proprietà che collega un *E42 Identifier* all’entità *E7 Activity* è la proprietà ***PI is identified by***. Di seguito viene mostrato lo scope di questa proprietà: *“Questa proprietà descrive la denominazione o identificazione di qualsiasi item mondiale reale per nome o qualsiasi altro identificatore”*;

- **Sottocampo DSCI**: indica la serie numerica o alfanumerica che identifica il bene catalogato all’interno dell’inventario di scavo. Viene mappato con l’entità CIDOC-CRM ***E42 Identifier***.

Esempio: 87, VC45/ 187, ecc.

La proprietà che collega un *E42 Identifier* all’entità *E7 Activity* è la proprietà ***PI is identified by***. Lo scope non viene ripetuto perché scritto precedentemente;

- **Sottocampo SCAN**: indica la denominazione ufficiale dello scavo archeologico o, in alternativa, quella maggiormente attestata nella bibliografia o nella letteratura scientifica, oppure una denominazione convenzionale utile ad individuare l’intervento di scavo a cui si fa riferimento. Viene mappato con l’entità CIDOC-CRM ***E41 Appellation***. Di seguito viene fornito lo scope di questa entità: *“Questa classe comprende segni, che abbiano significato o no, o disposizioni di segni che seguono una specifica sintassi, che vengono utilizzati o potrebbero essere utilizzati per identificare una specifica istanza di una qualche classe o categoria all’interno di un certo contesto”*.

Esempio: Himera, tempio B; Calvatone, area della domus dei signini.

## CAPITOLO 7 - PROGETTO FASE 3: MAPPING DELLA SCHEDA RA

La proprietà che collega un *E41 Appellation* all'entità *E7 Activity* è la proprietà *P1 is identified by*. Lo scope non viene ripetuto perché scritto precedentemente;

- **Sottocampo DSCF:** indica la denominazione dell'Ente responsabile dello scavo. Viene mappato con l'entità CIDOC-CRM *E39 Actor*. Lo scope non viene ripetuto perché scritto precedentemente.

La proprietà che collega un *E39 Actor* all'entità *E7 Activity* è la proprietà *P14 carried out by*. Lo scope non viene ripetuto perché scritto precedentemente;

- **Sottocampo DSCA:** indica il nome del responsabile scientifico dello scavo nella forma 'cognome, nome'. Viene mappato con l'entità CIDOC-CRM *E39 Actor*. Lo scope non viene ripetuto perché scritto precedentemente.

La proprietà che collega un *E39 Actor* all'entità *E7 Activity* è la proprietà *P14 carried out by*. Lo scope non viene ripetuto perché scritto precedentemente;

- **Sottocampo DSCT:** indica il motivo che ha determinato l'intervento di scavo. Viene mappato con l'entità CIDOC-CRM *E17 Type Assignment*. Lo scope non viene ripetuto perché scritto precedentemente.

Esempio: US 122, USM SG71, ecc.

La proprietà che collega un *E17 Type Assignment* all'entità *E7 Activity* è la proprietà *P32 used general technique*. Lo scope non viene ripetuto perché scritto precedentemente;

- **Sottocampo DSCM:** indica il sinteticamente il metodo di scavo adottato. Viene mappato con l'entità CIDOC-CRM *E55 Type*. Lo scope non viene ripetuto perché scritto precedentemente.

### 7.3 IL PARAGRAFO RE

Esempio: open area, trincea di sondaggio, ecc.

La proprietà che collega un *E55 Type* all'entità *E7 Activity* è la proprietà *P32 used general technique*. Lo scope non viene ripetuto perché scritto precedentemente;

- **Sottocampo DSCD:** indica l'epoca in cui è stato effettuato lo scavo archeologico. Viene mappato con l'entità CIDOC-CRM *E52 Time-Span*. Lo scope non viene ripetuto perché scritto precedentemente.

Esempio: 1999/07/11-1999/10/01, 2001/12/16, ecc.

La proprietà che collega un *E52 Time-Span* all'entità *E7 Activity* è la proprietà *P4 has time-span*. Lo scope non viene ripetuto perché scritto precedentemente;

- **Sottocampo DSCU:** indica l'identificativo numerico o alfanumerico dell'Unità Stratigrafica dalla quale proviene il bene catalogato, preceduta dalla sigla US. Viene mappato con l'entità CIDOC-CRM *E53 Place*. Viene riportato di seguito lo scope di questa classe: “*Questa classe comprende estensioni nello spazio, in particolare sulla superficie terrestre, nel puro senso fisico: indipendentemente da fenomeni temporali o altro*”.

Esempio: USR PT41, US CC1456, ecc.

La proprietà che collega un *E53 Place* all'entità *E7 Activity* è la proprietà *P7 took place at*. Lo scope di questa proprietà è: “*Questa proprietà descrive posizioni spaziali di un'istanza E4 Period*”. Anche in questo caso un *E53 Place*, seguendo la gerarchia, è sottoclasse di un *E4 Period*;

- **Sottocampo DSCS:** indica, qualora il ritrovamento del bene sia connesso ad una deposizione funeraria, l'identificativo numerico o

## CAPITOLO 7 - PROGETTO FASE 3: MAPPING DELLA SCHEDA RA

alfanumerico. Viene mappato con l'entità CIDOC-CRM *E53 Place*.  
Lo scope non viene ripetuto perché scritto precedentemente.

Esempio: Tomba 4, T115, ecc.

La proprietà che collega un *E53 Place* all'entità *E7 Activity* è la proprietà *P7 took place at*. Lo scope non viene ripetuto perché scritto precedentemente.

### 7.4. IL PARAGRAFO AU

In questo paragrafo vengono registrate le informazioni relative ai soggetti (persone singole o enti collettivi) intervenuti nella realizzazione del bene e/o all'ambito culturale di riferimento. La produzione è un processo molto complesso, che coinvolge vari oggetti, persone e luoghi. È di primaria importanza documentare nel modo corretto quando si tratta di reperti archeologici.

In un modo simile a quanto fatto per il ritrovamento (RE), viene definito un evento di produzione *E12 Production* in grado di mettere in relazione tra loro i vari luoghi e attori coinvolti. Di seguito viene descritto lo scope di questa entità:

*“Questa classe comprende le attività che sono progettate per, e riescono per, la creazione di uno o più item nuovi”.*

La produzione dell'oggetto archeologico è specificamente riferita dalla proprietà *P108 was produced by*. Di seguito lo scope per questa proprietà:

*“Questa proprietà identifica l'E24 Physical Man-Made thing che si verifica, ed esiste, come risultato di una E12 Production”.*

### 7.4.1. IL CAMPO AUT

Questo campo contiene i dati relativi all'autore (persona singola o ente collettivo) oppure, se l'autore non è noto, all'ambito a cui esso è riferibile. In questo caso, a differenza dei campi del paragrafo *RE*, è utile descriverli invece che lasciarli come nodi vuoti.

Questo campo è mappabile tramite l'entità CIDOC-CRM *E39 Actor*, di seguito viene fornito lo scope di questa entità:

*“Questa classe comprende persone, individui o gruppi, che hanno il potenziale per eseguire azioni intenzionali per le quali qualcuno potrebbe essere responsabile”.*

Similmente, la proprietà utilizzata per collegare un *E39 Actor* ad un *E12 Production* non è altro che la proprietà *P14 carried out by*, con scope:

*“Questa proprietà descrive i partecipanti attivi di un E39 Actor in un E7 Activity”.*

### 7.4.2. IL CAMPO ATB

L'ultimo campo che ho deciso di mappare relativo al paragrafo *AU* è quello relativo all'Ambito Culturale *ATB*. Questo campo descrive indicazioni relative all'ambito culturale a cui può essere riferito il bene catalogato. Anche qui, è utile mappare questo campo, perché porta informazioni aggiuntive già a questo livello.

Questo campo è stato mappato con l'entità CIDOC-CRM *E13 Attribute Assignment*, di seguito viene fornito lo scope di questa entità:

*“Questa classe comprende le azioni di fare asserzioni sulle proprietà di un oggetto o di una relazione tra due item o concetti”.*

## CAPITOLO 7 - PROGETTO FASE 3: MAPPING DELLA SCHEDA RA

La proprietà che collega un *E13 Attribute Assignment* ad una *E12 Production* è la ***P140 was attributed by***, di cui lo scope è:

*“Questa proprietà indica l’item al quale è assegnato un attributo o una relazione”.*

### 7.5. IL PARAGRAFO CO

In questo paragrafo vengono registrate le informazioni sullo stato di conservazione del bene, sulla base dell’osservazione autoptica. Non ci sono molte informazioni in questo paragrafo come nei precedenti, quindi passiamo subito alle entità CIDOC-CRM che mappano queste informazioni.

Il paragrafo *CO* viene mappato con l’entità CIDOC-CRM *E3 Condition State*, il quale scope relativo è:

*“Questa classe comprende gli stati degli oggetti caratterizzati da una certa condizione nel tempo. Un’istanza di questa classe descrive la condizione fisica prevalente di un qualsiasi oggetto o caratteristica durante un E52 Time-Span specifico”.*

Come solito, ricordiamo che l’entità Reperto Archeologico è identificata con l’entità CIDOC-CRM *E22 Man-Made Thing*, quindi, la proprietà che collega questa entità ad un *E3 Condition State* è la ***P44 has condition***. Di seguito viene presentato lo scope di questa proprietà:

*“Questa proprietà registra un E3 Condition State per un qualche E18 Physical Thing”.* Anche in questo caso, è possibile risalire da un *E22 Man-Made Thing* ad un *E18 Physical Thing*.

### 7.5.1. IL CAMPO STC

Questo campo contiene le informazioni sullo stato di conservazione del bene, nello specifico però è mappabile come nodo vuoto, in quanto l'informazione viene salvata sul suo sottocampo.

- **Sottocampo STCC:** indica lo stato di conservazione del bene catalogato, valutato in relazione al manufatto intero originale. Viene mappato con l'entità CIDOC-CRM *E55 Type*. Viene riportato di seguito lo scope di questa classe: *“Questa classe comprende concetti denotati da termini provenienti da thesauri e vocabolari controllati utilizzati per caratterizzare e classificare istanze di classi CRM”*.

Esempio: integro, mutilo, ecc.

La proprietà che collega un *E55 Type* all'entità *E3 Condition State* è la proprietà *P2 has type*. Lo scope di questa proprietà è: *“Questa proprietà permette di sotto-tipizzare entità CRM, come una forma di specializzazione, attraverso l'uso di una gerarchia di termini, o thesauri”*.

## 7.6. IL PARAGRAFO GP

Con la sigla *GP* si intende *Georeferenziazione Tramite Punto*, nello specifico in questo paragrafo vengono registrate informazioni relative alla georeferenziazione del bene, espresse mediante coordinate (geografiche o proiettate). Il paragrafo *GP* è molto simile al paragrafo *LC*, infatti anche nelle specifiche ICCD viene affermato che, qualora si sia fatto riferimento ad un contenitore fisico autorevole (campo *LDC*), si raccomanda, per la localizzazione fisica del bene, di non registrare nella scheda i dati di georeferenziazione, in quanto coincideranno con quella del contenitore fisico stesso.



## CAPITOLO 7 - PROGETTO FASE 3: MAPPING DELLA SCHEDA RA

In questo caso però, viene comunque dato un mapping di questo campo in quanto fornisce ulteriore chiarezza, e verrà comunque implementato nell'ontologia completa. Visto che l'informazione relativa ai punti non viene registrata a questo livello, si passerà direttamente a quelli successivi, però per raggiungere questa informazione è necessario effettuare un collegamento tra l'entità Reperto Archeologico e il luogo.

Nello specifico, verrà utilizzata la proprietà *P55 has current location* per collegare l'entità *E22 Man-Made Thing* e l'entità *E53 Place*. Quindi, il mapping che seguirà, considererà come punto di partenza un *E53 Place*.

### 7.6.1. IL CAMPO GPD

Questo campo registra le informazioni necessarie per la definizione del punto. Le coordinate spaziali definiscono i punti che identificano e georeferenziano il luogo in cui l'oggetto è tenuto o localizzato. Le coordinate spaziali si riferiscono anche a tutti gli altri tipi di posto descritti dall'ICCD.

Per questo motivo, le coordinate spaziali, descritte dai sottocampi *GDPX* e *GDPY* verranno collegate al luogo definito nella Tabella 8 come *PVC*.

- **Sottocampo GDPX:** indica la coordinata x (longitudine x), utilizzando il punto come separatore per i decimali. Viene mappato con l'entità CIDOC-CRM *E47 Spatial Coordinates*. Viene riportato di seguito lo scope di questa classe: “*Questa classe comprende informazioni testuali o numeriche per localizzare specifiche istanze di un E53 Place all'interno di schemi di identificazione spaziale*”.

Esempio: 12.49248, 654541.166, ecc.

La proprietà che collega un *E47 Spatial Coordinates* all'entità *AFO#Iccd\_PVC* è la proprietà *P87 is identified by*. Lo scope di

## 7.7 IL PARAGRAFO TU

questa proprietà è: “*Questa proprietà identifica un E53 Place tramite un E44 Place Appellation*”;

- **Sottocampo GDPY:** indica la coordinata y (latitudine Nord), utilizzando il punto come separatore per i decimali. La mappatura è analoga a quanto fatto per il mapping del sottocampo *GDPX*.

### 7.7. IL PARAGRAFO TU

Il quinto ed ultimo paragrafo da mappare per questo progetto è il Paragrafo *TU – Condizione Giuridica e Vincoli*. In questo paragrafo vengono registrate le informazioni inerenti l’acquisizione e la condizione giuridica del bene e i provvedimenti di tutela che lo riguardano. In caso di aggiornamento della scheda, vengono indicati qui gli eventi che hanno determinato mutamenti della titolarità del bene o della sua condizione materiale.

Anche in questo caso, non essendoci informazione in questo nodo, è stato deciso di tralasciarlo per mappare il campo figlio di *TU*.

#### 7.7.1. IL CAMPO ACQ

Questo campo registra indicazioni relative alla circostanza e/o titolo in base a cui il bene è pervenuto e si trova nelle attuali condizioni di proprietà o detenzione, le misure protettive che lo riguardano. Siccome le istituzioni potrebbero avere modi diversi per acquisire un Reperto Archeologico (esempio, dopo un ritrovamento o un acquisto, una donazione, uno scambio, ecc.), l’ICCD fornisce un vocabolario specifico per la definizione del tipo dell’acquisizione (*ACQT*).

L’evento acquisizione (*E8 Acquisition*), descritto nel Capitolo 7.3, viene considerato come cambio di proprietà del Reperto Archeologico attraverso un trasferimento del titolo dall’ex proprietario (*P22 transferred title to*) ad uno nuovo che lo riceve (*P23 transferred title*

## CAPITOLO 7 - PROGETTO FASE 3: MAPPING DELLA SCHEDA RA

from). Inoltre, il campo *ACQ* registra date e luoghi, e fornisce dettagli riguardanti gli attori (persone o gruppi) coinvolti nell'evento. Per quest'ultimo, la proprietà *P52 has current owner* viene utilizzata per definire l'ultimo ricevente, nel nostro caso l'istituzione che ha creato questo record (*E39 Actor*).

Dopo questa analisi, risulta ovvia la scelta dell'entità CIDOC-CRM *E10 Transfer to Custody* per identificare il campo *ACQ*. Di seguito viene riportato lo scope di questa classe:

*“Questa classe comprende trasferimenti di custodia fisica di oggetti tra istanze E39 Actor”.*

La proprietà invece che collega un *E22 Man-Made Object* ad un *E10 Transfer to Custody* è la proprietà *P30 custody transferred through*, il cui scope è:

*“Questa proprietà identifica uno o più item di E18 Physical Thing riguardanti un'attività E10 Transfer of Custody”.* Ora, si procederà al mapping dei sottocampi del campo *ACQ*.

- **Sottocampo ACQT:** indica le modalità secondo quali il bene è stato acquisito. Viene mappato con l'entità CIDOC-CRM *E55 Type*. Viene riportato di seguito lo scope di questa classe: *“Questa classe comprende concetti denotati da termini proveniente da thesauri e vocabolari controllati utilizzati per caratterizzare e classificare istanze di classi CRM”.*

Esempio: ipoteca, confisca, ecc.

La proprietà che collega un *E10 Transfer to Custody* all'entità *E55 Type* è la proprietà *P2 has type*. Lo scope di questa proprietà è: *“Questa proprietà permette di sotto-tipizzare entità CRM, come una*

*forma di specializzazione, attraverso l'uso di una gerarchia di termini, o thesauri”;*

- **Sottocampo ACQN:** indica il soggetto fonte dell'acquisizione, utilizzando il nome (quanto possibile nella forma 'cognome, nome'), la denominazione, oppure la forma aggettivata. Viene mappato con l'entità CIDOC-CRM **E82 Actor Appellation**. Viene riportato di seguito lo scope di questa classe: *“Questa classe comprende qualsiasi tipo di nome, numero, codice o simbolo caratterizzante utilizzato per identificare un E39 Actor”*.

Esempio: Bianchi, Mario; Banca popolare di Milano, ecc.

La proprietà che collega un *E39 Actor* all'entità *E82 Actor Appellation* è la **P131 is identified by**. Lo scope di questa proprietà è: *“Questa proprietà identifica il nome usato per identificare un E39 Actor”*.

È necessario però specificare che, per il collegamento da un'entità *E10 Transfer to Custody* ad un'entità *E39 Actor* bisogna utilizzare tre diverse proprietà, ognuna per tipo di proprietario:

- Ex proprietario: la proprietà è la **P23 transferred title from**, di cui lo scope: *“Questa proprietà identifica uno o più E39 Actor che cedono la proprietà legale come risultato di un E8 Acquisition”*;
- Proprietario corrente: la proprietà è la **P52 has current owner**, con scope: *“Questa proprietà identifica l'E21 Person, E74 Group o E40 Legal Body che fu proprietario di un'istanza di E18 Physical Thing”*;
- Nuovo proprietario: la proprietà è la **P22 transferred title to**, di cui lo scope: *“Questa proprietà identifica l'E39 Actor che*

## CAPITOLO 7 - PROGETTO FASE 3: MAPPING DELLA SCHEDA RA

*acquisisce la proprietà legale di un oggetto come risultato di un E8 Acquisition”.*

- **Sottocampo ACQD:** indica l’epoca in cui è avvenuta l’acquisizione del bene. Viene mappato con l’entità CIDOC-CRM **E52 Time-Span**. Viene riportato di seguito lo scope di questa classe: *“Questa classe comprende estensioni temporali astratte, nel senso della fisica Galileiana, avendo un inizio, una fine e una durata”.*

La proprietà che collega un *E10 Transfer to Custody* all’entità *E52 Time-Span* è la proprietà **P4 has time-span**. Lo scope di questa proprietà è: *“Questa proprietà descrive il confinamento temporale di un’istanza di un E2 Temporal Entity”;*

- **Sottocampo ACQL:** indica il luogo in cui è avvenuta l’acquisizione, nella forma ‘Regione/ sigla Provincia/ denominazione per esteso del Comune/ località o toponimo (le varie informazioni vanno divise da una barra ‘/’ seguita da uno spazio). Viene mappato con l’entità CIDOC-CRM **E53 Place**. Viene riportato di seguito lo scope di questa classe: *“Questa classe comprende estensioni nello spazio, in particolare sulla superficie terrestre, nel puro senso fisico: indipendentemente da fenomeni temporali o altro”.*

La proprietà che collega un *E10 Transfer to Custody* all’entità *E53 Place* è la proprietà **P7 took place at**. Lo scope di questa proprietà è: *“Questa proprietà descrive posizioni spaziali di un’istanza E4 Period”.*



## CAPITOLO 8

### PROGETTO FASE 4: CICDO

Il nome che è stato scelto per questa ontologia è *Central Institute for Cataloguing and Documentation Ontology CICDO*.

In questo capitolo viene presentata una tabella che descrive gli oggetti ICCD mappati su quelli CIDOC-CRM, per fornire una visione più grafica del lavoro fino a questo punto. Successivamente, come fatto nel capitolo precedente relativo all'ontologia del livello inventariale, verrà presentata un'ulteriore tabella relativa all'implementazione dell'ontologia sull'applicativo Protégé.

#### 8.1. TABELLA DEL MAPPING

In questo paragrafo viene presentata la tabella che rappresenta il primo mapping dell'ontologia CICDO. Vengono descritte unicamente le scelte fatte che riguardano il match dei paragrafi ICCD con le entità dell'ontologia scelta CIDOC-CRM. In questa fase non viene detto, né la tecnologia, né le scelte implementative che verranno studiate e descritte successivamente. I paragrafi vengono indicati con un colore di sfondo verde, i campi sono indicati con un colore di sfondo grigio, e l'obbligatorietà di un campo viene indicata con un asterisco.

RE – Modalità di reperimento	E7 Activity
RCG – Riconsezioni	

## 8.1 TABELLA DEL MAPPING

RCGA	Responsabile scientifico	E39 Actor
RCGE	Motivo	E17 Type Assignment
RCGM	Metodo	E55 Type
RCGD	Data	E52 Time-Span
DSC – Dati di scavo		
NCUN	Codice univoco	E42 Identifier
DSCI	Numero inventario di scavo	E42 Identifier
SCAN	Denominazione dello scavo	E41 Appellation
DSCF	Ente responsabile	E39 Actor
DSCA	Responsabile scientifico	E39 Actor
DSCT	Motivo	E17 Type Assignment
DSCM	Metodo	E55 Type
DSCD	Data	E52 Time-Span
DSCU	Unità stratigrafica	E53 Place
DSCS	Numero tomba	E53 Place
AU – Definizione Culturale		E12 Production
AUT	Autore	E39 Actor
ATB	Ambito culturale	E13 Attribute Assignment
CO – Conservazione		E3 Condition State
STCC	Stato di conservazione	E55 Type
GDPX – Georeferenziazione Tramite Punto		E47 Spatial Coordinates



## CAPITOLO 8 - PROGETTO FASE 4: CICDO

GDPY – Georeferenziazione Tramite Punto		E47 Spatial Coordinates
TU – Condizione Giuridica *		
ACQ – Acquisizione		E10 Transfer to custody
ACQT	Tipo di acquisizione	E55 Type
ACQN	Nome	E82 Actor Appellation
ACQD	Data acquisizione	E52 Time-Span
ACQL	Luogo acquisizione	E53 Place

Tabella 11: Mapping dell'ontologia CICDO

Come risulta dalla Tabella 11, le proprietà che collegano le classi sono state omesse, per non complicarne ulteriormente il formato, in quanto verranno descritte nello specifico nel prossimo paragrafo riguardante l'implementazione.

### 8.2. IMPLEMENTAZIONE SU PROTÉGÉ

In questo paragrafo verrà presentata una tabella completa dell'implementazione eseguita in maniera incrementale, sotto controllo da parte del reasoner Hermitt, dell'ontologia CICDO. Come fatto per l'ontologia del livello inventariale AFO, viene data una definizione delle classi in formato tabellare, con esempi per capirne il funzionamento in sintassi Turtle.

RE – Modalità di reperimento	<p>CICDO#Iccd_RE  rdfs:subClassOf <b>crm:E7_Activity</b>  (Class)</p> <p>La classe RE indica tutti i dati relativi alle indagini per il reperimento del bene.</p>
------------------------------	---

## 8.2 IMPLEMENTAZIONE SU PROTÉGÉ

		<p>Esempio di utilizzo: :mdr001 a Iccd_RE, crm:E7_Activity</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#includes rdfs:domain : Iccd_RE; rdfs:range : E8_Acquisition;</li> <li>• CICDO#occurs_during rdfs:domain : E8_Acquisition; rdfs:range : Iccd_RE;</li> </ul> <p>Esempio di utilizzo: mdr001 CICDO#includes crm:E8_Acquisition;</p>
RCG	<b>Ricognizioni</b>	
	Informazioni relative alla ricognizione che ha portato al reperimento del bene.	
RCGA	<b>Responsabile scientifico</b>	<p>CICDO#Iccd_RCGA rdfs:subClassOf :E21_Person; (Class)</p> <p>Esempio di utilizzo: :'Bianchi, Mario' a Iccd_RCGA, crm:E39_Actor</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#carried_out_by rdfs:domain : Iccd_RE; rdfs:range : Iccd_RCGA;</li> <li>• CICDO#performed rdfs:domain : Iccd_RCGA; rdfs:range : Iccd_RE;</li> </ul> <p>Esempio di utilizzo: 'Bianchi, Mario' CICDO#performed mdr001;</p>
	Informazioni su chi ha operato la ricognizione, 'cognome, nome'.	
RCGE	<b>Motivo</b>	<p>CICDO#Iccd_RCGE rdfs:subClassOf:E17_Type_Assignment; (Class)</p> <p>Esempio di utilizzo: : 'archeologia preventiva' a Iccd_RCGE, crm:E17_Type_Assignment</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#used_general_technique rdfs:domain : Iccd_RE;</li> </ul>
	Informazioni sul motivo che ha determinato la ricognizione.	

## CAPITOLO 8 - PROGETTO FASE 4: CICDO

		<p>rdfs:range : Iccd_RCGE;</p> <p>Esempio di utilizzo: mdr001 CICDO#was_motivated_by 'archeologia preventiva';</p>
RCGM	<p><b>Metodo</b></p> <p>Informazioni sul metodo di ricognizione adottato.</p>	<p>CICDO#Iccd_RCGM rdfs:subClassOf :E55_Type; (Class)</p> <p>Esempio di utilizzo: : 'occasionale' a Iccd_RCGE, crm:E55_Type</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#used_general_technique rdfs:domain : Iccd_RE; rdfs:range : Iccd_RCGM;</li> <li>• CICDO#was_technique_of rdfs:domain : Iccd_RCGM; rdfs:range : Iccd_RE;</li> </ul> <p>Esempio di utilizzo: 'archeologia preventiva' CICDO#was_technique_of mdr001;</p>
RCGD	<p><b>Data</b></p> <p>Informazioni sull'epoca in cui è stata effettuata la ricognizione.</p>	<p>CICDO#Iccd_RCGD rdfs:subClassOf :E52_Time_Span; (Class)</p> <p>Esempio di utilizzo: : '1998/11/06' a Iccd_RCGD, crm:E52_Time_Span</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#has_time_span rdfs:domain : Iccd_RE; rdfs:range : Iccd_RCGD;</li> <li>• CICDO#is_time_span_of rdfs:domain : Iccd_RCGD; rdfs:range : Iccd_RE;</li> </ul> <p>Esempio di utilizzo: '1998/11/06' CICDO#is_time_span_of mdr001;</p>
DSC	<p><b>Dati di scavo</b></p> <p>Informazioni relative all'intervento di scavo che ha consentito il reperimento del bene.</p>	

## 8.2 IMPLEMENTAZIONE SU PROTÉGÉ

NCUN	<p>Codice univoco</p>	<p>CICDO#Iccd_NCUN  rdfs:subClassOf :E42_Identifier;  (Class)</p> <p>Esempio di utilizzo:  : ncun001  a Iccd_NCUN, crm:E42_Identifier</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#identifies  rdfs:domain : Iccd_NCUN;  rdfs:range : Iccd_RE;</li> <li>• CICDO#is_identified_by  rdfs:domain : Iccd_RE;  rdfs:range : Iccd_NCUN;</li> </ul> <p>Esempio di utilizzo:  ncun001 CICDO#identifies mdr001;</p>
DSCI	<p>Numero inventario di scavo</p> <p>Informazioni sulla serie numerica o alfanumerica che identifica il bene all'interno dell'inventario di scavo.</p>	<p>CICDO#Iccd_DSCI  rdfs:subClassOf :E42_Identifier;  (Class)</p> <p>Esempio di utilizzo:  : dsci001  a Iccd_DSCI, crm:E42_Identifier</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#identifies  rdfs:domain : Iccd_DSCI;  rdfs:range : Iccd_RE;</li> <li>• CICDO#is_identified_by  rdfs:domain : Iccd_RE;  rdfs:range : Iccd_DSCI;</li> </ul> <p>Esempio di utilizzo:  dsci001 CICDO#identifies mdr001;</p>
SCAN	<p>Denominazione dello scavo</p> <p>Informazioni sulla denominazione ufficiale dello scavo archeologico.</p>	<p>CICDO#Iccd_SCAN  rdfs:subClassOf :E41_Appellation;  (Class)</p> <p>Esempio di utilizzo:  : scan001  a Iccd_SCAN, crm:E41_Appellation</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#identifies</li> </ul>

## CAPITOLO 8 - PROGETTO FASE 4: CICDO

		<p>rdfs:domain : Iccd_SCAN;  rdfs:range : Iccd_RE;  • CICDO#is_identified_by  rdfs:domain : Iccd_RE;  rdfs:range : Iccd_SCAN;</p> <p>Esempio di utilizzo:  scan001 CICDO#identifies mdr001;</p>
DSCF	<p>Ente responsabile</p> <p>Informazioni sulla denominazione dell'Ente responsabile dello scavo.</p>	<p>CICDO#Iccd_DSCF  rdfs:subClassOf :E21_Person;  (Class)</p> <p>Esempio di utilizzo:  : dscf001  a Iccd_DSCF, crm:E21_Person</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#carried_out_by  rdfs:domain : Iccd_RE;  rdfs:range : Iccd_DSCF;</li> <li>• CICDO#performed  rdfs:domain : Iccd_DSCF;  rdfs:range : Iccd_RE;</li> </ul> <p>Esempio di utilizzo:  dscf001 CICDO#performed mdr001;</p>
DSCA	<p>Responsabile scientifico</p> <p>Informazioni su chi ha operato la ricognizione, 'cognome, nome'.</p>	<p>CICDO#Iccd_DSCA  rdfs:subClassOf :E21_Person;  (Class)</p> <p>Esempio di utilizzo:  :'Bedeschi, Giuseppe'  a Iccd_DSCA, crm:E39_Actor</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#carried_out_by  rdfs:domain : Iccd_RE;  rdfs:range : Iccd_DSCA;</li> <li>• CICDO#performed  rdfs:domain : Iccd_DSCA;  rdfs:range : Iccd_RE;</li> </ul> <p>Esempio di utilizzo:  :'Bedeschi, Giuseppe' CICDO#performed mdr001;</p>
DSCT	Motivo	<p>CICDO#Iccd_DSCT  rdfs:subClassOf:E17_Type_Assignment;</p>

## 8.2 IMPLEMENTAZIONE SU PROTÉGÉ

	<p>Informazioni sul motivo che ha determinato la ricognizione.</p>	<p>(Class)</p> <p>Esempio di utilizzo: : 'opere private' a Iccd_DSCT, crm:E17_Type_Assignment</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#was_motivated_by rdfs:domain : Iccd_RE; rdfs:range : Iccd_DSCT;</li> </ul> <p>Esempio di utilizzo: mdr001 CICDO#was_motivated_by 'opere private';</p>
DSCM	<p>Metodo</p> <p>Informazioni sul metodo di ricognizione adottato.</p>	<p>CICDO#Iccd_DSCM rdfs:subClassOf :E55_Type;</p> <p>(Class)</p> <p>Esempio di utilizzo: : 'open area' a Iccd_DSCM, crm:E55_Type</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#used_general_technique rdfs:domain : Iccd_RE; rdfs:range : Iccd_DSCM;</li> <li>• CICDO#was_technique_of rdfs:domain : Iccd_DSCM; rdfs:range : Iccd_RE;</li> </ul> <p>Esempio di utilizzo: 'open area' CICDO#was_technique_of mdr001;</p>
DSCD	<p>Data</p> <p>Informazioni sull'epoca in cui è stata effettuata la ricognizione.</p>	<p>CICDO#Iccd_DSCD rdfs:subClassOf :E52_Time_Span;</p> <p>(Class)</p> <p>Esempio di utilizzo: : '2001/12/16' a Iccd_DSCD, crm:E52_Time_Span</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#has_time_span rdfs:domain : Iccd_RE; rdfs:range : Iccd_DSCD;</li> <li>• CICDO#is_time_span_of rdfs:domain : Iccd_DSCD;</li> </ul>

## CAPITOLO 8 - PROGETTO FASE 4: CICDO

		<p>rdfs:range : Iccd_RE;</p> <p>Esempio di utilizzo: '2001/12/16' CICDO#is_time_span_of mdr001;</p>
DSCU	<p>Unita stratigrafica</p> <p>Informazioni sull'identificativo numerico o alfanumerico dell'Unità Stratigrafica dalla quale proviene il bene catalogato.</p>	<p>CICDO#Iccd_DSCU rdfs:subClassOf :CICDO#Geographic_Localization; (Class)</p> <p>La classe CICDO#Geographic_Localization è la classe dell'ex livello inventariale AFO#Iccd_LC.</p> <p>Esempio di utilizzo: : 'US 122' a Iccd_DSCU, CICDO#Geographic_Localization</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#took_place_at rdfs:domain : Iccd_RE; rdfs:range : Iccd_DSCU;</li> <li>• CICDO#witnessed rdfs:domain : Iccd_DSCU; rdfs:range : Iccd_RE;</li> </ul> <p>Esempio di utilizzo: 'US 122' CICDO#witnessed mdr001;</p>
DSCS	<p>Numero tomba</p> <p>Informazioni relative all'identificatore numerico, qualora il ritrovamento del bene sia connesso ad una deposizione funeraria.</p>	<p>CICDO#Iccd_DSCS rdfs:subClassOf :CICDO#Geographic_Localization; (Class)</p> <p>Esempio di utilizzo: : 'Tomba 4' a Iccd_DSCS, CICDO#Geographic_Localization</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#took_place_at rdfs:domain : Iccd_RE; rdfs:range : Iccd_DSCU;</li> <li>• CICDO#witnessed rdfs:domain : Iccd_DSCU; rdfs:range : Iccd_RE;</li> </ul>

## 8.2 IMPLEMENTAZIONE SU PROTÉGÉ

		<p>Esempio di utilizzo: ‘Tomba 4’ CICDO#witnessed mdr001;</p>
AU – Definizione Culturale		<p>CICDO#Iccd_AU rdfs:subClassOf <b>crm:E12_Production</b> (Class)</p> <p>La classe AU indica tutti i dati relativi ai soggetti intervenuti nella realizzazione del bene e/o nell’ambito culturale di riferimento.</p> <p>Esempio di utilizzo: :au001 a Iccd_AU, crm:E12_Production</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#has_produced rdfs:domain : Iccd_AU; rdfs:range : Bene_Culturale;</li> <li>• CICDO#was_produced_by rdfs:domain : Bene_Culturale; rdfs:range : Iccd_AU;</li> </ul> <p>Esempio di utilizzo: au001 CICDO#has_produced bc001;</p>
AUT	Autore  Dati relativi all’autore.	<p>CICDO#Iccd_AUT rdfs:subClassOf :E21_Person; (Class)</p> <p>Esempio di utilizzo: : aut001 a Iccd_AUT, crm:E21_Person</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#carried_out_by rdfs:domain : Iccd_AU; rdfs:range : Iccd_AUT;</li> <li>• CICDO#performed rdfs:domain : Iccd_AUT; rdfs:range : Iccd_AU;</li> </ul> <p>Esempio di utilizzo: au001 CICDO#carried_out_by aut001;</p>
ATB	Ambito culturale	<p>CICDO#Iccd_ATB rdfs:subClassOf :E13_Attribute_Assignment; (Class)</p>



## CAPITOLO 8 - PROGETTO FASE 4: CICDO

		<p>Esempio di utilizzo: : atb001 a Iccd_ATB, crm:E13_Attribute_Assignment</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#assigned_attribute_to rdfs:domain : Iccd_ATB; rdfs:range : Iccd_AU;</li> <li>• CICDO#was_attribute_by rdfs:domain : Iccd_AU; rdfs:range : Iccd_ATB;</li> </ul> <p>Esempio di utilizzo: atb001 CICDO#assigned_attribute_to au001;</p>
<p><b>CO – Conservazione</b></p>		<p>CICDO#Iccd_CO rdfs:subClassOf <b>crm:E3_Condition_State</b> (Class)</p> <p>La classe CO indica tutti i dati allo stato di conservazione del bene culturale.</p> <p>Esempio di utilizzo: :co001 a Iccd_CO, crm:E3_Condition_State</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#has_condition rdfs:domain : Bene_Culturale; rdfs:range : Iccd_CO;</li> <li>• CICDO#is_condition_of rdfs:domain : Iccd_CO; rdfs:range : Bene_Culturale;</li> </ul> <p>Esempio di utilizzo: co001 is_condition_of bc001;</p>
<p><b>STCC</b></p>	<p><b>Stato di conservazione</b></p> <p>Informazioni sullo stato di conservazione del bene catalogato, con dati presi da un vocabolario chiuso.</p>	<p>CICDO#Iccd_STCC rdfs:subClassOf :E55_Type; (Class)</p> <p>Esempio di utilizzo: : stacc001 a Iccd_STCC, crm:E55_Type</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#has_type rdfs:domain : Iccd_CO; rdfs:range : Iccd_STCC;</li> </ul>

## 8.2 IMPLEMENTAZIONE SU PROTÉGÉ

		<ul style="list-style-type: none"> <li>• CICDO#is_type_of rdfs:domain : Iccd_STCC; rdfs:range : Iccd_CO;</li> </ul> <p>Esempio di utilizzo: stacc001 CICDO#is_type_of co001;</p>
GDPX – Georeferenziazione tramite Punto		<p>CICDO#Iccd_GDPX rdfs:subClassOf <b>crm:E47_Spatial_Coordinates</b> (Class)</p> <p>La classe GDPX indica la coordinata X (Longitudine Est).</p> <p>Esempio di utilizzo: :12.49248 a Iccd_GDPX, crm:E47_Spatial_Coordinates</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#identifies rdfs:domain : CICDO#Actual_Localization; rdfs:range : Iccd_GDPX;</li> <li>• CICDO#is_identified_by rdfs:domain : Iccd_GDPX; rdfs:range : CICDO#Actual_Localization;</li> </ul> <p>La classe CICDO# Actual_Localization è la classe dell'ex livello inventariale AFO#Iccd_PVC</p> <p>Esempio di utilizzo:</p> <ul style="list-style-type: none"> <li>• 12.49248 CICDO#identifies al001;</li> </ul>
GDPY – Georeferenziazione tramite Punto		/* Non diversa alla classe GDPX, se non per la sigla del campo */
TU	Condizione Giuridica	
	Informazioni relative all'acquisizione e la condizione giuridica del bene e i provvedimenti di tutela che lo riguardano	
ACQ	Acquisizione  Informazioni relative alla circostanza e/o titolo in base a cui il bene è pervenuto e	<p>CICDO#Iccd_ACQ rdfs:subClassOf :E10_Transfer_of_Custody; (Class)</p> <p>Esempio di utilizzo: : acq001 a Iccd_ACQ, crm:E10_Transfer_of_Custody</p>

## CAPITOLO 8 - PROGETTO FASE 4: CICDO

	<p>si trova nelle attuali condizioni di proprietà o detenzione.</p>	<p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#custody_transferred_through rdfs:domain : Bene_Culturale; rdfs:range : Iccd_ACQ;</li> <li>• CICDO#transferred_custody_of rdfs:domain : Iccd_ACQ; rdfs:range : Bene_Culturale;</li> </ul> <p>Esempio di utilizzo:</p> <ul style="list-style-type: none"> <li>• acq001 CICDO#transferred_custody_of bc001;</li> </ul>
ACQT	<p><b>Tipo di acquisizione</b></p> <p>Informazioni che riguardano le modalità secondo le quali il bene è acquisito.</p>	<p>CICDO#Iccd_ACQT rdfs:subClassOf :E55_Type; (Class)</p> <p>Esempio di utilizzo: : 'ipoteca' a Iccd_ACQT, crm:E55_Type</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#has_type rdfs:domain : Iccd_ACQ; rdfs:range : Iccd_ACQT;</li> <li>• CICDO#is_type_of rdfs:domain : Iccd_ACQT; rdfs:range : Iccd_ACQ;</li> </ul> <p>Esempio di utilizzo: 'ipoteca' CICDO#is_type_of acq001;</p>
ACQN	<p><b>Nome</b></p> <p>Informazioni sul soggetto fonte dell'acquisizione, nel formato 'cognome, nome'.</p>	<p>CICDO#Iccd_ACQN rdfs:subClassOf :E82_Actor_Appellation; (Class)</p> <p>Esempio di utilizzo: : 'Bedeschi, Sara' a Iccd_ACQN, crm:E82_Actor_Appellation</p> <p>Proprietà:</p> <ul style="list-style-type: none"> <li>• CICDO#acquired_title_through rdfs:domain : Iccd_ACQN; rdfs:range : Iccd_ACQ;</li> <li>• CICDO#has_current_owner rdfs:domain : Iccd_ACQ; rdfs:range : Iccd_ACQN;</li> </ul> <p>Esempio di utilizzo:</p>

## 8.2 IMPLEMENTAZIONE SU PROTÉGÉ

		'Bedeschi, Sara' CICDO#acquired_title_through acq001;
ACQD	Data  Informazioni sull'epoca in cui è avvenuta l'acquisizione del bene.	CICDO#Iccd_ACQD rdfs:subClassOf :E52_Time_Span; (Class)  Esempio di utilizzo: : acqd001 a Iccd_ACQD, crm:E52_Time_Span  Proprietà: • CICDO#has_time_span rdfs:domain : Iccd_ACQ; rdfs:range : Iccd_ACQD; • CICDO#is_time_span_of rdfs:domain : Iccd_ACQD; rdfs:range : Iccd_ACQ;  Esempio di utilizzo: acqd001 CICDO#is_time_span_of acq001;
ACQL	Luogo	CICDO#Iccd_ACQL rdfs:subClassOf :CICDO#Geographic_Localization; (Class)  Esempio di utilizzo: : acql001 a Iccd_ACQL, CICDO#Geographic_Localization  Proprietà: • CICDO#took_place_at rdfs:domain : Iccd_ACQ; rdfs:range : Iccd_ACQL; • CICDO#witnessed rdfs:domain : Iccd_ACQL; rdfs:range : Iccd_ACQ;  Esempio di utilizzo: acql001 CICDO#witnessed acq001;

Tabella 12: CICDO in mapping RDF

### 8.3. CICDO E FABIO

In questo paragrafo viene descritto il modo in cui è stato possibile l'integrazione dell'ontologia *FRBR-aligned Bibliographic Ontology* (FaBiO) all'ontologia di progetto *Central Institute for Cataloguing and Documentation Ontology CICDO*. Come è stato spiegato nel capitolo *FRBR-aligned Bibliographic Ontology (FaBiO)*, ma senza elencare nuovamente i vantaggi di questo import, è stato necessario individuare le entità ICCD equivalenti, o simili quanto più possibile, alle entità descritte dall'ontologia FaBiO.

Definisco quindi un mapping nella prossima tabella, dove possibile e con significato, tra gli elementi dell'ontologia CICDO e quelli di FaBiO che interessano.

FaBiO Entity	CIDOC Entity	CICDO Entity
Work	E28 Conceptual Object	Cultural Heritage (Conceptual Object)
Expression	/* Non mappato */	/* Non mappato */
Manifestation	E22 Man-Made Object	Cultural Heritage (Physical Object)
Item	/* Non mappato */	/* Non mappato */

Tabella 13: Mapping tra FaBiO e CICDO

Come è possibile notare dalla Tabella 13 non è stato necessario, né utile, mappare alcune entità dell'ontologia FaBiO. Questo perché, nell'ambito di questa tesi e nello specifico di un reperto archeologico non c'è necessità di esprimere una *FaBiO:Expression* né un *FaBiO:Item*. Nell'ontologia CICDO considereremo un *CICDO:Cultural Heritage (Physical Object)* la manifestazione di un *CICDO:Cultural Heritage (Conceptual Object)*, con semantica di FaBiO.

### 8.3 CICDO E FABIO

Il costrutto semantico che è stato scelto per definire questo mapping è l'*equivalentTo*. Con questo costrutto, che si applica a classi, object property e datatype property, si definiscono come equivalenti gli individui di uno di questi gruppi elencati. Nello specifico, dopo aver importato l'ontologia FaBiO nel progetto CICDO, nelle entità create *Cultural Heritage (Physical Object)* e *Cultural Heritage (Conceptual Object)* sono stati aggiunti due *equivalentTo* con le entità *FaBiO:Manifestation* e *FaBiO:Work*. Le due classi CICDO saranno poi collegate dalla proprietà **CICDO:hasManifestation** che collegherà l'idea del Bene Culturale, *Cultural Heritage (Conceptual Object)*, alla sua manifestazione in oggetto fisico, *Cultural Heritage (Physical Object)*.

## CONCLUSIONI E SVILUPPI FUTURI

Il mapping di questo schema così complesso ha già dimostrato, almeno dal punto di vista logico e in parte implementativo, la coerenza tra CIDOC-CRM, FRBR FaBiO e lo schema RA dell'ICCD. Resta comunque molto lavoro da fare. ICCD sta ancora completando il suo modello, e verrà rilasciata una versione 4.00 più razionale delle raccomandazioni per la catalogazione. D'altro canto, anche CIDOC-CRM è in continua evoluzione. Il mapping della scheda RA fornisce sicuramente un buon punto di partenza per una possibile convergenza dei due modelli, e su di esso, concetti ed elementi come i paragrafi in comune, renderanno il mapping con gli altri schemi più facili e veloci.

L'uso di standard nella descrizione di attività archeologiche, sia a livello di scavo che di catalogazione del bene, rappresenta una forte richiesta che non può essere limitata, come nel caso dell'ICCD italiano, per garantire una corretta conservazione e gestione degli archivi, soprattutto nel campo dei reperti archeologici. Fino ad oggi non esiste un vocabolario standard internazionale. Questo perché è complicato, vista l'esistenza di norme nazionali e/o locali. Con l'aggiunta delle tecnologie del Semantic Web, la situazione viene ancora più articolata, a causa delle molteplici possibilità di creare adattamenti ai sistemi locali.

I benefici dell'utilizzo di un'ontologia che può essere CIDOC-CRM o più nello specifico CIDDO sono evidenti: da un lato "forza" gli archeologi a concettualizzare e formalizzare il metodo di scavo e registrazione del bene; dall'altro permette ai ricercatori di mantenere un sistema che garantisce il riutilizzo dei dati e interoperabilità. Inoltre,

## CONCLUSIONI E SVILUPPI FUTURI

permette anche di fornire un quadro generale versatile ad implementazioni più specifiche a livello locale.

In conclusione, questo progetto di Tesi propone uno studio e un'implementazione dei concetti, e delle tecniche, per creare un'ontologia, con standard moderni e import di ontologie pubblicate, in cui sia possibile trasferire le proprie informazioni, o crearne di nuove, e agevola notevolmente le successive estensioni, i futuri aggiornamenti e crea un fondamento, sia a livello teorico che implementativo, aperto a qualsiasi sviluppo futuro inerente al campo dei Beni Culturali e alle tecnologie del Semantic Web.

Uno sviluppo futuro, partendo dal lavoro di questa Tesi e dall'ontologia definita in Protégé, potrebbe essere quello di applicare direttamente, a delle informazioni salvate in formato differente dall'RDF (o comunque basato su un'ontologia, ad esempio un database XML), il mapping da un dominio all'altro seguendo quanto definito.

Un altro sviluppo futuro potrebbe essere quello di estensione e di import di altre ontologie non di dominio (ad esempio, Geonames, DBpedia, ecc.) per aggiungere informazione e interoperabilità; per esempio, con un import di Geonames, collegare esplicitamente le definizioni di luogo per arricchire l'informazione restituita e l'esperienza dell'utilizzatore.



## BIBLIOGRAFIA

1. **Harpring, Murtha Baca and Patricia.** *Categories for description of works of art.* Los Angeles : Getty Publications, 2009.
2. **Murtha Baca, Patricia Harpring, Elisa Lanzi, Linda McRae, and Ann Whiteside.** *ataloging Cultural Objects. A guide to describing works and their images.* s.l. : American Library Association, 2008.
3. **Isaac, Bernhard Haslhofer and Antoine.** *the Europeana linked open data pilot. In International Conference on Dublin Core and Metadata Applications.* DC : s.n., 2011.
4. **Doerr, M.** *The CIDOC CRM—an ontological approach to semantic interoperability of metadata.* s.l. : AI Magazine, 2003.
5. **Doerr, Martin.** *Ontologies for cultural heritage.*
6. **Patrick Le Boeuf, Martin Doerr, Christian Emil Ore, Stephen Stead.** *Definition of the CIDOC Conceptual Reference Model 6.2.2.* s.l. : ICOM/CIDOC Documentation Standards Group, 2011.
7. **Olivia Madison, John Byrum, Suzanne Jouguelet, Dorothy McGarry, Nancy Williamson, and.** *Functional requirements for bibliographic records. Final Report. International.* 2009.
8. **K.G. Saur, München.** *Functional Requirements for Authority Data – A Conceptual Model.* s.l. : Glenn E. Patton, 2009.
9. **MarciaLeiZeng, MajaZumer, and Athena Salaba.** *Functional Requirements for Subject.* s.l. : International Federation of Library Associations and Institutions (IFLA), 2010.
10. **Chryssoula Bekiari, Martin Doerr, and Patrick Le Boeuf.** *FRBR object-oriented definition and mapping to FRBR(version 1.0).* s.l. : International Federation of Library Associations and Institutions (IFLA), 2009.
11. **Erin Coburn, Richard Light, Gordon McKenna, Regine Stein, and Axel Vitzthum.** *LIDO – Lightweight Information Describing Objects. Version 1.0.* s.l. : ICOM-CIDOC Working Group Data Harvesting and Interchange, 2010.
12. **R. Kishore, R.Ramesh, and R. Sharman.** *Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems.* s.l. : Springer–Verlag, 2006.

13. **Visser, Ubbo.** *Intelligent information integration for the Semantic Web.* s.l. : Springer–Verlag, 2004.
14. **Riikka Henriksson, Tomi Kauppinen, and Eero Hyvönen.** *Core geographical concepts: Case finnish geo-ontology.* In *Location and theWeb (LocWeb) 2008 workshop, 17th InternationalWorld WideWeb ConferenceWWW2008, volume 300 ofACMInternational ConferenceProceeding Serie.*
15. **Relph, E.** *Place and placeness.* London, : Pilon, 1976.
16. **Eero Hyvönen, Thea Lindquist, Juha Törnroos, and Eetu Mäkelä.** *History on the semantic web as linked data – an event gazetteer and timeline for the World War I. web as linked data – an event gazetteer and timeline for the Worl War I.* In *Proceedings of CIDOC 2012 – Enriching Cultural Heritage.* Helsinki, : CIDOC, 2012.
17. **Thea Lindquist, Eero Hyvönen, Juha Törnroos, and Eetu Mäkelä.** *Leveraging linked data to enhance subject access - a case study of the University of Colorado Boulder’s World War I collection online.* s.l. : International Federation of Library Associations and Institutions (IFLA, 2012.
18. **Willem Robert van Hage, Véronique Malaisé, Roxane Segers, Laura Hollink, and Guus Schreiber.** *Design and use of the simple event model (SEM) Web Semantics: Science, Servicesand Agents on the World Wide Web.* 2011.
19. **Fawcett, Bruce G. Robertson.** *A toolkit to begin an historical semantic web.* *Digital Studies / Le champ numerique.* 2009.