

ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA
SCUOLA DI LETTERE E BENI CULTURALI

Corso di Laurea Magistrale in
Scienze Storiche

La gestione delle relazioni tra soggetto produttore
e risorse archivistiche nel semantic web.

Un modello ontologico

TESI DI LAUREA IN
ARCHIVISTICA INFORMATICA

Relatore:
Prof.ssa *Francesca Tomasi*

Correlatore:
Dott. *Silvio Peroni*

Correlatore:
Dott.ssa *Francesca Ricci*

Presentata da:
Marilena Daquino

Sessione seconda
Anno accademico 2012-2013

INDICE

| | |
|-----------------------------|---|
| Indice delle immagini | 8 |
|-----------------------------|---|

| | |
|--------------------|----|
| Introduzione | 11 |
|--------------------|----|

CAPITOLO PRIMO

Il campo d'indagine: l'archivistica e il soggetto produttore d'archivio

| | |
|---|----|
| Premessa | 17 |
| 1. Il percorso degli standard per la descrizione del soggetto produttore | 19 |
| 1.1 Le ragioni dello sviluppo di ISAAR(CPF) | 19 |
| 1.1.1 La gestione dell'archivio | 20 |
| 1.1.2 I punti di accesso alla ricerca | 21 |
| 2. L'Authority Control nelle discipline umanistiche: un confronto | 22 |
| 3. Descrizione di entità e relazioni tra oggetti | 24 |
| 3.1 Le relazioni negli standard di ICA | 25 |
| 3.1.1 Le relazioni in ISAAR(CPF) | 28 |
| 3.1.2 Le relazioni in ISDF | 29 |
| Conclusioni | 30 |

CAPITOLO SECONDO

Il riorientamento dell'archivistica informatica nella descrizione del soggetto produttore

| | |
|---|----|
| Premessa | 33 |
| 1. Introduzione ai linguaggi di markup per l'archivistica | 35 |
| 2. EAC(CPF): uno standard rappresentativo per ISAAR(CPF) | 37 |
| 2.1 Gli elementi fondamentali di EAC(CPF) | 40 |
| 2.2 Le relazioni in EAC(CPF) | 41 |
| 2.2.1 <functionRelation> e <resourceRelation> | 42 |

| | |
|---|----|
| 3. Il semantic web e le applicazioni nell'archivistica | 44 |
| 3.1 Uno sguardo sulle tecnologie per il semantic web | 48 |
| 3.1.1 Resource Description Framework (RDF) e RDF(S) | 48 |
| 3.1.2 Web Ontology Language (OWL) e Linked Open Data (LOD) | 48 |
| 3.2 L'importanza della modellizzazione per la rappresentazione di un dominio di conoscenza | 50 |
| 3.2.1 I benefici del semantic web nell'ambito del cultural heritage | 51 |
| 3.3 I data model per i beni culturali | 53 |
| 3.3.1 Dublin Core (DC)..... | 54 |
| 3.3.2 Europeana Data Model (EDM)..... | 55 |
| 3.3.3 CIDOC-CRM | 57 |
| 3.4 Progetti per l'archivistica | 58 |
| 3.4.1 Reload | 59 |
| 4. EAC-CPF Ontology..... | 60 |
| 4.1 Le classi | 63 |
| 4.1.1 <i>eac-cpf:entity</i> | 63 |
| 4.1.2 <i>eac-cpf:controlArea</i> | 65 |
| 4.1.3 <i>eac-cpf: descriptionArea</i> | 66 |
| 4.2 Le relazioni | 68 |

CAPITOLO TERZO

Un percorso nella gestione delle relazioni

del soggetto produttore: l'approccio al full text

| | |
|--|----|
| Premessa | 71 |
| 1. Il dibattito interdisciplinare..... | 74 |
| 1.1 “Le fonti dello storico nell'era del computer” | 74 |
| 1.2 Gli archivi fuori dagli archivi | 75 |
| 1.2.1 I punti di accesso nelle descrizioni archivistiche | 77 |
| 1.2.2 Ruoli e funzioni: perché utili added access point | 78 |
| 1.3 Le edizioni digitali: quali spunti dal full text | 79 |
| 2. La descrizione dei ruoli politici | 81 |

| | |
|--|----|
| 2.1 Gli archivi storici delle relazioni politiche | 81 |
| 2.2 Dalle carte al soggetto: il filo rosso delle relazioni politiche | 82 |
| 3. Una proposta eclettica: un modello ontologico per fonti, ruoli ed entità archivistiche | 84 |
| 3.1 Le fonti | 85 |
| 3.2 I ruoli indicizzati | 85 |
| 3.3 I ruoli politici | 86 |

CAPITOLO QUARTO

Una proposta di ontologia: Political Roles (PROles)

| | |
|---|-----|
| Premessa | 89 |
| 1. Introduzione | 91 |
| 2. Le classi principali | 98 |
| 2.1 Agent (<i>foaf:Agent</i>) | 98 |
| 2.2 Political role (<i>proles:PoliticalRole</i>) | 99 |
| 2.3 Source (<i>proles:Source</i>) | 100 |
| 2.4 Time interval (<i>ti:TimeInterval</i>) | 100 |
| 2.5 Political role in time (<i>pro:PoliticalRoleInTime</i>) | 101 |
| 2.6 Event (<i>participation:Event</i>) | 102 |
| 2.7 Participation with political role (<i>proles:ParticipationWithPoliticalRole</i>) | 102 |
| 3. Le object property | 103 |
| 3.1 describes agent (<i>proles:describesAgent</i>) | 103 |
| 3.2 describes situation (<i>proles:describesSituation</i>) | 104 |
| 3.3 has related agent (<i>proles:hasRelatedAgent</i>) | 105 |
| 3.4 has related source (<i>proles:hasRelatedSource</i>) | 105 |
| 3.5 had primary source (<i>prov:hadPrimarySource</i>) | 106 |
| 3.6 was attributed to (<i>prov:wasAttributedTo</i>) | 107 |
| 3.7 has setting (<i>situation:hasSetting</i>) | 107 |
| 3.7.1 holds political role in time (<i>proles:holdsPoliticalRoleInTime</i>) | 107 |
| 3.7.2 is related to role in time | |

| | |
|---|---------|
| (<i>pro:isRelatedToRoleInTime</i>) | 108 |
| 3.7.2.1 is document context for (<i>pro:isDocumentContextFor</i>) | 108 |
| 3.7.2.2 is organization context for (<i>pro:isOrganizationContextFor</i>) | 109 |
| 3.7.2.3 is personal context for (<i>pro:isPersonalContextFor</i>) | 109 |
| 3.7.2.4 is place context for (<i>proles:isPlaceContextFor</i>) | 109 |
| 3.7.3 is political role in (<i>proles:isPoliticalRoleIn</i>) | 110 |
| 3.7.4 is included in participation political role (<i>proles:isIncludedInParticipationWithPoliticalRole</i>) | 110 |
| 3.8 is setting for (<i>situation:isSettingFor</i>) | 110 |
| 3.8.1 at time (<i>tvc:atTime, tisit:atTime</i>) | 111 |
| 3.8.2 is held by (<i>pro:isHeldBy</i>) | 111 |
| 3.8.3 relates to entity (<i>pro:relatesToEntity</i>) | 112 |
| 3.8.4 with role (<i>pro:withRole</i>) | 112 |
| 3.8.5 includes agent (<i>proles:includesAgent</i>) | 112 |
| 3.8.6 participation includes political role in time (<i>proles:participationIncludesPoliticalRoleInTime</i>) | 113 |
| 4. Data property e Annotation property | 113 |
| 5. Applicazioni ed esempi | 114 |
| 5.1 Relazioni politiche tra persone | 115 |
| 5.2 Ruoli politici indicizzati | 116 |
| 5.3 Relazioni e ruoli politici in un contesto complesso | 117 |
| 6. PROles ed EAC-CPF Ontology | 119 |
| 6.1 L'allineamento tra PROles e EAC-CPF Ontology | 120 |
| Legenda | 124 |
| Bibliografia | 126 |
| Sitografia | 130 |

INDICE DELLE IMMAGINI

CAPITOLO SECONDO

| | |
|--|----|
| Fig. 2.1 Classi, proprietà e visualizzazione grafica di EAC-CPF Schema Ontology | 61 |
| Fig. 2.2 Classi, proprietà e visualizzazione grafica di EAC-CPF Ontology | 62 |

CAPITOLO QUARTO

| | |
|--|-----|
| Fig. 4.1 Rappresentazione grafica dell'estensione di PRO in PROles | 91 |
| Fig. 4.2 Rappresentazione grafica dei modelli riusati in PROles | 94 |
| Fig. 4.3 Livelli rappresentativi delle relazioni di contesto in PROles | 96 |
| Fig. 4.4 Rappresentazione grafica delle proprietà di PROles per l'allineamento con le altre ontologie | 97 |
| Fig. 4.5 Rappresentazione grafica delle catene di proprietà di <i>proles:describesAgent</i> | 104 |
| Fig. 4.6 Rappresentazione grafica delle catene di proprietà di <i>proles:hasRelatedAgent</i> | 105 |
| Fig. 4.7 Rappresentazione grafica delle catene di proprietà di <i>proles:hasRelatedSource</i> | 106 |
| Fig. 4.8 Rappresentazione grafica della classe <i>proles:PoliticalRoleInTime</i> applicata alle relazioni politiche | 115 |
| Fig. 4.9 Rappresentazione grafica della classe <i>proles:PoliticalRoleInTime</i> applicata ai ruoli politici | 117 |
| Fig. 4.10 Rappresentazione grafica della classe <i>proles:ParticipationWithPoliticalRole</i> | 118 |
| Fig. 4.11 Rappresentazione grafica del mapping tra EAC-CPF e PROles | 120 |
| Fig. 4.12 Rappresentazione grafica dei ruoli politici come possibili sottoclassi di <i>eac-cpf:relation</i> | 121 |
| Fig. 4.13 Rappresentazione grafica di un'applicazione combinata di PROles e EAC-CPF Ontology | 122 |

INTRODUZIONE

La ricerca nel campo del *cultural heritage management* ha adottato negli ultimi decenni le tecnologie web quali strumenti privilegiati per stabilire i nuovi approcci e indirizzi nella valorizzazione della conoscenza.

Questa tesi si colloca nell'ambito interdisciplinare tra le scienze umanistiche e informatiche e si fonda sulla consapevolezza del reciproco arricchimento che può derivare dal continuo confronto, le une disponendo di mezzi più espressivi e popolari per divulgare il proprio patrimonio e le altre usufruendo di “materia prima” autorevole (ossia dati strutturati di qualità e alto livello di fiducia) in fase di sperimentazione. Lo studio dei punti di tangenza tra le discipline muove da due ambiti precisi, ovvero le applicazioni informatiche nel campo dell'archivistica e gli sviluppi del *semantic web* nel settore delle *digital humanities*.

Il campo archivistico si è dimostrato in questa sede un interessante attore culturale che ha saputo tradurre la propria tradizione teorica in strumenti informativi digitali, mantenendo saldi i presupposti oggettivi di organizzazione del suo patrimonio, ed ha sviluppato metodi efficienti per la sua valorizzazione facendo uso delle tecnologie per il *semantic web*.

L'obiettivo qui proposto è quello di continuare a testare le potenzialità del web di dati nel dominio archivistico e proporre spunti per ulteriori ricerche, nell'ottica di formalizzazione di un modello ontologico archivistico che tenga conto delle numerose potenzialità descrittive offerte da approcci informatici

e studi in altri settori umanistici. La modellizzazione ontologica riveste infatti un ruolo cruciale nell'espressione della struttura di un dominio di conoscenza, poiché l'annotazione semantica di risorse sul web fondata su un modello di riferimento permette una maggiore interoperabilità (e quindi condivisione) dei dati ed è un presupposto per migliorare la ricerca.

L'input alla sperimentazione è giunto dalla recente formalizzazione dell'ontologia *Encoded Archival Context – Corporate bodies, Persons and Families (EAC-CPF)*¹, per la rappresentazione delle entità (i soggetti produttori e non) legate ai complessi archivistici, e delle ontologie *Semantic Publishing and Referencing (SPAR)*², per la descrizione delle entità bibliografiche: di queste ultime, *Publishing Roles Ontology (PRO)*³ è sembrata quella più congeniale allo sviluppo di un ragionamento su nuovi aspetti descrittivi e relazionali da integrare al modello archivistico EAC-CPF.

EAC-CPF Ontology fornisce un modello per la descrizione di persone, famiglie, enti che hanno relazioni con la documentazione archivistica: a differenza degli *authority systems* del mondo biblioteconomico, che si concentrano principalmente sulle forme autorizzate del nome delle entità al fine di identificarle univocamente, lo standard EAC-CPF⁴ su cui si basa l'ontologia omonima prevede l'approfondimento di ulteriori aspetti inerenti la “natura” del soggetto, ovvero descrive elementi che identificano il contesto più ampio in cui è coinvolto il soggetto produttore d'archivio. L'obiettivo dello standard EAC-CPF non è solo la disambiguazione dell'identità del soggetto, ma una sua descrizione articolata che assume valore di fonte autonoma.

Le relazioni che intercorrono tra soggetti, risorse e attività assumono un ruolo fondamentale nella rappresentazione formale del contesto di produzione

.....
1 Cfr. Mazzini S., Ricci F., *EAC-CPF Ontology and Linked Archival Data*, in «Proceedings of the 1st International Workshop on Semantic Digital Archives, Berlin 29/9/2011 (SDA 2011)», <http://ceur-ws.org/Vol-801/paper6.pdf>; Regesta.exe, "EAC-CPF Ontology", <http://labs.regesta.com/progettoReload/lontologia-eac-cpf/>

2 Cfr. The Semantic Publishing And Referencing ontologies (SPAR), <http://purl.org/spar>

3 Cfr. Silvio Peroni, David Shotton, Fabio Vitali, *Scholarly publishing and the Linked Data: describing roles, statuses, temporal and contextual extents*, in H. Sack, T. Pellegrini, *Proceedings of the 8th International Conference on Semantic Systems*, ACM, New York, 2012, <http://speroni.web.cs.unibo.it/publications/peroni-2012-scholarly-publishing-linked.pdf>; Publishing Roles Ontology (PRO), <http://purl.org/spar/pro>

4 Cfr. Staatsbibliothek zu Berlin, "Encoded Archival Context - Corporate bodies, persons, families (EAC-CPF)", <http://eac.staatsbibliothek-berlin.de/>

di un complesso archivistico; l'*International Council on Archives (ICA)*⁵ ha promosso la redazione di ulteriori standard per l'arricchimento degli scenari relazionali, tra i quali *International Standard for Describing Functions (ISDF)*⁶ su cui ci si è soffermati in questa sede. La descrizione delle funzioni di un ente è utile innanzitutto alla ricostruzione del percorso che ha portato alla produzione di un documento, nonché a stabilire legami tra entità e risorse sulla base di un concetto: un documento viene prodotto nell'espletamento di una funzione amministrativa di un ente e/o è relazionato a più figure che rivestono un ruolo nei suoi confronti.

In questa sede ci si è soffermati sull'analisi dei ruoli delle persone piuttosto che sulle funzioni, le quali sono maggiormente vincolate al contesto istituzionale, per approfondire le relazioni tra entità e risorse ed arricchire le descrizioni di entrambe.

PRO descrive i ruoli editoriali che una persona/ente intrattiene nei confronti di una documentazione: è un modello facilmente estendibile, per cui è stato possibile ampliare lo spettro di ruoli (non esclusivamente legati alla pubblicazione del documento) semplicemente aggiungendo nuove istanze. Il focus si è spostato dal contesto editoriale a quello politico, per poter allargare il raggio di relazioni tra persone/enti – ruoli/relazioni politiche – fonti. L'estensione di *PRO* ha interessato in principio la formalizzazione di istanze per descrivere ruoli e relazioni politiche ed in seguito ha richiesto un arricchimento degli scenari per rappresentare il contesto di svolgimento di quei ruoli/relazioni. In fase di modellizzazione è stata data una notevole importanza alla provenienza e all'autorevolezza delle asserzioni fatte sull'effettivo esercizio di un ruolo politico, perciò, ad ogni *statement* su un ruolo politico indicizzato temporalmente e contestualizzato, è stato associato un meta-contesto per definire la fonte e il responsabile culturale (un “editore”) dell'affermazione estratta dalla fonte.

Il risultato è *PROles*⁷, un'ontologia in OWL 2 DL per descrivere le relazioni politiche tra persone, estrapolate dal testo pieno di fonti che ne attestano in-

5 Cfr. International Council on Archives (ICA), <http://www.ica.org/>

6 Cfr. CBPS - Sub-Committee on Descriptive Standards, "International Standard for Describing Functions (ISDF)", 2011, <http://www.ica.org/10208/standards/isdf-international-standard-for-describing-functions.html>

7 Cfr. Marilena Daquino, Silvio Peroni, Francesca Tomasi, "Political Roles Ontology (PROles)", <http://www.essepuntato.it/2013/10/politicalroles>

formazioni. Per la sua formalizzazione è stato utilizzato *Protégé editor*⁸ e per la visualizzazione HTML del suo element set il tool *Live Owl Documentation Environment (LODE)*⁹. Vengono forniti dei casi d'uso ed esempi per mostrare le potenziali applicazioni del modello, tratti da uno studio sui dati riversati in xDams dall'*Istituto per i beni artistici, culturali e naturali dell'Emilia Romagna (IBC-ER)*¹⁰ inerenti le carte di Andrea Costa.

Infine, *PROles* viene allineato ad *EAC-CPF Ontology*: il mapping tra le due ontologie non ha richiesto una comparazione di tutti gli elementi dei due modelli, ma solo alcune relazioni gerarchiche tra proprietà generiche e l'equivalenza fra classi sostanzialmente uguali.

L'intento è dimostrare i benefici che l'archivistica può trarre dall'estensione del proprio modello ontologico con altre soluzioni, ideate per rappresentare nodi concettuali appartenenti allo stesso dibattito teorico ma sviluppate in altri campi di studio delle *digital humanities*. La rappresentazione formale dei ruoli esercitati da entità in ambito archivistico è all'ordine del giorno nel dibattito internazionale, ma fatica a trovare un seguito nella sperimentazione italiana: è dall'ambito biblioteconomico (le *SPAR Ontologies* ne sono un esempio) che giungono stimoli ad una implementazione di maggiori informazioni di contesto.

L'auspicio è che questo primo tentativo di confronto del modello EAC-CPF con formalizzazioni già in uso in ambiti affini, ampli gli orizzonti della discussione e dia un contributo al percorso di conservazione, valorizzazione e “creazione” di conoscenza.

.....
8 Cfr. Stanford Center for Biomedical Informatics Research, "Protégé", <http://protege.stanford.edu/>

9 Cfr. Silvio Peroni, David Shotton, Fabio Vitali, *The Live OWL Documentation Environment: A Tool for the Automatic Generation of Ontology Documentation*, in A. Teije, J. Völker, S. Handschuh, H. Stuckenschmidt, M. d' Aquin, A. Nikolov, ... N. Hernandez, *Proceedings of the 18th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2012), Lecture Notes in Computer Science*, Berlin, Heidelberg, Germany, Springer, 2012; Live Owl Documentation Environment (LODE), www.essepuntato.it/lode

10 Cfr. Argelli B., *Una regione in rete: gli archivi storici dell'Emilia-Romagna e il progetto IBC*, in «DigItalia, Rivista del digitale nei beni culturali», III, 2, 2008, <http://digitalia.sbn.it/article/view/449>

CAPITOLO PRIMO

Il campo d'indagine: l'archivistica e il soggetto produttore d'archivio

PREMESSA

Il percorso di ricerca portato avanti dall'archivistica nell'ambito della conservazione del patrimonio culturale internazionale dagli anni '80 ad oggi è unanimemente uno dei più inclini alla ricezione e alla sistematica articolazione dei problemi inerenti l'organizzazione della memoria e la sua valorizzazione.

Benché si fondi su un dominio di conoscenza fortemente strutturato, intendendo qui la necessità di un "vincolo" non solo materiale ma anche strettamente concettuale tra gli oggetti culturali, la disciplina archivistica ha portato avanti un processo di elaborazione di standard metodologici predisposto ad essere modello d'eccellenza e di riferimento per più settori umanistici.

Questa considerazione discende innanzitutto dalla consapevolezza del notevole livello di analiticità richiesto dallo studio di documentazioni che hanno un ruolo legale e storico, unitamente alle proposte concrete di ordinamento e all'attenzione posta sulla fruibilità dei complessi archivistici da parte di utenti che non intendono la struttura del complesso stesso e devono quindi ricorrere a strumenti di corredo.

Da queste premesse si muove la ricerca archivistica, attraverso stimoli e proposte che aprono sia alle esigenze di lunga durata dell'operatore culturale sia a quelle contingenti dell'utente: tra questi due soggetti lo scarto significativo di conoscenza viene colmato grazie all'adeguatezza degli strumenti di ricerca

che vengono forniti in sede di consultazione dei fondi e che rivestono di conseguenza un ruolo centrale nell'elaborazione teorica degli archivisti.

Proprio il dualismo tra astrazione organizzativa e pragmatismo informativo ha permesso di sintetizzare in questo campo del cultural heritage management alcune tappe importanti di un iter volto a restituire agli oggetti culturali tutto il loro potenziale comunicativo.

Fra i contributi maggiori che questa disciplina è stata in grado di dare, qui viene presa in esame la gestione dei record di autorità archivistici, ossia la formalizzazione di elementi standardizzati per la descrizione della natura dei soggetti produttori, passo fondamentale per aprire gli archivi stessi alle reti comunicative che travalicano il campo settoriale della materia.

In questo capitolo viene data una sintetica panoramica del processo di definizione degli standard per la descrizione archivistica, con particolare attenzione alle scelte adottate per la descrizione del soggetto produttore d'archivio e alle relazioni che esso intrattiene con altri oggetti reali: l'intento è fornire un background del dibattito teorico che è stato alla base delle riflessioni e delle proposte sviluppate nei capitoli successivi.

1. IL PERCORSO DEGLI STANDARD PER LA DESCRIZIONE ARCHIVISTICA DEL SOGGETTO PRODUTTORE

Tra la fine degli anni '80 e il decennio successivo su proposta del Consiglio Internazionale degli Archivi (ICA) vengono elaborati da una commissione ad hoc due fondamentali documenti di riferimento per la descrizione archivistica: *International Standard of Archival Description (General) (ISAD(G))*¹ e *International Standard Archival Authority Record for Corporate bodies, Persons and Families (ISAAR(CPF))*².

ISAD(G) viene approvato nel 1994 e pubblicato l'anno seguente con lo scopo di dar vita a norme generali per la descrizione del contesto e del contenuto di documentazione archivistica; formalizzando 26 elementi descrittivi autoesplicativi, lo standard mira a rendere le descrizioni di complessi archivistici accessibili ai fruitori e condivisibili tra istituzioni conservatrici.

A seguito dell'attenta analisi di queste finalità nasce anche *ISAAR(CPF)*: pubblicato da ICA nel 1996 e revisionato nel 2004, lo standard per l'elaborazione di record di autorità archivistici prevede la descrizione separata e approfondita di entità produttrici di archivi tramite 27 elementi descrittivi.

La decisione di separare gli *authority record* dalla descrizione del contesto di produzione degli archivi è il risultato di un proficuo dibattito che ha posto in essere una prassi volta a promuovere la gestione di informazioni d'autorità come punti privilegiati, quindi separati, di accesso agli archivi stessi³.

1.1 LE RAGIONI DELLO SVILUPPO DI ISAAR(CPF)

Le necessità di una standardizzazione autonoma della descrizione dei soggetti produttori sono molteplici. Come spiegato nelle linee guida redatte da ICA a presentazione della seconda edizione italiana dello standard:

.....

1 Cfr. CBPS - Sub-Committee on Descriptive Standards, "International standard of archival description (general) (ISAD(G))", [http://www.icacds.org.uk/eng/ISAD\(G\).pdf](http://www.icacds.org.uk/eng/ISAD(G).pdf)

2 Cfr. CPBS, "International standard archival authority record for corporate bodies, persons and families (ISAAR(CPF)), Second Edition", http://www.icacds.org.uk/eng/isaar2ndedn-e_3_1.pdf

3 Cfr. International Council on Archives (ICA), "History of ICA/CDS", <http://www.icacds.org.uk/eng/history.htm>

I record di autorità archivistici possono essere utilizzati:

- a. per descrivere enti, persone o famiglie come elementi di un sistema di descrizione archivistico e/o
- b. per sottoporre a controllo d'autorità l'elaborazione e l'uso di chiavi d'accesso alle descrizioni archivistiche;
- c. per documentare le relazioni fra differenti soggetti produttori e fra questi e la documentazione da essi prodotta e/o altre risorse ad essi relative o ascrivibili⁴.

Lo scopo è fornire un'adeguata definizione del soggetto produttore attraverso quattro aree informative: area dell'identificazione, area della descrizione, area delle relazioni, area di controllo. Gestendo separatamente queste informazioni rispetto a quelle dei complessi archivistici è possibile dar vita ad una nuova risorsa informativa autonoma, che vede la sua ragion d'essere proprio nello svolgere un ruolo primario di chiave d'accesso alla documentazione collegata: è allora possibile la condivisione di questa singola descrizione normalizzata fra più istituzioni conservatrici di documentazione relativa allo stesso soggetto produttore, o detentrici di risorse collegate ad esso a vario titolo e di varia natura, consentendo in ultima analisi sia una migliore gestione dell'archivio sia una semplificazione della ricerca.

Ne derivano allora benefici in termini di scambio ed in termini di accessibilità, identificando univocamente un'entità produttrice e offrendo maggiori competenze all'utente finale, che necessita principalmente (o almeno in prima battuta) di indagare il contesto dei documenti per capire quanto questi siano utili ai suoi fini.

1.1.1 La gestione dell'archivio

Una parte cospicua della tradizione archivistica del '900 vede nel rapporto monolitico tra descrizione del soggetto produttore e descrizione della documentazione collegata il presupposto teorico della disciplina stessa: questa

.....
⁴ Stefano Vitali (trad. italiana di), *ISAAR(CPF): Standard internazionale per i record d'autorità archivistici di enti, persone, famiglie, traduzione italiana di International Standard Archival Authority Records for Corporate Bodies, Persons and Families, seconda edizione*, 2004, in «Rassegna degli Archivi di Stato», 2003, p. 205, <http://www.icar.beniculturali.it/biblio/pdf/standard/isaar%202.pdf>

tendenza, che ha dato vita ad una visione statica della descrizione degli archivi, è stata messa in discussione a partire dagli anni '70 da una generale presa di coscienza sulla complessità intrinseca della natura del soggetto produttore.

Il dibattito sul cambiamento di approccio alla descrizione ha avuto ripercussioni nelle scelte di ordinamento, nella riflessione sugli schemi interpretativi (per rispecchiare l'organizzazione della memoria voluta dall'ente e/o quella voluta dall'archivista) e sul fruitore di archivi, il quale ha una comprensione mirata a ciò che realmente gli è utile piuttosto che ad una piena coscienza dell'organizzazione complessiva dell'archivio.

A dar conto efficacemente del modo in cui sia le vicissitudini della trasmissione documentaria che gli intrecci di competenze e strutture nelle moderne organizzazioni si ripercuotono nelle forme di sedimentazione della documentazione contemporanea, non sembra adeguato un modello di rapporto fra soggetti produttori e archivi come quello tradizionale.

(...) Ciò significa che, nel rappresentare le relazioni che intercorrono fra gli archivi e i loro produttori, bisogna costruire un modello che preveda la possibilità che a un fondo archivistico corrispondano più soggetti produttori e, viceversa, che ad un soggetto produttore corrispondano più fondi archivistici. Un modello quindi che disegni un rapporto plurilineare, multidimensionale e dinamico in prospettiva diacronica⁵.

Come spiega Stefano Vitali, tra i promotori e membro della commissione che ha rivisto lo standard, ISAAR(CPF) ha il merito di proseguire la strada tracciata da questa prospettiva, che non ha ricadute solo prettamente organizzative, ma più diffusamente sull'approccio teorico degli archivisti.

1.1.2 I punti di accesso alla ricerca

Rendere indipendente la descrizione del soggetto produttore da quelle della documentazione correlata e degli enti conservatori ha dato modo alle comunità

.....
⁵ Cfr. Stefano Vitali, *La seconda edizione di ISAAR(CPF) e il controllo d'autorità nei sistemi di descrizione archivistica*, in Mauro Guerrini e Barbara B. Tillett (a cura di), *Authority Control: definizione ed esperienze internazionali*, Firenze, University Press, 2003, p. 3, http://www.sba.unifi.it/ac/relazioni/vitali_ita.pdf

archivistiche, che spesso interagiscono a livello internazionale, di scoprire le potenzialità della comunicazione partendo da una base di per lo scambio di informazioni autoesplicative, normalizzate e riusabili.

La descrizione del contenuto dei complessi archivistici può essere rintracciata sempre facendo riferimento alle singole risorse collegate, mentre un unico contesto storico-politico-istituzionale diviene valido per tutto quanto sia collegabile all'entità produttrice; divenendo così una risorsa a sua volta, il contesto unico di provenienza dei complessi documentali, si pone come *hub* di infinite possibilità relazionali. Questo stimolo offerto dallo studio degli *authority record* è condiviso da altri settori umanistici, in particolar modo dalla biblioteconomia, verso la quale l'archivistica è debitrice di quella riflessione sulle intestazioni d'autorità ma che quest'ultima sviluppa in maniera più articolata, ambendo a fornire chiavi di accesso più complesse e complete dei soli record bibliografici.

2. L'AUTHORITY CONTROL NELLE DISCIPLINE UMANISTICHE: UN CONFRONTO

Come si evince dall'introduzione allo standard ISAAR(CPF), i record d'autorità archivistici presentano numerose convergenze con il controllo d'autorità dei nomi degli autori all'interno dei cataloghi di biblioteca.

I record d'autorità archivistici sono simili a quelli bibliotecari in quanto entrambi intendono rendere possibile l'elaborazione di chiavi d'accesso normalizzate alle descrizioni. Di tali chiavi d'accesso, il nome del soggetto produttore è una delle più importanti.

(...) I record d'autorità archivistici tuttavia, richiedono di soddisfare un insieme molto più ampio di requisiti rispetto ai record d'autorità bibliografici. Questi requisiti aggiuntivi sono connessi al rilievo che nei sistemi di descrizione archivistica hanno le informazioni che documentano i soggetti produttori e il contesto di produzione della documentazione. Perciò, i record d'autorità archivistici sono più approfonditi ed in genere contengono molte più informazioni rispetto a quelli bibliotecari⁶.

.....

⁶ Vitali S., *ISAAR (CPF)*, ivi, p. 10

Partendo da queste affermazioni si può allora desumere che, date le sue precise attitudini all'articolazione del contesto delle *authority*, l'archivistica svolge una funzione più costruttiva rispetto al mero controllo delle intestazioni e ha più possibilità di diventare punto di riferimento e di appoggio alla ricerca, oltre che alla catalogazione.

David Bearman, archivista americano, già nel 1989 sosteneva che nell'ambito archivistico gli *authority record* hanno un maggiore portato informativo e sono naturalmente più propensi ad arricchire la terminologia utilizzabile nel processo di formulazione degli *access point*, piuttosto che a guidarli e limitarli nella definizione⁷: la possibilità di creare *reference file* invece di *authority file* punta cioè a generare fonti di supporto alla ricerca che privilegino il *recall* piuttosto che la *precision*, per parafrasare il concetto espresso da Bearman.

È prassi comune nello sviluppo di sistemi descrittivi svolgere un'analisi dell'utenza e dei percorsi di ricerca da loro effettuabili: da questi studi sulle ricerche degli utenti si evince quali differenti risultati si auspichino archivisti o creatori di *authority files* rispetto ai ricercatori.

Mentre i primi puntano alla *precision* (termine mutuato dall'ambito dell'*information retrieval*) dei risultati forniti dalla ricerca, i secondi preferiscono reperire (*recall*) più materiale possibile, anche a discapito della pertinenza, in modo da poter vagliare ogni possibilità e non rischiare di perdere qualche informazione rilevante trascurata dall'alto grado di *precision*⁸.

In definitiva, una disciplina, quella biblioteconomica, si attiene alla ricerca di precisione nell'intestazione dei nomi come riferimento principale di qualità informativa dell'*authority file* (da corredare successivamente con relazioni ad altre risorse); l'archivistica, oltre a questo, deve formalizzare una cospicua serie di punti di accesso. Si proietta cioè in un contesto descrittivo già intrinsecamente multidimensionale, a cui poi fanno capo altre risorse collegate. Più esplicitamente,

(...) gli *authority record* dei soggetti produttori, quindi, sono destinati a comprendere un insieme assai più complesso di informazioni rispetto ai tradizionali record

.....

7 Cfr. David Bearman, *Authority Control Issues and Prospects*, in «American Archivist», Vol. 52, 1989, <http://archivists.metapress.com/content/g562600um1063123/fulltext.pdf>

8 Cfr. Maurizio Savoia, *Descrizione e liste d'autorità: alcune proposte nordamericane*, in *Standard, vocabolari controllati, liste d'autorità. Atti del seminario, Milano, 25 maggio 1994*, Milano: Regione Lombardia, 1995, pp. 71-94

d'autorità bibliografici. Nella "Description area" convergono così elementi cruciali per la descrizione del soggetto produttore di archivi, quali: le date di esistenza (5.2.1); la storia o la biografia del soggetto (5.2.2); le aree geografiche nelle quali ha operato (5.1.3); la condizione giuridica (5.1.4); le funzioni, occupazioni e attività (5.1.5); il mandato e le fonti dell'autorità esercitata (5.1.6); la struttura interna (5.1.7); informazioni sul contesto politico, sociale, culturale all'interno del quale il soggetto ha operato (5.1.8); nonché tutte le altre informazioni di possibile rilevanza non comprese nei precedenti elementi (5.1.9)⁹.

Si può quindi dire, seguendo il discorso di Vitali, che l'archivistica e ISAAR(CPF) nello specifico pongono l'accento non tanto (e non solo) sulla formulazione di denominazioni d'autorità disambigue, ma sulla gestione integrale di descrizioni di entità: ancora più precisamente, l'aspetto "denominazione" è conseguente alla necessità di identificare univocamente il contesto della documentazione.

3. DESCRIZIONE DI ENTITÀ E RELAZIONI TRA OGGETTI

Nella prima edizione di ISAAR(CPF) gli scopi proposti dallo standard non erano molto lontani da quelli della catalogazione bibliografica: creare liste di autorità di soggetti produttori di cui si fornivano anche descrizioni.

Nella seconda edizione si sostiene un procedimento inverso: fornire regole per la descrizione di entità di cui vengono fornite le varianti di intestazione d'autorità e relazionate in modo nuovo, in un'apposita area descrittiva. Questo cambio di visuale ha aperto nuovi orizzonti relazionali che hanno preso forma in un network di collegamenti ideati per gestire non solo rapporti di "paternità culturale" tra soggetti e documentazione prodotta, ma tra soggetti e soggetti, soggetti e risorse collegate, tra soggetti, funzioni e risorse. L'apertura di uno scenario simile ha dato avvio a "quella che potrebbe definirsi l'uscita dei sistemi archivistici da loro stessi"¹⁰, l'apertura dei dati a nuovi legami, nuovi dati, nuove potenzialità comunicative, rendendosi al contempo promotori di una precisa prassi metodologica.

.....
⁹ S. Vitali, *La seconda edizione di ISAAR(CPF)*, ivi, p. 5

¹⁰ S. Vitali, *La seconda edizione di ISAAR(CPF)*, ivi, p. 6

3.1 LE RELAZIONI NEGLI STANDARD DI ICA

Le riflessioni degli archivisti si sono estese, in aggiunta alle liste d'autorità, verso altri elementi, sempre specifici della descrizione archivistica, che si prestano a forme controllate e sono accessibili tramite gli *authority file*: in primo luogo l'aspetto temporale, fondamentale per documentare i cambiamenti delle entità descritte e relazionarne dinamicamente i “nomi”, ma anche i luoghi (coinvolti nell'espletamento delle funzioni del soggetto produttore) e le relazioni con risorse esterne al sistema informativo a cui appartiene l'entità produttrice.

Nell'ultimo decennio sono stati formalizzati da ICA accanto a ISAD(G) e ISAAR(CPF) altri due standard che concorrono al processo di separazione e apertura delle descrizioni archivistiche, fermi restando però i due poli fondamentali di queste, cioè i soggetti produttori e il materiale archivistico, e il riconoscimento del metodo della provenienza come linea guida: gli standard in questione sono *International Standard for Describing Functions (ISDF)*¹¹ e *International Standard for Describing Institutions with Archival Holdings (ISDIAH)*¹².

ISDF, edito nel 2007, provvede a descrivere le funzioni svolte da un istituzione creatrice di complessi archivistici, definendone anche sotto-funzioni, branche amministrative e attività: lo scopo è esplicitare il nesso tra un soggetto produttore e le sue carte tramite una funzione, una finalità amministrativa, che ha dato vita alla documentazione stessa.

ISDIAH, edito nel 2008, si propone invece come standard per la descrizione di enti conservatori di archivi e/o altre istituzioni non esplicitamente archivistiche (come musei, biblioteche, istituti) detentrici di documentazione archivistica. Questa standardizzazione è fondamentale soprattutto per il reperimento del materiale archivistico, al quale la descrizione ISDIAH viene collegata tramite link alla descrizione ISAD(G), e per conoscere di quali soggetti produttori detiene gli archivi, tramite link alla descrizione ISAAR(CPF). Vengono forniti anche contatti dell'ente e informazioni sui servizi offerti, al fine di facilitare i movimenti degli utenti.

.....

11 Cfr. CBPS - Sub-Committee on Descriptive Standards, "International Standard for Describing Functions (ISDF)", <http://www.ica.org/10208/standards/isdf-international-standard-for-describing-functions.html>

12 Cfr. CBPS, "International Standard for Describing Institutions with Archival Holdings (ISDIAH)", <http://www.ica.org/?lid=10198>

Dalla congiunzione di questi quattro standard prende forma un *archival information system*, dove aumentano i nodi da descrivere (complessi archivistici, soggetti produttori, funzioni, enti conservatori, risorse collegate) e da relazionare tramite forme standardizzate che diano effettivamente idea della complessità della gestione a monte: così i soggetti produttori di più fondi conservati in diversi istituti, di varia natura e collocazione, possono essere facilmente rintracciabili tramite un'unica descrizione del soggetto, o risorse archivistiche simili conservate in più luoghi possono fare capo ad una sola descrizione di contenuto direttamente linkata a risorse affini e rispettivi enti conservatori.

Le potenzialità di questo sistema di interconnessioni sono notevoli se si considera la natura spesso viscosa del materiale archivistico, che raramente rispecchia effettivamente un'istituzione produttrice in tutte le sue fasi di vita ed altrettanto difficilmente è ordinato con uno schema classificatorio astrattamente preciso e coerente col metodo storico, mentre più spesso è depositato in luoghi diversi e rispecchia più esigenze pratiche del produttore¹³.

Per offrire una normalizzazione dei legami tra entità archivistiche di vario tipo all'interno di un *archival descriptive system*, sono state formalizzate da ICA alcune regole fondamentali per descrivere la natura delle relazioni¹⁴.

All'interno di ciascuna area specifica dedicata alle relazioni (negli standard di descrizione dei record archivistici e dei record d'autorità) è consigliabile fornire tre elementi formalizzati per definire univocamente la relazione:

<identifier/name>

Un codice di riferimento o una forma autorizzata del nome dell'entità o risorsa che è relazionata con quella descritta.

<nature/description>

Una categorizzazione normalizzata delle relazioni prescritta dagli standard internazionali che può essere accompagnata da una descrizione testuale sulla

.....
¹³ Cfr. Claudio Pavone, *Di carte e d'altro*, in Isabella Zanni Rosiello (a cura di), *Intorno agli archivi e alle istituzioni. Scritti di Claudio Pavone*, Ministero per i beni e le attività culturali, 2004, pp. 37-89, http://www.archivi.beniculturali.it/dga/uploads/documents/Saggi/Saggi_84_I.pdf

¹⁴ Cfr. CBPS, "Relationship in archival descriptive systems", <http://www.ica.org/13149/standards/cbps-relationship-in-archival-descriptive-systems.html>

natura della relazione. Le convenzioni per i nomi delle relazioni sono:

For the relationships between corporate bodies, persons or families and archival materials

- Creator of/created by.
- Author of/authored by.
- Custodian of/held by.
- Owner of/owned by.
- Controller of/controlled by.
- Copyright owner of/copyright owned.
- Subject of/concern.

For the relationships between corporate bodies, persons or families and functions:

- Controller of/controlled by.
- Owner of/owned by.
- Performer of/performed by.

For the relationships between functions and archival materials

- Creator of/created by.
- Subject of/concern.
- Controller of/controlled by¹⁵.

<dates>

Una singola data o un *range* di date che includano l'arco temporale della durata della relazione.

Ai fini dello studio delle relazioni che intercorrono tra soggetti produttori e risorse collegate, verranno presi in considerazione solo due dei quattro standard, ovvero ISAAR(CPF) e ISDF, che rispondono meglio alle esigenze di gestione delle relazioni e che si pongono come mediatori tra le risorse piuttosto che come approdi finali (quali i contenuti dei record di ISAD(G) e i conservatori dei record di ISDIAH).

.....
15 CBPS, "Relationship in archival descriptive systems", ivi, pp. 3-4

3.1.1 Le relazioni in ISAAR(CPF)

Come spiegato precedentemente, ISAAR(CPF) nasce con l'intento di creare record d'autorità archivistici, intesi come forme autorizzate del nome del soggetto produttore combinate con altri elementi informativi che descrivono l'entità e la relazionano con altri record.

Gli elementi di descrizione del record di autorità archivistico sono raggruppati in quattro aree di interesse:

- *area dell'identificazione* (5.1 ISAAR): comprende tutti i dati necessari a definire l'entità e fornisce le chiavi di accesso normalizzate principali al record, tra cui la forma autorizzata del nome, le forme parallele e le forme normalizzate appartenenti ad altri standard. Da qui si dipanano le relazioni principali, quelle cioè tra denominazioni d'autorità, specularmente a quanto avviene nei file d'autorità bibliografici.
- *area della descrizione* (5.2): qui vengono fornite le informazioni rilevanti sulla natura, il contesto e le attività dell'entità descritta, includendo date di esistenza, storia, luoghi rilevanti, funzioni e struttura, contesto generale. Questi elementi, frutto del lavoro a monte di ordinamento e di ricerca sulle carte del soggetto produttore, sono l'elemento distintivo dell'archivistica rispetto alle altre scienze umanistiche, essendo *access point* supplementari al solo nome e già risorsa informativa. All'interno di questa sezione si aprono potenziali collegamenti con altre risorse (che verranno esaminati in seguito), quali ad esempio i legami con le funzioni, la cui definizione però afferisce a ISDF.
- *area delle relazioni* (5.3): vengono indicate e descritte le relazioni con altri enti, persone o famiglie (identificati da codici o nomi normalizzati), corredate da date. Le tipologie di relazioni possibili sono state normalizzate: gerarchica, cronologica, familiare, associativa.
- *area di controllo* (5.4): è l'area che definisce univocamente il record d'autorità, tramite un codice identificativo per il record e un codice per l'agenzia responsabile dell'elaborazione del record.

Esiste infine una zona dedicata ai collegamenti tra il soggetto produttore e la sua documentazione archivistica e altre risorse collegate (6): questa area descrittiva è, insieme alla scelta della separazione della descrizione del soggetto produttore

dal contenuto delle sue risorse e alla possibilità di avere intestazioni plurime del soggetto produttore, la fonte principale di innovazioni nello standard.

3.1.2 *Le relazioni in ISDF*

Come ISAAR(CPF), questo standard ha lo scopo di fornire informazioni normalizzate in elementi descrittivi che vadano ad arricchire il contesto di produzione dei record archivistici.

L'importanza delle funzioni come punto d'accesso ad un'istituzione produttrice d'archivi si spiega sia per i suoi risvolti organizzativi, essendo le funzioni più stabili nel tempo rispetto alle strutture amministrative e più determinanti nel descrivere e classificare i record archivistici, sia per i risvolti nella ricerca dell'utente finale, che nell'ambito dei suoi studi fa riferimento più alle attività di un'organizzazione che non alla sua struttura interna.

Lo scopo che muove una gestione separata della descrizione delle funzioni dalle rispettive descrizioni della documentazione e dei soggetti produttori è quello di, come già detto, fornire chiavi di accesso sempre più precise, ma anche più flessibili dal punto di vista dell'utenza. La descrizione è organizzata in aree, specularmente agli altri standard:

- *area dell'identificazione* (5.1): viene indicata la tipologia di funzione e la forma autorizzata del nome unitamente alle altre forme parallele. Le forme normalizzate di nomi di funzioni insieme alla tipologia (*subfunction, business process, activity, task, transaction...*) diventano access points per la ricerca.
- *area del contesto* (5.2): fornisce un quadro generale delle condizioni in cui viene svolta la funzione e informazioni sulla sua natura, incluse le date di riferimento, descrizione della funzione, storia e norme giuridiche annesse.
- *area delle relazioni* (5.3): qui vengono collegate le funzioni descritte alle altre funzioni o attività che adempiono o svolgono le strutture amministrative a cui si riferiscono. Le relazioni tra funzioni vengono categorizzate, di modo da poter creare una classificazione di queste: le relazioni tipizzate sono quelle gerarchiche, temporali e associative. Da prassi, ogni relazione ha un *range* di date di esistenza, elemento fondamentale per la descrizione di istituzioni complesse e longeve che subiscono mutamenti notevoli nel corso del tempo.

Così facendo, è possibile tenere traccia di tutti i cambiamenti intercorsi all'interno della struttura organizzativa di un soggetto produttore consentendo una maggiore comprensione all'utente.

- *area di controllo* (5.4): definisce i criteri univoci utilizzati per la descrizione della funzione, quali un identificativo unico per il nome della funzione e per l'istituzione responsabile della creazione di questo record.

Anche qui è presente in conclusione allo standard una sezione (6) dedicata alle relazioni: tra funzioni e istituzioni, funzioni e materiali archivistici, funzioni e risorse collegate.

In questa parte dello standard emerge chiaramente l'intento che esso si propone: evidenziare ancora più concretamente tramite nuove relazioni il principio di provenienza delle documentazioni e arricchirlo in favore di una migliore interpretazione da parte di utenti "inesperti" della struttura organizzativa di un soggetto produttore, di un archivio o di un ente conservatore, creando un medium interpretativo e informativo del percorso delle carte.

CONCLUSIONI

Questi brevi aspetti introduttivi sono utili ad inquadrare il punto di partenza degli studi che hanno preso piede nell'ultimo decennio nell'ambito interdisciplinare delle scienze umane legate alle scienze informatiche: queste ultime, punto di nuove riflessioni e per buona parte motore dei cambiamenti nell'approccio degli umanisti, stanno modificando il modo stesso di concepire la ricerca storica e le modalità di accesso ad essa.

L'applicazione continuativa e incrementale di supporti informatici nel contesto delle pubbliche amministrazioni non ha potuto non coinvolgere il settore archivistico e ne ha modificato il punto di vista, dato per consolidato, sull'organizzazione della memoria.

Produrre descrizioni archivistiche in ambiente digitale e comunicarle attraverso Internet non è la medesima cosa che farlo con i tradizionali supporti cartacei. Anche in questo, come in altri casi, il mezzo condiziona fortemente la struttura e i contenuti

delle informazioni e spinge a riconsiderare le forme di organizzazione delle conoscenze e le modalità di venirne in possesso¹⁶.

L'alto livello di relazionalità offerto dall'informatizzazione ha segnato solchi profondi nelle stesse concezioni cardine dell'archivistica, quali le ricadute pratiche del metodo storico nella descrizione delle risorse archivistiche o le possibilità offerte dall'*information retrieval*.

L'attitudine con cui l'archivistica si appresta alle novità dell'informatica e le sue caratteristiche scientifiche peculiari già considerate, conducono alla conclusione per cui,

(...) come custodi della testimonianza senza uguali su cui sono basate le biografie e le storie e con l'obbligo professionale di descrivere i soggetti produttori, unicamente gli archivisti sono deputati a giocare un ruolo maggiore nello sviluppo di uno standard per la descrizione del produttore¹⁷.

Da qui la precisa scelta di questo campo umanistico per indagare il sempre più esteso spazio informatico del web di dati, che stimolato dalle continue innovazioni, auspica la creazione di modelli d'eccellenza e linee guida per essere accessibile ad un vasto *range* di utenti e fornire un'informazione di qualità.

16 Stefano Vitali, *Passato digitale. Le fonti dello storico nell'era del computer*, Bruno Mondadori, Milano, 2004, pp. 75-76

17 Daniel V. Pitti, *Descrizione del soggetto produttore. Contesto archivistico codificato*, in Mauro Guerrini e Barbara B. Tillett (a cura di), *Authority Control: definizione ed esperienze internazionali*, Firenze, University Press, 2003, p. 155, http://www.sba.unifi.it/ac/relazioni/pitti_ita.pdf

CAPITOLO SECONDO

Il riorientamento dell'archivistica informatica nella descrizione del soggetto produttore

PREMESSA

In questo capitolo vengono presentate le fasi salienti del percorso che ha portato l'archivistica a confrontarsi con le tecnologie informatiche, dalla nascita del markup e le sue applicazioni da parte della comunità archivistica, sino alle soglie del *semantic web*.

Senza pretese di completezza nella descrizione di questo processo, che ha coinvolto l'intero ambito dei beni culturali e un numero considerevole di attori, si vuole dare risalto a due dei risultati più rappresentativi di questo connubio interdisciplinare: la formalizzazione dello standard rappresentativo EAC (CPF) Schema e la sua recente traduzione in ontologia.

La decisione di questa restrizione di campo è propedeutica alla discussione successiva sulle proposte di arricchimento del modello archivistico ontologico con nuove potenzialità, che ad oggi sono state messe in secondo piano dagli operatori culturali, al fine di concentrarsi unicamente sugli ambiti e i concetti utili alla descrizione archivistica tramite le nuove tecnologie del web of data.

Ciò che qui si vuole dimostrare è come il dibattito e le proposte offerte dall'archivistica informatica siano, benché in perenne *work in progress*, solide fondamenta per una più ampia discussione nel merito della modellizzazione di dati: ovvero, si rivelino utile punto di riferimento per altre comunità culturali che si pongono l'obiettivo di aprire i propri dati al web e cercano una metodologia di

lavoro e di descrizione confacente alla creazione di metadati efficienti e di qualità.

Per questi motivi si è seguito il percorso inerente la descrizione delle entità produttrici di archivio, tralasciando gli sviluppi degli altri standard per l'archivistica, confermandolo come modello utilmente riusabile in altri contesti per collegare risorse e descrizioni di *authority record* e, nel caso in questione, definizioni di *access point* per la ricerca.

1. INTRODUZIONE AI LINGUAGGI DI MARKUP PER L'ARCHIVISTICA

Nel momento in cui le risorse archivistiche vengono rese disponibili in ambiente elettronico è (...) auspicabile che il contenuto informativo risieda in un formato standard, ossia una specifica norma condivisa da una comunità, in grado di favorire la condivisione di oggetti, metodi e processi finalizzati all'interoperabilità degli stessi. In tal modo viene garantita sia la comunicazione tra hardware e software diversi (portabilità interpiattaforma), sia la stabilità e riusabilità delle risorse stesse¹.

Il *World Wide Web Consortium (W3C)*², organizzazione ufficialmente deputata allo sviluppo di standard tecnologici per il web, ha sviluppato nel 1998 un meta-linguaggio descrittivo per la codifica di dati afferenti a documenti sul web, l'*Extensible Markup Language (XML)*³: il fine dichiarato è quello di rendere interoperabili, portabili e accessibili le informazioni logiche di un documento, ossia permettere il reperimento e l'accesso a qualsiasi tipo di risorsa descritta in un ambiente comune (XML appunto), indipendentemente dalla piattaforma tecnologica e senza perdita di informazioni.

Tentativi precedenti di sviluppare uno standard per lo scambio di dati in forma *machine-readable*, cioè comprensibili alla macchina piuttosto che all'uomo, come ad esempio lo *Standard Generalized Markup Language (SGML)*⁴, sono stati soppiantati nel processo di uniformazione, per cui si è optato per l'adozione di unico standard accettato unanimemente dalla comunità informatica.

XML si è infatti imposto come base per la codifica e lo scambio di dati intellegibili all'utente e alle macchine, utilizzato per creare sintassi (*Document Type Definition*) per la descrizione di documenti appartenenti alla stessa tipologia e separare i dati informativi dalla visualizzazione sul web.

Un documento XML si compone, essenzialmente, di tre elementi principali:

1. la DTD (Document Type Definition) o l'XML Schema (opzionale), un documento

.....

1 Cfr. Regesta.exe, "XML", <http://www.regesta.com/la-tecnologia-xdams/xml/>

2 Cfr. World Wide Web Consortium (W3C), <http://www.w3.org/>

3 Cfr. Liam Quin, "Extensible Markup Language (XML)", <http://www.w3.org/XML>

4 Cfr. Dan Connolly, "Standard Generalized Markup Language (SGML)", <http://www.w3.org/MarkUp/SGML/>

che permette di definire ex novo un insieme di tag che rispettino le esigenze di codifica di un documento o di un insieme di documenti specifici. In altre parole si realizza un modello dati che ricalca la struttura logica e le componenti semantiche dell'informazione trattata;

2. l'istanza di documento, ossia uno o più documenti testuali in cui risiede il contenuto informativo, e che, eventualmente, rispetta/no il modello dati implementato;
3. il foglio di stile XSL (XML Stylesheet Language) per la rappresentazione del documento e la sua visualizzazione⁵.

Anche nell'ambito dell'archivistica XML viene utilizzato come formato per lo scambio di dati, al fine di favorire politiche di conservazione e valorizzazione del patrimonio documentario, codificando la descrizione dei record archivistici e associandoli a metadati descrittivi.

La decisione di sottoscrivere l'utilizzo dei linguaggi di *markup* da parte della comunità archivistica deriva dal bisogno comune e crescente all'interno della Pubblica Amministrazione di aderire a regole condivise, a livello nazionale e internazionale, per la gestione corrente dei documenti ed il mantenimento degli archivi storici: gestione che richiede una maggiore razionalizzazione e formalizzazione di procedure per ottenere una conservazione a lungo termine dei documenti, integrandoli e rendendoli interoperabili e quindi condivisibili.

Nel contesto delle procedure per gli archivi storici (ovvero enti che non producono nuova documentazione, tantomeno documentazione nativa digitale), sono stati elaborati dalla *Society of American Archivists (SAA)*⁶ due standard rappresentativi basati sulla tecnologia XML, ovvero due DTD poi divenute XML Schema (XSD)⁷, che traducono gli standard descrittivi ISAD(G) e ISAAR(CPF): *Encoded Archival Description (EAD)*⁸, per la rappresentazione dei contenuti di documentazione, ed *Encoded Archival Context (EAC-CPF)*⁹, per la rappresenta-

.....
5 Ilaria Barbanti, *La codifica informatica dei dati di archivio: EAD ed EAC*, in «Scrinia», 2-3/2005, http://media.regesta.com/dm_o/regesta/regestacms/generic/documentazione/Barbanti-EADEAC2005.pdf, p. 3

6 Cfr. Society of American Archivists (SAA), <http://www2.archivists.org>

7 Cfr. L. Quin, "XML Schema", <http://www.w3.org/standards/xml/schema>

8 Cfr. Library of Congress, "Encoded Archival Description (EAD)", 2006, <http://www.loc.gov/ead/eaddev.html>

9 Cfr. Staatsbibliothek zu Berlin, "Encoded Archival Context – Corporate bodies, Persons, Families (EAC-CPF)", <http://eac.staatsbibliothek-berlin.de/>

zione e descrizione dei soggetti produttori d'archivio.

Essendo uno dei principi basilari della descrizione archivistica la multi-livellarità, ciascuno dei due standard ha l'obiettivo comune di definire porzioni della struttura logica di un inventario archivistico (o comunque di uno strumento di corredo) ed è stato

(...) progettato in raccordo con gli altri così da completarsi, accrescersi e supportarsi reciprocamente, al fine di rendere completamente accessibili in tutti i loro dettagli e tutti i loro aspetti, i fondi conservati¹⁰.

A differenza degli standard per il contenuto dei dati (ISAD(G) e ISAAR(CPF)), questi strumenti non forniscono indicazioni su come esprimere il contenuto informativo dei dati, ma solo la struttura in cui devono essere inseriti; l'approccio di XML è gerarchico, consente di creare un albero di elementi al cui interno ne vengono annidati altri più specifici. Da qui discendono alcuni dei maggiori problemi di formalizzazione, poiché l'approccio della descrizione archivistica è invece pluri-livellare, non solo meramente gerarchico, e codificare ogni elemento necessario a esplicitare la dinamicità del contesto di produzione dei documenti archivistici ha richiesto continue revisioni degli standard al fine di ottenere lo scenario più soddisfacente.

Essendo il focus di questo studio la formalizzazione della descrizione del contesto di produzione di complessi documentari, maggiore attenzione verrà data allo Schema EAC(CPF), dove la dinamicità degli elementi costitutivi offre maggiori spunti di riflessione, mentre per una argomentazione approfondita dello Schema EAD si rimanda alla descrizione della *Tag Library* di EAD Schema 2002¹¹.

2. EAC(CPF): UNO STANDARD RAPPRESENTATIVO PER ISAAR(CPF)

Encoded Archival Context (EAC) si pone in discontinuità con la consolidata tradizione di lavoro incentrata sulla sola descrizione di documenti e risorse. Questa

.....

10 Francesca Ricci (a cura di), *Le linee guida per l'elaborazione e la presentazione di strumenti di ricerca*, in «Rassegna degli Archivi di Stato», anno LXIII – n. 1, 2003, <http://www.icar.beniculturali.it/biblio/pdf/standard/guidelines.pdf>, p. 1

11 Cfr. Library of Congress, EAD Tag Library, <http://www.loc.gov/ead/tglib/index.html>

iniziativa nasce dalla consapevolezza che se descrivere persone, famiglie e organizzazioni produttrici di documentazione è utile alla conservazione stessa dei beni culturali, gestire ed identificare univocamente queste entità è fondamentale per l'accesso alle risorse.

Attraverso collegamenti alle descrizioni dei record, le descrizioni dei produttori possono servire da gateway verso i record stessi. Le descrizioni dei produttori possono anche fungere da risorsa indipendente per gli utenti che cercano informazioni su individui, famiglie e organizzazioni. Importante quanto questi benefici, forse più eccitante per gli archivisti, è l'opportunità presentata dalle tecnologie digitali di descrivere e controllare registrazioni archivistiche in maniera più efficace e accurata di quanto sia possibile con il mezzo a stampa¹².

Sino all'avvento dei mezzi informatici la mole di informazioni inerenti documentazione e soggetti produttori veniva gestita contemporaneamente all'interno dello stesso strumento di corredo, precisamente nelle introduzioni agli inventari; adesso è possibile mantenere e incrementare questo lavoro semplicemente separando e collegando queste informazioni complesse, evitando dispendi di energie, ripetizioni finalizzate alla comprensione, risparmiando risorse economiche e fornendo un mezzo informativo più efficiente e dinamico.

Il filo conduttore che muove l'archivistica informatica in questa scelta non è tanto mettere a disposizione sul web "copie anastatiche" degli strumenti di corredo cartacei, ma far interagire i singoli elementi concettuali legati alla descrizione del soggetto produttore, che danno vita a complessi network di relazioni all'interno di sistemi descrittivi flessibili. Le relazioni che intercorrono tra registrazioni, soggetti produttori, funzioni e attività sono dinamiche e multilivellari e richiedono un'attitudine altrettanto dinamica nel connettere questi nodi descrittivi, di modo da renderli "parti di un tutto" ma tendenzialmente autonome, al fine di arricchirsi vicendevolmente e offrire punti di accesso più elastici. Per le peculiarità della struttura di EAC, il modello descrittivo è riproponibile (riusabile) ad altre comunità che optano per la normalizzazione di descrizioni di *authority record*: ad oggi è l'unico schema che approfondisce

12 D. V. Pitti, *Descrizione del soggetto produttore. Contesto archivistico codificato*, ivi, p. 157

analiticamente il contesto di produzione di risorse culturali in una piattaforma indipendente da quelle che descrivono i contenuti delle risorse e gli *authority record*. Sulle potenzialità di questo approccio alle descrizioni si apre

(...) un'interessante prospettiva: quella di utilizzare il modello dati EAC non solo per la codifica dei soggetti produttori, ma anche di tutto ciò che, oltre ai soggetti produttori, potrebbe assumere valore di chiave di accesso alle risorse archivistiche (nomi di ente, persona o luogo citati, o descrittori relativi al contenuto).

In altre parole, se il contesto è importante per completare la descrizione archivistica, è altrettanto importante che la suddetta descrizione possa sfruttare le possibilità offerte dalla codifica di altre tipologie di chiavi di accesso (non limitata soltanto all'identificazione della corretta intestazione), e delle relazioni che intercorrono tra loro. Pur tenendo presente il rispetto del principio di provenienza e del contesto di produzione della documentazione d'archivio, grazie ai sistemi di ricerca che è possibile implementare oggi per l'accesso alle fonti, un numero molto maggiore di chiavi d'accesso rese sino ad ora disponibili nei tradizionali strumenti di ricerca potrebbe aggiungersi a quelle relative ai soggetti produttori, anche nell'ottica di poter condividere risorse appartenenti a diverse tipologie.

Dal punto di vista informatico, il trattamento delle informazioni assumerebbe una veste logica differente, secondo cui non è tanto importante il ruolo (soggetto produttore) che l'oggetto descritto assume, bensì la sua natura, dunque la tipologia della voce, in questo caso persona, famiglia o ente, a cui, parlando di chiavi di accesso, si aggiungono le tipologie di soggetto o luogo¹³.

Di fatto, altre comunità hanno indagato sulla formalizzazione del ruolo dei "creatori di *cultural heritage*", come Dublin Core¹⁴, che ne hanno formalizzato gli aspetti fondamentali ma non sufficienti a rappresentare la complessità di informazioni delle entità: su queste basi EAC continua a proporsi come il modello potenzialmente più efficiente per la gestione separata di descrizioni di autorità.

.....
13 I. Barbanti, *La codifica informatica dei dati d'archivio*, ivi. p. 13

14 Cfr. ASIS&T, "Dublin Core Metadata Initiative (DCMI)", 1995, <http://dublincore.org/>

2.1 GLI ELEMENTI FONDAMENTALI DI EAC(CPF)¹⁵

Lo Schema EAC(CPF) ha un *document element* denominato *<eac-cpf>* contenente due elementi obbligatori, *<control>* e *<cpfDescription>*, da cui discendono a grappolo tutti gli altri sotto-elementi descrittivi, la maggior parte dei quali sono opzionali.

<control>

Specularmente all'area di controllo di ISAAR(CPF) (5.4), contiene tutti gli elementi necessari ad identificare univocamente il record d'autorità (*<recordId>*) che si va a descrivere e l'agenzia responsabile del record in questione (*<maintenanceAgency>*).

In questa sezione vanno indicate le fasi della creazione del record, il linguaggio e le convenzioni/standard utilizzati, nonché le fonti (*<sources>*) consultate per creare la nota biografica/contestuale e le altre forme autorizzate del nome.

<cpfDescription> (o *<multipleIdentities>*)

Rispetto allo standard descrittivo ISAAR (CPF), più articolato, questo elemento funge da macroarea che accentra sia gli aspetti identificativi, sia quelli descrittivi che relazionali.

Al suo interno si trovano infatti i complessi elementi *<identity>*, che riproduce le informazioni contenute nell'area identificazione di ISAAR (CPF) (5.1), la forma autorizzata del nome e quelle parallele; *<alternativeSet>*, che prevede la possibilità di collegare alla descrizione EAC altre descrizioni della stessa entità appartenenti a diversi authority systems; *<description>* e *<relations>*, per la descrizione dell'entità e delle relazioni che intercorrono tra esso e risorse collegate, authorities e funzioni.

Nel caso in cui ad una entità corrispondano più identità viene utilizzato l'elemento *<multipleIdentities>* o vengono correlate più istanze di EAC.

.....
¹⁵ Cfr. Staatsbibliothek zu Berlin, "EAC-CPF Tag Library", 2010,
<http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html>

<description>

Rinvio della corrispondente area descrizione di ISAAR (CPF) (5.2), definito nello Schema EAC come sotto-elemento di *<cpfDescription>*, contiene sia elementi formalizzati con forme controllate del nome (*<existDates>*, *<place>*, *<legalStatus>*) sia descrizioni testuali inerenti la natura del soggetto produttore (*<place>*, *<function>*, *<occupation>*, *<mandate>*, *<structureOrGenealogy>*, *<generalContext>* e *<biogHist>*).

2.2 LE RELAZIONI IN EAC(CPF)

Come già osservato analizzando ISAAR (CPF), le relazioni tipizzate dallo standard per la descrizione del soggetto produttore hanno come nodi fondamentali: soggetti produttori – funzioni – risorse.

L'elemento *<relations>* di EAC, differentemente dalla corrispondente area delle relazioni di ISAAR(CPF) (5.3), si presenta come contenitore delle relazioni intercorribili tra il soggetto produttore e qualsiasi altro tipo di risorsa collegata.

In ISAAR (CPF) infatti, le relazioni sono dislocate in più aree descrittive: nell'area relazioni (5.3) è possibile collegare *authorities* fra di loro; nell'area descrizione (5.2) è prevista l'enucleazione delle funzioni e delle attività svolte dall'entità; nell'ultima sezione (6) sono previsti i link a risorse esterne collegate e alla documentazione prodotta dall'entità.

Astrattamente in EAC-CPF Schema tutte le relazioni vengono equiparate e raccolte in un unico elemento a grappolo che le definisce con i sotto-elementi *<cpfRelation>*, *<functionRelation>*, *<resourceRelation>*.

<cpfRelation>

Il soggetto produttore descritto può essere relazionato a persone, famiglie o istituzioni, normalmente (ma non obbligatoriamente) descritte in un'altra istanza EAC.

Le relazioni tra *authorities* vengono qui descritte fornendo una forma controllata del nome dell'entità collegata, le date e i luoghi della relazione e una nota discorsiva che esplicita la natura del legame. L'elemento include l'attributo opzionale *@cpfRelationType* per tipizzare la relazione con possibili valori

controllati: *identity, hierarchical, hierarchical-parent, hierarchical-child, temporal, temporal-earlier, temporal-later, family, associative*.

<functionRelation>

L'elemento viene utilizzato per normalizzare i tipi di funzioni collegati ad un soggetto produttore all'interno di un *archival descriptive system*, ossia in un sistema che prevede una descrizione separata delle funzioni (sulla base delle regole ISDF) collegate ai record EAC. Come per <cpfRelation>, possono essere indicate date e luoghi rilevanti per identificare la relazione e si utilizzano come valori opzionali dell'attributo @functionRelationType: *controls, owns, performs*.

<resourceRelation>

Le risorse che tipicamente vengono correlate ad un soggetto produttore spaziano tra collezioni, raccolte documentali, risorse bibliografiche e oggetti museali: lo standard non definisce univocamente quali tipologie di risorse sono contemplate, lasciando perciò aperte tutte le possibilità di contenuti e supporti di tali risorse. L'attributo @resourceRelationType che definisce il legame tra l'entità e la risorsa collegata ha come valori plausibili: *creatorOf, subjectOf, other*.

2.2.1 <functionRelation> e <resourceRelation>

Facendo un confronto diretto tra EAC(CPF) e ISAAR(CPF) nel merito delle relazioni tipizzate, vanno fatte qui alcune considerazioni preliminari su due punti che sono stati gli input per questa ricerca, ovvero le definizioni di <functionRelation> e <resourceRelation>.

Riprendendo il documento ufficiale che descrive gli elementi di EAC, alla voce <function> si legge:

A single <function> element may be encoded directly within <description>.

Alternatively, multiple <function> elements may be grouped within a <functions> wrapper that facilitates manipulating them as a group.

Functions may alternatively be described in discursive form in <biogHist> (ISAAR (CPF) History 5.2.2).

The <function> element should be used whenever separate semantic processing of information about functions is required.

As a further alternative, descriptions of functions may form discrete components in an archival descriptive system. Such descriptions should be compiled in accordance with the International Standard for Describing Functions (ISDF) and will typically be described in another encoding language. In such a system, use <functionRelation> to point from the EAC-CPF entity to the related function description. In these systems, the <function> element will probably become redundant¹⁶.

Questo elemento, che in ISAAR(CPF) apparteneva all'area descrizione, dopo la formalizzazione dello standard ISDF per la descrizione delle funzioni e la definizione dell'XML Schema di EAC, non ha ancora trovato una coerente applicazione. Come spiegano le indicazioni dello schema, le funzioni, a seconda del livello di complessità che bisogna rappresentare, possono essere descritte con metodi alternativi che variano da:

- la semplice descrizione testuale all'interno della biografia/nota storica senza ricorso all'elemento <function>;
- l'utilizzo dell'elemento a grappolo <functions> all'interno di <description>, in cui vengono enucleati una serie di sotto-elementi <function> per ciascuna attività o funzione da descrivere;
- la forma più complessa, ovvero l'utilizzo di <functionRelation> che è ascritta alla esistenza di un *archival descriptive system* e prevede descrizioni separate delle singole funzioni codificate in altri documenti e relazionati con l'istanza EAC.

Quest'ultima possibilità, che sarebbe la più auspicabile, è in realtà la meno formalizzata, essendo lo standard ISDF ancora allo status di documento di raccomandazioni per la preparazione di descrizioni e non definito in uno Schema XML integrato in EAC¹⁷. Ciò significa che per quanto vengano fornite delle buone norme per la descrizione delle funzioni e delle attività, queste faticano a farsi posto tra gli *access point* normalizzati della ricerca.

.....

¹⁶ EAC-CPF Tag Library, ibidem

¹⁷ È in corso in Italia il dibattito per la realizzazione di uno schema EAC-F(unctions), che per ora vede solo come effettivo il progetto personale di Salvatore Vassallo, reperibile su <https://github.com/svassallo/EAC-F/blob/master/eac-f.rng>

Per quanto concerne invece *<resourceRelation>* le tipizzazioni offerte dallo standard si limitano a forme controllate sulla paternità del soggetto nei confronti della risorsa (creatore, autore, custode, proprietario, controllore, detentore di copyright) e su un generico richiamo associativo ad esse (soggetto di).

Non sono contemplate per esempio le relazioni con risorse correlate quali le fonti storiche utilizzate per la descrizione del soggetto produttore, che vengono invece richiamate nell'elemento *<control>* tramite l'elemento *<sources>*: è ovviamente possibile considerarle come risorse collegate, ma il generico *subjectOf* non dà rilievo al ruolo che il soggetto produttore ha nei confronti della risorsa e di quali informazioni sono state estratte da queste risorse ai fini della sua descrizione. Non si dice insomma quale ruolo o funzione dell'*authority* viene attestata in una risorsa, né in quale veste il soggetto è richiamato nella risorsa stessa. Per ogni funzione svolta da un soggetto dovrebbe esservi invece un contesto che la definisce univocamente, come un documento, una fonte; e per ogni risorsa che è relazionata ad un soggetto dovrebbe esservi un *range* maggiore di forme controllate per definire le relazioni che intercorrono tra queste due entità, magari estrapolate dal contenuto della risorsa stessa, invece che desunte dai soli rapporti di paternità. Nonostante quindi l'esigenza di standardizzare le funzioni/ruoli di un soggetto produttore in uno standard apposito, questi nodi fondamentali della descrizione del contesto archivistico restano ancora in un *work in progress* di definizione.

L'intento è esplorare ulteriormente le possibilità offerte dalla standardizzazione delle relazioni in EAC, per estrapolare il massimo di informazioni dalle relazioni stesse, descrivendole in una forma normalizzata e confacente alle necessità della ricerca, il tutto alla luce delle innovazioni offerte dalle nuove tecnologie del *semantic web*, che ha fornito mezzi tecnologici e concettuali più elastici ed espressivi a tal fine.

3. IL SEMANTIC WEB E LE APPLICAZIONI NELL'ARCHIVISTICA

Il dibattito sulla standardizzazione della descrizione archivistica, intesa come strumento di comunicazione formalizzata, ha fatto sorgere nuove problematiche inerenti i linguaggi da adottare e le tecniche di rappresentazione delle in-

formazioni; dall'esigenza di adottare *software* più elastici ai complessi scenari all'utilizzo di strumenti che amplino il potenziale espressivo della descrizione, il *semantic web*¹⁸ si è imposto nell'ultimo decennio come luogo più proficuo dove implementare sistemi informativi adeguati alle necessità di comunicazione comuni alle discipline umanistiche.

(...) iniziativa che si propone di promuovere l'interoperabilità nella rete internet, trasformandola da rete di documenti a rete di dati, inseriti in un contesto e arricchiti di semantica, ed elaborabili dalle macchine. Il Web può essere cioè definito Semantico quando i computer diventano in grado di compiere operazioni logiche su dati di natura differente collegati tra di loro, così da inferire nuova conoscenza a partire da quella nota. Per far ciò è necessario strutturare l'informazione in maniera tale che i documenti non siano "isole di dati" ma piuttosto "basi di dati aperti" da cui un programma sia in grado di attingere automaticamente informazioni¹⁹.

Questa iniziativa del W3C, che propone un *framework* in cui definire i metadati nei loro aspetti sintattici, strutturali e semantici, adotta come linguaggi di riferimento quelli del *markup* dichiarativo (XML) e tecnologie per la specifica del contesto semantico: *Resource Description Framework (RDF)*²⁰ e *Ontology Web Language (OWL)*²¹ per la modellizzazione dei metadati, *SPARQL*²² per l'interrogazione.

L'evoluzione del web auspicata dal fondatore del progetto di web semantico, Tim Berners-Lee, prevede un'estensione dell'attuale web di documenti in un web di dati multilivellare, o meglio di metadati che descrivano il contenuto dei documenti presenti in rete, consentendo un'interazione di dati basata su

.....

18 Cfr. W3C, "Semantic Web", <http://www.w3.org/standards/semanticweb>

19 Stella di Fazio, "Web semantico e descrizione archivistica: un applicativo semantico per gli archivi", http://www.san.beniculturali.it/web/san/dettaglio-notizia-san?p_p_id=56_INSTANCE_X7Qi&articleId=2237879&p_p_lifecycle=1&p_p_state=normal&groupId=10704&viewMode=normal

20 Cfr. Brian McBride, Dan Brickley, Eric Miller, "Resource Description Framework (RDF)", <http://www.w3.org/RDF/>

21 Cfr. Michael K. Smith, Chris Welty, Deborah L. McGuinness, "Web Ontology Language (OWL)", <http://www.w3.org/OWL/>

22 Cfr. Eric Prud'hommeaux, Andy Seaborne, "Protocol and RDF Query Language (SPARQL)", <http://www.w3.org/TR/rdf-sparql-query/>

un'organizzazione formale di questi e la condivisione tra comunità che decidono di aprire e connettere i propri dati, *Linked Data*²³, al network.

Con linked data, in altre parole, ci si riferisce a dati pubblicati sul web in una modalità *leggibile e interpretabile da una macchina*, il cui significato sia esplicitamente definito tramite una stringa costituita da parole e marcatori. Si costruisce così un reticolo di *dati collegati* (linked data, appunto) appartenenti a un *dominio* (che costituisce il contesto di partenza), collegato a sua volta ad altri set di dati esterni, ovvero fuori dal dominio, in un contesto di relazioni sempre più estese²⁴.

Alla base dei *linked data* vi sono *Universal Resource Identifier* (URI) deferenziabili²⁵, una sintassi utilizzata per definire nomi e indirizzi di risorse sul web.

L'utilizzo combinato di queste tecnologie semantiche (URI, XML, RDF, OWL) permette di rappresentare un dominio di conoscenza, cioè un ambito conoscitivo definito, basandosi sulla sua concettualizzazione e su una rappresentazione ontologica di questa astrazione, che sia egualmente comprensibile all'uomo e alla macchina: i riflessi più vistosi di questo approccio innovativo si evincono da una maggiore possibilità di condivisione di informazioni strutturate e qualitativamente connotate e da una maggiore potenzialità di precisione nel richiamo di dati nella ricerca. Dalla nascita del *semantic web*, infatti la mole di dati condivisi è aumentata esponenzialmente, spesso a discapito però della qualità delle informazioni. Come spiegano Mauro Guerrino e Tiziana Possemato

La crescita e l'uso esponenziale dell'informazione disponibile in rete non coincide nemmeno con la crescita di fiducia nelle notizie: il grado di loro affidabilità è basso. Gli utenti devono selezionare tra un mare di informazioni restituite per arrivare a una notizia attendibile. Sulla base di quale criterio scegliere? L'autorevolezza della fonte diventa l'elemento discriminante, la selezione avviene a monte, preferendo scegliere la risorsa sulla base dell'autorevolezza di chi l'ha creata, anziché a valle,

.....
23 Cfr. Tim Berners-Lee, "Linked Data", <http://www.w3.org/DesignIssues/LinkedData.html>

24 Mauro Guerrini, Tiziana Possemato, "Linked data: un nuovo alfabeto del web semantico", p. 7, <http://www.bibliotecheoggi.it/content/201200300701.pdf>

25 Cfr. Universal Resource Identifiers (URI), <http://www.w3.org/TR/uri-clarification/>

scegliendo acriticamente sulla base del *ranking* delle notizie che appaiono sulla pagina. La qualità della fonte, la certezza della provenienza diventano, dunque, elementi determinanti nel percorso esplorativo del ricercatore²⁶.

In un simile contesto la qualità delle informazioni da immettere in rete diventa una caratteristica fortemente ricercata da chi si propone di incrementare un dominio del sapere. In forza di ciò, i dati dell'ambiente dei beni culturali (dai musei, alle biblioteche agli archivi, in veste della loro tradizione di creatori di dati fortemente strutturati, attenti alla qualità delle informazioni e di alto livello di fiducia) sono fortemente ambiti dagli studi sull'integrazione in rete di informazioni significative e qualitativamente evolute; d'altro canto è lo stesso settore del *cultural heritage* ad essere interessato a dare maggiore rilievo e visibilità ai propri dati in un contesto facilmente accessibile agli utenti.

Nel caso specifico delle descrizioni archivistiche, come spiega Salvatore Vassallo

(...) l'utilizzo degli strumenti e delle tecnologie del semantic web potrebbe consentire una maggior integrazione dei dati e delle descrizioni archivistiche e dei servizi collegati, in particolar modo nel caso delle pubblicazioni di fonti in rete (...) è possibile dimostrare come sia attuabile una gestione delle descrizioni archivistiche direttamente nel web semantico con una sufficiente granularità. Una volta appurato che le descrizioni archivistiche possono essere gestite direttamente nei termini del web semantico, occorre interrogarsi su quali strumenti sia possibile costruire su queste basi e se un simile approccio comporti benefici in termini di gestione, ricerca e presentazione dei dati²⁷.

.....
26 M. Guerrini, T. Possemato, "Linked data", *ivi*, p. 10

27 Salvatore Vassallo, *Descrizioni archivistiche e web semantico: un connubio possibile?*, in «JLIS», 1, n.1, 2010, p. 186, <http://leo.cilea.it/index.php/jlis/article/view/27/37>

3.1 UNO SGUARDO SULLE TECNOLOGIE PER IL SEMANTIC WEB

3.1.1 *Resource Description Framework (RDF) e RDF(S)*²⁸

RDF è lo strumento proposto dal W3C per la codifica, lo scambio e il riuso di metadati strutturati: ogni “oggetto” descritto in *RDF* è detto risorsa e può essere identificato da un URI, modellizzato in un sistema che gli attribuisce delle proprietà a cui associa dei valori (anche questi definiti da URI) e quindi relazionato con altri oggetti. Il data model *RDF* fornisce le regole per formare gli statements (o triple *RDF*: soggetto – predicato – oggetto, ovvero risorsa – proprietà – valore) ma non dice come dichiarare e specificare queste proprietà, né descrive la semantica (fornisce solo una base per poterla esprimere), compito assolto da *RDF Schema (RDF(S))* che definisce specifici vocabolari (insiemi di proprietà semantiche) per i metadati.

In *RDF(S)* ogni statement è relazionato con altri statements e va così a formalizzare proprietà comuni di un concetto, creare classi di risorse simili e gerarchizzare sia classi che proprietà, definendone ove necessario vincoli sul dominio e sui valori. Nonostante le ampie possibilità offerte da *RDF(S)* nel descrivere formalmente un dominio di conoscenza, gli sono preclusi alcuni mezzi espressivi di notevole importanza, utili quando si vuole rappresentare un contesto complesso: la limitata possibilità di inferenza (cioè di ricavare nuova conoscenza da conoscenza già formalizzata) ha infatti posto la necessità di incrementare le capacità di arricchimento semantico.

3.1.2 *Web Ontology Language (OWL)*²⁹ e *Linked Open Data (LOD)*

Lo scopo preposto ad *OWL* è il medesimo di *RDF(S)*, ossia fornire un vocabolario (ontologia) per definire classi, proprietà e relazioni che descrivano un dominio di conoscenza: per farlo *OWL* riusa buona parte del vocabolario di *RDF(S)* implementandolo con nuove proprietà semanticamente più espressive.

.....
28 Cfr. Resource Description Framework (*RDF*), <http://www.w3.org/RDF/> e Resource Description Framework Schema (*RDF(S)*), <http://www.w3.org/TR/rdf-schema/>

29 Cfr. Web Ontology Language (*OWL*), <http://www.w3.org/OWL/>

Come in RDF, è possibile circoscrivere le risorse in classi che raggruppano individui con caratteristiche simili, relazionare le risorse tramite *object property* (ovvero proprietà che collegano individui tra classi) o *data property* (che relazionano un individuo con un data value). In aggiunta a RDF(S), è possibile indicare vincoli di cardinalità e transitività sulle proprietà, eguagliare classi e individui e creare combinazioni tra classi (unioni, complementarità, disgiunzione). Questo insieme di caratteristiche permette di collegare *dataset* descritti da comunità diverse tra loro, ovvero creare di *Linked Open Data* (LOD), evitando ridondanze e ripetizioni di informazioni, allineandole semanticamente.

(...) la logica del link è quella di rompere l'autoreferenzialità dei dati, moltiplicando le relazioni con altre fonti di dati che, ad esempio, forniscano informazioni di contesto circa l'identità di una persona o il luogo in cui egli vive. Inoltre, il fatto che essi puntino a URI differenti, per riferirsi a una medesima "cosa" del mondo reale o a un medesimo concetto astratto, consente di documentare e di esprimere la polisemia e la pluralità di punti di vista che intorno ad essi esistono³⁰.

Per fare un esempio pratico delle possibilità offerte dalle reti di ontologie e LOD:

La creazione di LOD richiama il concetto di rete di ontologie, che permette di muoversi da un dataset concettuale all'altro, scoprendo i nuovi universi di contenuti collegati a dataset di partenza. Supponiamo, per esempio, di avere, da una parte, un Archivio Istituzionale che pubblica i dati aperti relativi ai lavori scientifici del proprio staff accademico; dall'altra, la piattaforma con i curriculum dei medesimi autori e, in terzo luogo, l'anagrafe dei prodotti della ricerca di Ateneo che valuta l'efficienza e l'efficacia delle attività scientifica dell'istituzione. Collegare (linkare) questi dataset potrebbe essere di grande utilità, sia ai produttori dei lavori scientifici che al Nucleo di valutazione di Ateneo e agli utenti. Si potrebbe ottenere un dataset comune facilmente navigabile, da cui si potrebbero ottenere interessanti "viste". Ad esempio, tale dataset comune potrebbe: offrire un servizio personalizzato per gli autori depositanti il materiale nell'archivio (permettendo l'aggiornamento

.....
30 Gianfranco Crupi, *Oltre le colonne d'Ercole: Linked Data e cultural heritage*, in «JLIS», 4, n.1, 2013, p. 9, <http://leo.cilea.it/index.php/jlis/article/view/8587/7888>

automatico dei loro curriculum in base alle pubblicazioni e altre attività di ricerca inserite nell'archivio); contribuire ad un miglior censimento delle competenze di tali autori da parte del Nucleo, nonché il censimento dei nuovi risultati scientifici ottenuti in base all'aggiornamento dei curriculum professionali (partecipazione ai seminari, convegni, conferenze; specializzazioni professionali ecc.). Inoltre si potrebbe creare servizio di valore aggiunto nell'Archivio istituzionale stesso, in cui gli utenti potranno consultare in modo aperto non solo le pubblicazioni dei vari autori, ma anche visionare le schede complete dei loro curriculum, evidenziando i loro percorsi professionali, interessi di ricerca ecc³¹.

3.2 L'IMPORTANZA DELLA MODELLIZZAZIONE PER LA RAPPRESENTAZIONE DI UN DOMINIO DI CONOSCENZA

Il concetto di ontologia viene introdotto negli anni '90 in ambito informatico nel contesto dell'Intelligenza Artificiale per creare rappresentazioni formali di domini di conoscenza appartenenti a comunità ristrette, per condividerle e, ove possibile, riusarle in altri contesti. Gli sviluppi di queste tecnologie, che sono l'avanguardia nell'evoluzione della comunicazione della conoscenza umana, hanno potenzialità rivoluzionarie, con ricadute immediate sull'approccio degli utenti alla rete e sul reperimento di informazioni sempre più pertinenti alle loro ricerche³².

Like Gutenberg, the developers of this new technology have no way of envisioning the ultimate ramifications of their work. They are, however, united by the conviction that creating the ability to capture knowledge in machine understandable form, to publish knowledge online, to develop agents that can integrate that knowledge and reason about it, and communicate the results both to people and to other agents, will do nothing short of revolutionize the way people disseminate and utilize information³³.

.....
31 Iryna Solodovnik, "Comunicazione e ricerca semantica di contenuti informativi: tra metadati, Linked open Data e Ontologie", p. 6, <http://eprints.rclis.org/15966/1/Metad-LOD-Ontologie.pdf>

32 Cfr. John Davies, Rudi Studer, Paul Varren, *Semantic Web Technologies: Trends and Research in Ontology-based Systems*, Wiley Ed., 2006

33 J. Davies, *Semantic Web Technologies*, ivi, p. XII

Insomma, invece di effettuare ricerche *on line* che restituiscono semplicemente liste di risultati non sempre pertinenti, si possono ottenere gruppi di informazioni indicizzati per significato (che per l'uomo è un significato permeato dal linguaggio naturale, per la macchina è un dato etichettato con un linguaggio intellegibile) o concentrare informazioni provenienti dal maggior numero di documenti possibili evitando le ridondanze. Alla base di questi meccanismi, come già accennato, vi è l'applicazione dei Linked Data, che come spiega Gianfranco Crupi

(...) si configura dunque come un'applicazione dei principi del web finalizzata a un nuovo e più flessibile paradigma editoriale dei dati. Il risultato è uno spazio di dati globale – il web di dati, appunto – basato su standard aperti e costituito da un numero incalcolabile di asserzioni RDF provenienti dalle fonti più disparate e coprono i più svariati argomenti. È qui che risiede la fortuna che la tecnologia dei linked data sta cominciando a riscuotere in ogni settore dell'interazione sociale nel web e, più specificatamente, nell'ambito del cultural heritage e della comunicazione scientifica³⁴.

Le ontologie si sono affermate come ambito privilegiato per la modellizzazione della conoscenza, essendo ad oggi il mezzo più potente ed espressivo per esprimere le infinite relazioni di un dominio, e hanno attirato l'attenzione di numerosi settori culturali interessati a sistematizzare i propri dati all'interno del www in modo da poter essere consultati facilmente dal numero più ampio possibile di utenti.

3.2.1 I benefici del semantic web nell'ambito del cultural heritage³⁵

In generale sono notevoli i benefici derivanti dall'utilizzo di queste tecnologie e l'ambito del *cultural heritage* ha tratto da questa offerta i presupposti per modificare radicalmente il proprio rapporto con gli utenti, avviando un processo di uscita dall'autoreferenzialità dei propri sistemi informativi e mettendosi a disposizione delle esigenze pragmatiche-culturali dei suoi fruitori.

.....
34 G. Crupi, *Oltre le colonne d'Ercole*, ivi. pp. 10-11

35 Cfr. G. Crupi, *Oltre le colonne d'Ercole*, Ibidem

Per un'istituzione culturale, modellare e integrare i propri dati con quelli di altri soggetti fa sì che questi diventino più completi qualitativamente, fornendo una maggiore offerta informativa che può essere riusata anche in altri contesti, diversi da quelli di partenza: ovvero, ponendosi l'obiettivo di descrivere un certo dominio si può inferire nuova conoscenza, che diviene autonoma e ri-contestualizzabile in altri ambiti conoscitivi. Questo procedimento non è dissimile dall'obiettivo che la disciplina archivistica si pone quando decide di separare le descrizioni dei complessi archivistici da quelle dei soggetti produttori, le quali, per la precisione metodologica offerta dalle istituzioni creatrici dei record, possono essere riusate come dati di alto valore qualitativo e strutturale (nonché di fiducia) in ambiti diversi dai sistemi informativi archivistici.

Inoltre, integrando i dati di un ambiente culturale nel *semantic web*, un riflesso conseguente all'arricchimento dei dati è la possibilità di offrire nuovi servizi per gli utenti per navigare questi dati: considerando la flessibilità e le possibilità di esplicazione dei legami tra entità culturali modellizzate con i linguaggi semantici, per un utente poco esperto dell'organizzazione e del lessico specifico delle istituzioni culturali, queste innovazioni semplificano notevolmente l'interpretazione degli strumenti offerti per l'esplorazione delle risorse (il riferimento è ai cataloghi bibliografici e agli strumenti di corredo archivistici, che possono mostrare difficoltà di interpretazione complessiva per i “non addetti ai lavori”).

Di questo percorso si è avvantaggiato notevolmente il settore biblioteconomico, che ha aperto i propri dati e di conseguenza anche i suoi strumenti di descrizione delle risorse al contesto del web, traducendo i suoi schemi bibliografici e di classificazione (*ISBD*³⁶, *FRBR*³⁷, *RDA*³⁸, *LC Classification*³⁹) in *linked data*. Anche il campo archivistico ha intrapreso il percorso di apertura dei record archivistici ai LOD, con alcuni progetti di eccellenza che hanno visto la creazione di ontologie basate sugli standard per la descrizione ISAD, ISAAR e ISDIAH.

.....
36 Cfr. Carlo Bianchini, Giuliano Genetasio, Mauro Guerrini, ..., Patrizia Martini (trad. a cura di), "International Standard Bibliographic Description (ISBD)", 2009, http://www.ifla.org/files/assets/cataloguing/isbd/isbd-cons_2007-it.pdf

37 Cfr. IFLA Study Group, "Functional Requirements for Bibliographic Records (FRBR)", 2009, http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf

38 Cfr. Tom Desley, "Resource Description and Access (RDA)", <http://www.rda-jsc.org/rda.html>

39 Cfr. Library of Congress, "LC Classification", <http://www.loc.gov/catdir/cpsolcco/>

Prima di passare all'analisi delle ricadute pratiche del *semantic web* sulla formalizzazione di schemi e ontologie rappresentativi degli standard archivistici va sottolineata una questione rilevante: la pratica del “riuso” di *data model* o di parti di essi per generare nuove ontologie. Come spiegato precedentemente, una delle fondamentali innovazioni nell'uso di modelli che descrivono con lo stesso linguaggio sia le risorse sia le proprietà che le relazionano, è quella di poter riutilizzare frammenti o interi modelli nel contesto di nuovi modelli, senza doverne creare di ulteriori che sostanzialmente ripropongano le stesse funzionalità, ma semplicemente apportando modifiche dove necessario o allineando gli elementi con rapporti gerarchici o di equivalenza. Questo principio basilare della modellizzazione ontologica è la garanzia dell'immissione in rete di dati che siano sempre implementabili e arricchibili semanticamente, ma non ridondanti.

Il campo del *cultural heritage* ha fatto ampio uso di modelli già implementati, al fine di integrare al meglio i propri metadati, e si è fatto promotore a sua volta di modelli riutilizzabili da altri settori della conoscenza, dando un contributo ad una migliore definizione delle proprietà e delle risorse esistenti e creandone di nuove, ma soprattutto aprendo la propria base di conoscenza al web.

Lo scambio tra scienze umane e scienze dell'informazione nell'ambiente condiviso del *semantic web*, alla luce dei vantaggi reciproci, si configura perciò come vitale per la stessa persistenza del ruolo di mediatori delle istituzioni culturali tra la “cultura” e gli utenti, a cui va dedicata maggiore attenzione, proprio alla luce delle nuove potenzialità comunicative offerte dal web

(...) pena la morte dei cataloghi, abbandonati dagli utenti a favore di altri strumenti di reperimento dell'informazione, come i motori di ricerca⁴⁰.

3.3 I DATA MODEL PER I BENI CULTURALI

Senza voler dare un quadro esaustivo dell'offerta di *data model* riutilizzati dall'intero ambito dei beni culturali (all'interno del quale si potrebbe spaziare dai modelli per la biblioteconomia, per le fonti digitali, i musei, l'arte e l'architettura,

40 M. Guerrini, T. Possemato, *ivi*, p. 11

nonché ovviamente per l'archivistica), vi sono *element set* elaborati da alcune comunità che sono stati di fondamentale importanza per lo sviluppo della ricerca di modelli adeguati alle esigenze delle *digital humanities*.

In questa sede ne verranno presi in considerazione tre: *Dublin Core (DC)*⁴¹, *Europeana Data Model (EDM)*⁴² e *CIDOC-CRM*⁴³, utili a inquadrare gli approcci e le possibilità già in uso nel campo e a porre qualche questione propedeutica alla discussione successiva.

3.3.1 *Dublin Core (DC)*

Lo standard è stato sviluppato a partire dal 1995 per descrivere risorse in formato elettronico: fornendo una base semplice e interoperabile, utilizzabile da qualsiasi operatore culturale (anche non esperto) per indicizzare gli elementi fondamentali di una risorsa, *DC* ha il merito di aver creato una “piattaforma” all'interno della quale numerosissimi settori della ricerca sulle tecnologie del *semantic web* hanno dibattuto e dato un contributo, implementando e arricchendo costantemente lo standard e traendo da esso spunti per ulteriori modellizzazioni.

Le proprietà del suo *element set* sono tra le più riusate e con l'evoluzione semantica offerta dall'implementazione di *DCterms*⁴⁴, l'uso di vocabolari controllati e la creazione di un *Abstract Model*⁴⁵, le possibilità descrittive sono notevolmente aumentate. Quando si va a descrivere ambienti più complessi le sole proprietà di *DC* non sono però sufficienti. Ad esempio, *dc:relation* e le sue sotto-proprietà formalizzano relazioni tra una risorsa e un'entità (la quale è responsabile a vario titolo della risorsa descritta) e tra risorse collegate, ma non ammettono mediazioni tra gli eventuali *dcterms:Agent* e le risorse, come potrebbero essere i ruoli svolti da un agente o altri fattori⁴⁶. A loro volta,

.....

41 Cfr. Dublin Core (DC), <http://dublincore.org>

42 Cfr. Europeana Data Model (EDM), <http://pro.europeana.eu/edm-documentation>

43 Cfr. CIDOC-CRM, <http://www.cidoc-crm.org/>

44 Cfr. Dublin Core Terms (DCTerms), <http://dublincore.org/documents/dcmi-terms/>

45 Cfr. Dublin Core Abstract Model (DCAM), <http://dublincore.org/documents/abstract-model/>

46 Cfr. Nigel Ward, "Using the Dublin Core Relation Element", <http://dublincore.org/workshops/dc4/straw/relation1.html>

dcterms:creator e *dcterms.subject* sono proprietà generali che definiscono rapporti tra risorse e entità, senza tipizzare ulteriori sotto-relazioni (eccetto quelle per stabilire la proprietà intellettuale sulla risorsa) che definiscano le tipologie di rapporti intercorrenti tra i due soggetti: ma esistono solo relazioni di paternità diretta tra un *dcterms:Agent* e una *rdf:resource*? Un individuo di *dcterms:Agent* può essere genericamente relazionato con *dc:subject* ad una *rdf:resource*, ma cosa viene attestato in questa risorsa che va a descrivere la natura di *dcterms:Agent*? Come viene descritto?

L'impostazione dello standard è documento-centrica, per cui una descrizione separata del documento dal suo contesto non è tra gli obiettivi che DC si prefissa, né lo è una ricerca finalizzata ad estrapolare dalle risorse collegate informazioni ulteriori che definiscano il soggetto descritto. Per questa serie di motivazioni, le sole proprietà di DC, per quanto pilastri della descrizione di risorse, non sono sufficienti a descrivere un dominio di conoscenza in cui si vogliono attestare relazioni multiple tra entità – risorse – funzioni, quali invece sono alla base della descrizione archivistica.

3.3.2 *Europeana Data Model (EDM)*

Europeana è un network per l'immissione e lo scambio di dati afferenti a collezioni o singoli oggetti culturali, le cui descrizioni sono consultabili tramite un portale di navigazione⁴⁷. Essendo numerosi i settori culturali rappresentati in *Europeana*, (dai musei agli archivi, alle collezioni audiovisive, alle biblioteche) questo adotta altrettanti standard per la descrizione del loro materiale (*EAD* per gli archivi, *Lightweight Information Describing Objects (LIDO)*⁴⁸ per i musei e *Metadata Encoding and Transmission Standard (METS)*⁴⁹ per le librerie digitali), mappandoli e integrandoli con un set di elementi (*ESE elements*⁵⁰) appositamente ideati

.....
47 *Europeana Portal*, <http://europeana.eu/portal/>

48 Cfr. Erin Coburn, Richard Light, Gordon McKenna, Regine Stein, Axel Vitzthum, "Lightweight Information Describing Objects (LIDO)", 2010, www.lido-schema.org/schema/v1.0/lido-v1.0-specification.pdf

49 Cfr. Libray of Congress, "Metadata Encoding & Transmission Standard (METS)", <http://www.loc.gov/standards/mets/>

50 Cfr. Antoine Isaac, "Europeana Data Model Primer", <http://pro.europeana.eu/documents/900548/770bdb58-c60e-4beb-a687-874639312ba5>

all'interno del modello EDM⁵¹ per far confluire agevolmente gli altri modelli in una piattaforma comune. Al fine di una migliore interoperabilità EDM riusa anche altri *data model*, come *Simple Knowledge Organization System (SKOS)*⁵², *DC*, *CIDOC-CRM*, e *Object Reuse and Exchange (ORE)*⁵³.

Come dichiarato nella documentazione relativa al Data Model, gli approcci alle descrizioni degli oggetti culturali sono due⁵⁴:

- l'approccio *object-centric*: al centro della descrizione vi è un oggetto culturale (*edm:providedCHO*) di cui vengono fornite informazioni intrinseche della risorsa (tramite proprietà DC e le specifiche EDM) e le informazioni sui *provider* (proprietà ORE), senza però collegarle ad altri oggetti o eventi. In una battuta, lo stesso approccio del *metadata set* di DC.
- l'approccio *event-centric*: descrive gli eventi che ruotano attorno al *cultural heritage object (CHO)*, ricostruendone la storia e il contesto ed inserendolo in un network di relazioni con altre entità (eventi, agenti, luoghi, archi temporali, concetti). Questo approccio, che convive con quello puramente descrittivo delle risorse, è in realtà meno sviluppato del primo: riusa solo una piccola parte del vocabolario di CIDOC-CRM e utilizza una proprietà cardine, *edm:hasMet* con le sue sotto-proprietà, per collegare persone, istituzioni, risorse in un contesto (evento – luogo – date).

Di fatto, come dichiara lo stesso documento, EDM Data Model non è costruito per descrivere e “catturare” la vasta complessità del contesto, come fa invece CIDOC-CRM: adopera delle scelte progettuali per favorire lo scambio di informazioni su oggetti culturali, fornendo loro un contesto utile a questa finalità, ma senza proporsi di esaminare tutti i possibili scenari relazionali.

Il progetto Europeana, che ad oggi è tra le più grandi *digital library* nel panorama del *cultural heritage* ed è in continua implementazione, ammette delle lacune nella creazione di *access point* per la ricerca: non si discosta dalla prassi di DC, per cui un'entità, ad esempio una persona, può essere collegata

51 Cfr. A. Isaac, *Europeana Data Model Primer*, Ibidem

52 Cfr. Alistair Miles, Sean Bechhofer, "Simple Knowledge Organization System (SKOS)", 2009, <http://www.w3.org/TR/2009/REC-skos-reference-20090818/>

53 Cfr. Carl Lagoze, Herbert Van de Sompel, Pete Johnston, Michael Nelson, Robert Sanderson, Simeon Warner, "Open Archive Initiative – Object Reuse and Exchange (OAI-ORE)", 2008, <http://www.openarchives.org/ore/1.0/vocabulary>

54 A. Isaac, *Europeana Data Model Primer*, ivi, pp. 12-17

ad una risorsa solo se:

- una forma autorizzata del suo nome compare nel titolo dato dal modello alla risorsa (*title*);
- è il creatore della risorsa (*creator*);
- è il soggetto in senso lato della risorsa (*subject*);
- ha una data o un luogo relazionato ad entrambi (*dates, place*).

La possibilità di una mediazione tra le risorse e le altre entità non è stata ancora implementata, come del resto non è prevista una separazione concettuale della “persona” dal suo “ruolo” (*dc:creator, dc:subject* non sottintendono la presenza di un *edm:Agent*⁵⁵): volendo fare un'esempio pratico, è possibile reperire sul portale tutte le risorse collegate a Leonardo Da Vinci solo se il suo nome compare nel titolo di un'opera, se è il soggetto (citato?) o l'autore di quest'opera, o se ha un luogo o una data collegabili ad entrambi, mentre non è possibile effettuare ricerche sulle opere commissionate a lui ma eseguite da uno dei suoi apprendisti bottegai, perché non esiste una *tag* per descrivere questa funzione/ mediazione tra risorse.

Questa potenzialità espressiva non è ancora stata indagata dal modello, per i motivi sopra citati, unitamente alla decisione di non descrivere i creatori o i conservatori degli oggetti culturali.

3.3.3 CIDOC-CRM

The CIDOC CRM is intended to promote a shared understanding of cultural heritage information by providing a common and extensible semantic framework that any cultural heritage information can be mapped to. It is intended to be a common language for domain experts and implementers to formulate requirements for information systems and to serve as a guide for good practice of conceptual modeling. In this way, it can provide the "semantic glue" needed to mediate between different sources of cultural heritage information, such as that published by museums, libraries and archives⁵⁶.

.....
55 Cfr. Silvio Peroni, Francesca Tomasi, Fabio Vitali, *Reflecting on the Europeana Data Model*, in *Digital Libraries and Archives, 8th Italian Research Conference, IRCDL 2012. Bari, Italy, February 2012. Revised Selected Papers*, Berlin-Heidelberg, Springer, 2013, pp. 228-240

56 Cfr. Martin Doerr, Chair, Heraklion, "CIDOC-CRM", <http://www.cidoc-crm.org/index.html>

Il modello vuole essere un punto di riferimento per tutta la comunità del *cultural heritage* per definire e strutturare i dati impliciti ed espliciti, nonché le relazioni, dei beni culturali: consiste di 86 classi e 137 proprietà, ed ha un approccio fortemente evento-centrico, focalizzandosi soprattutto sull'aspetto temporale delle attività e degli eventi, relazionati alla partecipazione di attori, luoghi, date e oggetti in generale. Ogni evento può essere concepito come un incontro tra risorse, che con le specifiche proprietà va a definire il tipo di partecipazione a questo evento. Ad oggi ne fanno largo uso sia la comunità bibliografica che quella museale e nel 2011 si è conclusa l'operazione di mapping tra EAD e CIDOC-CRM⁵⁷. Per la complessità e la mole ingente di elementi, è più semplice riusare singole porzioni del suo vocabolario piuttosto che adottarlo in toto, o quanto meno è pensabile allinearne gli elementi più importanti ai fini della descrizione del dominio di conoscenza in oggetto; è di notevole interesse lo studio sulle definizioni delle potenziali relazioni tra gli oggetti.

Considerare ogni relazione tra risorse come (parte di) un evento può avere i suoi indiscussi vantaggi, ma il modello non prevede istanziazioni o proprietà che ne definiscano la natura: i rapporti biologici, sociali, culturali tra persone non sono tipizzati⁵⁸ con sottoproprietà specifiche e di conseguenza, le proprietà generiche che considerano un ruolo esercitato da una persona nei confronti di un'altra, hanno dei limiti di espressività⁵⁹.

Un eventuale allineamento tra CIDOC-CRM ed EAC-CPF allora, per quanto propedeutico ai fini dell'interoperabilità, non sarebbe di grande supporto nell'estensione delle possibilità relazionali attorno agli *authority record* e alle loro descrizioni con le risorse collegate e le funzioni/ruoli.

3.4 PROGETTI PER L'ARCHIVISTICA

Da molti anni nel settore dell'archivistica si è iniziato a parlare di *semantic web*

.....

57 Cfr. Lina Bountouri, Manolis Gergatsoulis, "Mapping Encode Archival Description in CIDOC-CRM", <http://eprints.rclis.org/15844/1/02.Bountouri.pdf> e M. Doerr, "CIDOC-CRM Mappings, Specialization and Data Example", http://www.cidoc-crm.org/crm_mappings.html

58 Cfr. Christian-Emil Ore, "Social and family relations and the CIDOC-CRM – some thoughts", 2006, http://www.cidoc-crm.org/docs/frbr_oo/frbr_docs/meeting_presentations/10th_meeting_presentations/Fam_relation_CRM.rtf

59 S. Peroni, F. Tomasi, F. Vitali, *Reflecting on the Europeana Data Model*, ivi, p. 237

e di necessità di collegamenti tra risorse afferenti a diversi domini. In questo contesto sono stati diversi i progetti meritevoli di avere posto le basi per un'integrazione dei dati archivistici all'interno della rete: dall'applicazione dello standard EAC in *Social Network and Archival Context Project (SNAC)*⁶⁰, all'applicazione di EAD in *Archives Portal Europe (APENet)*⁶¹ o nel già citato *Europeana*, alle proposte italiane di integrazione degli standard in *Ontologie archivistiche*⁶², all'implementazione di software di descrizione archivistica come *ICA-AtoM*⁶³ e *xDams*⁶⁴, al progetto di implementazione di ontologie *Repository for Linked Open Archival Data (ReLoad)*⁶⁵.

In questa sede non è possibile fornire uno scenario completo del percorso di apertura dei dati archivistici al network offerto dalle sperimentazioni delle metodologie del *semantic web*, per il quale si rimanda alle pagine ufficiali dei progetti in corso basati sull'applicazione degli standard SAA⁶⁶; chiudendo il cerchio delle proposte al quadro nazionale italiano degli ultimi anni e restringendo il campo della ricerca alle applicazioni inerenti allo standard EAC-CPF, il progetto più stimolante per lo sviluppo di ulteriori riflessioni finalizzate all'arricchimento semantico proviene da *ReLoad*, che ha sviluppato un set di ontologie per formalizzare gli standard per la descrizione archivistica.

3.4.1 *ReLoad*

ReLoad nasce dalla collaborazione dell'*Archivio Centrale dello Stato*, dell'*Istituto dei Beni culturali dell'Emilia Romagna* e da *Regesta.exe* per promuovere la raccolta e la condivisione di un set di descrizioni archivistiche dettagliate al fine di renderle disponibili sotto forma di rete di dati archivistici collegati: lo

.....

60 Cfr. The Social Networks and Archival Context Project (SNAC), <http://socialarchive.iath.virginia.edu/>

61 Cfr. Archives Portal Europe (APENet), <http://www.archivesportaleurope.net/web/guest;jsessionid=A4BEE3118A1158AE6BAF8027C67B754>

62 Cfr. *Ontologie archivistiche, analisi e descrizione*, <http://www.archivi.beniculturali.it/index.php/strumenti/sussidi/item/559-ontologie-archivistiche>

63 Cfr. International Council on Archives - Access to Memory (ICA-AtoM), <https://www.ica-atom.org/>

64 Cfr. Regesta.exe, "xDams", <http://www.xdams.org/>

65 Cfr. Repository for Linked Open Archival Data (ReLoad), <http://labs.regesta.com/progettoReload/>

66 Cfr. EAC-CPF in use, Finalized Projects, <http://eac.staatsbibliothek-berlin.de/eac-cpf-in-use/finalized-projects.html>; EAD, <http://www.loc.gov/ead/>

scopo dell'iniziativa è prettamente sperimentale, non intende cioè creare portali per la navigazione della raccolta di dati (benché sia possibile interrogarli) ma valutare le migliori metodologie di lavoro per ottenere il massimo potenziale dall'integrazione di queste collezioni con dati appartenenti a domini esterni.

La prima fase del progetto ha visto la realizzazione di una prima ontologia per la descrizione archivistica, *OAD*⁶⁷, che esplicita formalmente gli elementi informativi necessari alla descrizione di un record archivistico, facendo uso di concetti rappresentativi degli elementi fondamentali di ISAD(G). Per favorire l'interoperabilità l'ontologia riusa le proprietà fondamentali di SKOS, DC e FOAF e allinea alcuni elementi a *dataset* esterni, come GeoNames, DBPedia e VIAF.

Per integrarsi con gli altri standard per l'archivistica, è previsto il collegamento della descrizione dell'unità archivistica rappresentata in OAD con record d'autorità per i soggetti produttori, gli enti conservatori e altre entità, a loro volta rappresentati con i rispettivi vocabolari EAC-CPF e OCSA.

*OCSA*⁶⁸, diversamente da *OAD* e *EAC-CPF*, nasce con l'intento di aprire dati non strettamente archivistici, inerenti all'anagrafica degli enti conservatori di documentazione archivistica (siano essi archivi o meno), al fine di facilitare l'accesso diretto alla documentazione e la pubblicazione delle descrizioni archivistiche in formato *Linked Open Data*.

Nel paragrafo successivo viene descritta l'ontologia *EAC-CPF*⁶⁹, ad oggi pubblicata in versione 1.2, suscettibile di potenziali cambiamenti e implementazioni. Ogni riferimento ad essa rimane perciò parziale, contingente, e vuole offrire spunti per ulteriori riflessioni, partendo dall'analisi delle possibilità offerte da altri domini e dal presupposto che il *data model* *EAC-CPF* possa essere esportato a sua volta come modello di riferimento.

4. EAC-CPF ONTOLOGY

La realizzazione dell'ontologia di *EAC-CPF* muove dalla sperimentazione su

.....

67 Cfr. *OAD Ontology*, <http://labs.regesta.com/progettoReload/oad-ontology/>

68 Cfr. Archivio Centrale dello Stato, Regesta.exe, "*OCSA Ontology*", <http://labs.regesta.com/progettoReload/wp-content/uploads/2013/08/ocsa.html>

69 Cfr. Silvia Mazzini, Francesca Ricci, "*EAC-CPF Ontology*", <http://labs.regesta.com/progettoReload/wp-content/uploads/2013/10/eac-cpf.html>

schede descrittive nella piattaforma IBC-xDams (prodotte secondo la versione beta dello standard EAC-CPF o ad esso migrate) col fine di pubblicarle nel *web of data* come *Linked Open Data*⁷⁰.

La prima versione dell'ontologia si pone fundamentalmente come una rappresentazione dello standard, riprendendo gli elementi della *Tag Library* dello schema e ridefinendoli alla luce della struttura del *data model* RDF.

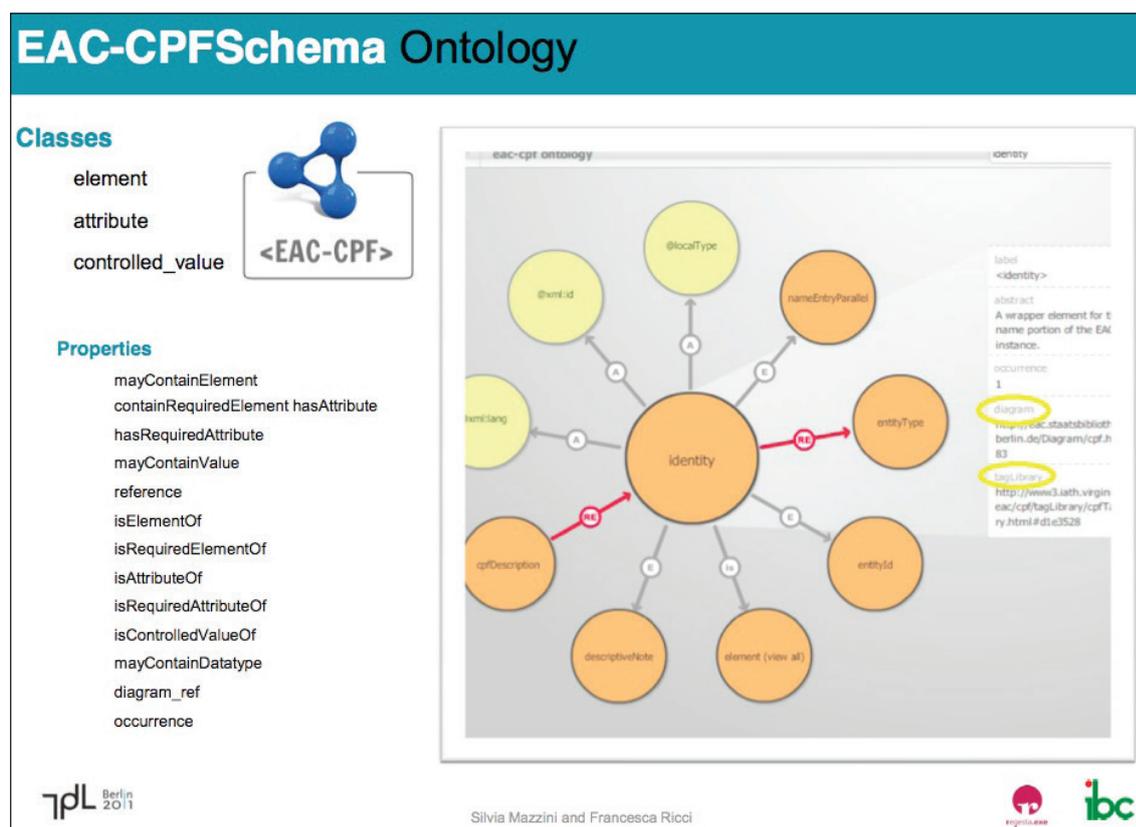


Fig. 2.1 Classi, proprietà e visualizzazione grafica di EAC-CPF Schema Ontology⁷¹

A questa prima formalizzazione, che definisce unicamente la struttura degli elementi informativi dello standard e non ha l'obiettivo di aprire i dati archivistici al web, è seguita una seconda ontologia finalizzata proprio all'apertura delle descrizioni archivistiche di entità per il *Linked Open Data*⁷².

70 Cfr. S. Mazzini, F. Ricci, *Linked archival authority data. Una sperimentazione sui linked open data del sistema informativo regionale IBC Archivi*, in «DigItalia, Rivista del digitale nei beni culturali», VI, 2, 2011, <http://digitalia.sbn.it/article/view/484/334>

71 Cfr. S. Mazzini, F. Ricci, "EAC-CPF Schema Ontology", <http://sda2011.dke-research.de/images/slides05.pdf>

72 Cfr. S. Mazzini, F. Ricci, "EAC-CPF Ontology", *Ibidem*

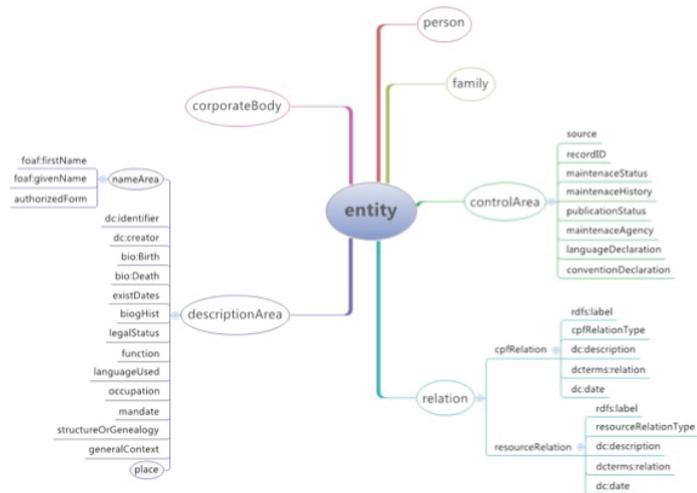
EAC-CPF Ontology for Linked Archival Data

Classes

entity, person, corporateBody,
family, controlArea,
descriptionArea, nameArea,
language, place, relation

Properties

authorizedForm, biogHist, control,
conventionDeclaration, cpfRelation,
cpfRelationType, description,
existDates, function, generalContext,
languageDeclaration, languageUsed,
legalStatus, localTypeDeclaration,
maintenanceAgency,
maintenanceHistory,
maintenanceStatus, mandate,
nameEntry, occupation,
publicationStatus, recordID,
resourceRelation,
resourceRelationType, source,
structureOrGenealogy



Silvia Mazzini and Francesca Ricci



Fig. 2.2 Classi, proprietà e visualizzazione grafica di EAC-CPF Ontology⁷³

Per consentire l'apertura, gli elementi dello schema XSD sono stati analizzati e debitamente trasformati in classi e proprietà, mantenendo gli stessi nomi (*labels*).

Ai fini dell'interoperabilità, questa seconda ontologia prevede il riuso di metadati descrittivi “classici” come quelli di *DC*, *DCTerms*, *FOAF*, *Bio*⁷⁴, *SKOS* e l'utilizzo ove possibile di URI provenienti da altri siti e *authority system*, come *GeoNames*⁷⁵, *VIAF*⁷⁶ e *DBpedia*⁷⁷ (allineati agli URI locali tramite l'uso della proprietà *owl:sameAs*).

Vengono qui riprese solo le classi principali del dominio dell'ontologia, con appunti sulle classi di altri domini e le proprietà correlate. La classe *eac-*

73 Cfr. S. Mazzini, F. Ricci, "EAC-CPF Ontology for Linked Data", <http://sda2011.dke-research.de/images/slides05.pdf>

74 Cfr. Ian Davis, David Galbraith, "BIO: a vocabulary for biographical information", <http://vocab.org/bio/0.1/.html>

75 Cfr. Bernard Vatant, Marc Wick, "GeoNames ontology", <http://www.geonames.org/ontology/documentation.html>

76 Cfr. Virtual International Authority File (VIAF), <http://viaf.org/>

77 Cfr. Christopher Sahnwaldt, "DBpedia", <http://dbpedia.org/About>

cpf:relation e le sue proprietà sono descritte separatamente, per riprendere le considerazioni fatte per la corrispondente area dello Schema EAC (CPF).

4.1 LE CLASSI

L'ontologia ripropone l'assetto dello Schema con le dovute modifiche semantiche: gli elementi cardine dell'XSD vengono formalizzati come classi e i sotto-elementi vengono riproposti come proprietà aventi dominio nelle nuove classi.

4.1.1 *eac-cpf:entity*

La classe definisce la tipologia del soggetto produttore d'archivio e degli altri “attori” a vario titolo relazionati con l'entità descritta: prevede come sottoclassi *eac-cpf:person*, *eac-cpf:family*, *eac-cpf:corporateBody*, rispettando la standardizzazione precedente (5.1.1 ISAAR (CPF)). Ogni individuo di queste classi è un soggetto/entità a cui afferisce una descrizione archivistica formalizzata con l'ausilio delle aree descrittive di ISAAR(CPF) che sono state “trasformate” a loro volta in classi, *eac-cpf:descriptionArea* e *eac-cpf:controlArea*; in EAC-CPF Ontology queste vengono inferite grazie all'uso di un ragionatore come sottoclassi di *eac-cpf:entity*.

Secondo la definizione dell'elemento *<entityType>* della *Tag Library* a cui rimanda la definizione della classe *eac-cpf:entity*, si ha:

Summary: The type of entity being described. Values available are: person, corporateBody, or family.

Description

Within *<identity>* this mandatory element specifies the type of entity being described in the EAC-CPF instance. Values available are: person, corporateBody, or family⁷⁸.

La classe *eac-cpf:entity* vuole raggruppare l'insieme di individui caratterizzati da una comune natura (cioè l'essere un'entità), ma disgiunti dalla specifica della tipologia (persone, famiglie o enti): si potrebbe considerare in futuro una defini-
.....

78 Cfr. EAC-CPF Tag Library, ibidem

zione univoca di *eac-cpf:entity* come unione delle sue sottoclassi *eac-cpf:person*, *eac-cpf:family*, *eac-cpf:corporateBody*, tramite lo statement *owl:equivalentClass* e la restrizione sulla proprietà *owl:unionOf*.

Per quanto riguarda le classi *eac-cpf:descriptionArea* e *eac-cpf:controlArea* invece, dichiarate sottoclassi di *owl:Thing* (classe “madre” di tutte le classi) ma inferite come sottoclassi di *eac-cpf:entity*, formalmente si presenta un'ambiguità: ogni sottoclasse di *eac-cpf:entity* è un insieme di entità con alcune specifiche caratteristiche, quindi *eac-cpf:descriptionArea* o *eac-cpf:controlArea* se dichiarate tali diventano entità *perfettamente equivalenti ai concetti di* persona, famiglia o ente. La traduzione degli elementi definiti nello Schema XML in classi OWL (*owl:Class*) ha implicazioni diverse sulla definizione della natura stessa della risorsa, poiché le classi in un *data model* non sono strutturate gerarchicamente allo stesso modo di un *document element* di un XML Schema. Un prossimo passaggio di revisione dell'ontologia potrebbe riconsiderare la definizione delle classi *eac-cpf:descriptionArea* e *eac-cpf:controlArea*, affinché non siano più definite come sottoclassi né ristrette da un insieme di data property aventi come dominio *eac-cpf:entity*. Ciò consentirebbe di definire *eac-cpf:entity*, *eac-cpf:descriptionArea* e *eac-cpf:controlArea* come classi al medesimo livello.

Un'ultima considerazione va fatta sulla compresenza all'interno dell'ontologia di classi simili: *dcterms:Agent*, *dcterms:AgentClass* e *eac-cpf:entity*. Le prime due vengono mutate da *DCTerms*, mentre *eac-cpf:entity* è “nativa” *eac-cpf:dcterms:AgentClass*, dichiarata come sottoclasse di *rdfs:Class* ha come istanza (non sottoclasse) la classe *dcterms:Agent*. Lo scopo della classe *dcterms:AgentClass* dovrebbe essere quello di definire un gruppo di “agenti” della stessa tipologia, ma che non sono espressamente delle entità. Emerge una poco intuitiva questione terminologica: in cosa differisce un'istanza di “entity” da una di “Agent”? In prima battuta verrebbe da dire che un'entità, termine appartenente alla tradizione archivistica, è un soggetto produttore d'archivio, mentre un agente è un'altra “entità” che non ha lo stesso ruolo di soggetto produttore.

In realtà, ormai da tempo nell'ambito dell'archivistica non si distingue più nella descrizione EAC il soggetto produttore da altri “soggetti non produttori”⁷⁹, poiché le relazioni stesse con le risorse collegate esplicitano questa natura/ruolo;

79 Cfr. Giovanni Michetti, *EAC: Elementi per un Approccio Critico*, in «Archivi & Computer», 1/2008, pp. 40-55

quindi si può dare per assodato che non sia questa la differenza tra le due classi.

La specifica di *dcterms:Agent* allora vuole essere nei confronti di quegli attori culturali che sono gli enti conservatori: tra gli intenti di EAC-CPF Ontology vi è infatti il collegamento della descrizione dei soggetti produttori (e non) alle istituzioni che conservano materiale afferente a loro, nell'ottica di integrare le descrizioni. Non è però lineare il passaggio concettuale, che se è comprensibile quando si fa una distinzione tra questi soggetti sulla base del linguaggio naturale, lo è meno quando si formalizza in elementi ontologici: infatti "Agent", nelle sue varie declinazioni dei più disparati *data model* in circolazione, ha un valore volutamente generico.

A titolo esemplificativo, DCTerms definisce la classe *Agent* come

A resource that acts or has the power to act⁸⁰.

Tramite questo espediente, che restituisce una definizione molto lasca del concetto, la classe *eac-cpf:entity* potrebbe essere considerata o una classe simile (se non equivalente) o una sua sottoclasse.

È una scelta progettuale, quella di EAC-CPF Ontology, con radici nel dibattito sul lessico (e non solo) da adottare per etichettare i concetti pregnanti della descrizione archivistica. In fase di revisione dell'ontologia anche questo aspetto meriterebbe di essere specificato, in modo da essere più intuitivo per chi la utilizza.

4.1.2 *eac-cpf:controlArea*

La classe *controlArea* corrisponde all'area per il controllo di ISAAR(CPF) (5.4) e all'elemento *<control>* dello Schema⁸¹. Come accennato nel paragrafo precedente, la classe è definita come sottoclasse della restrizione sull'insieme di *data property* (necessarie ma non sufficienti a definire univocamente i membri della classe) che hanno come dominio *eac-cpf:entity* ed è relazionata a quest'ultima tramite la proprietà *eac-cpf:control*. Rispetto al suo riferimento allo Schema non

.....
80 Cfr. DCMi Metadata Terms, "dcterms:Agent",
<http://dublincore.org/documents/2012/06/14/dcmi-terms/?v=terms#terms-Agent>

81 Cfr. EAC-CPF Tag Library, ibidem

presenta discrepanze, eccetto per l'esclusione della dichiarazione del linguaggio usato (*eac-cpf:languageDeclaration*) dal set di proprietà che la definiscono, direttamente associata a *eac-cpf:entity*.

Per le considerazioni in merito all'inferenza di sottoclasse di *eac-cpf:entity* si rimanda al paragrafo precedente e a quello successivo su *eac-cpf:descriptionArea*, per la quale valgono i medesimi ragionamenti.

4.1.3 *eac-cpf:descriptionArea*

La classe *eac-cpf:descriptionArea* richiama l'area della descrizione di ISAAR (CPF) (5.2) e l'elemento *<cpfDescription>*⁸² dello Schema.

È codominio della proprietà *eac-cpf:description*, avente come dominio *eac-cpf:entity*, ed è legato a proprietà rappresentative dei sotto-elementi dello schema: la classe è definita da restrizioni su *data property* e *object property* che riproducono gli elementi fondamentali già presenti nello schema (come *eac-cpf:nameEntry*, *eac-cpf:biogHist*, *eac-cpf:existDates*, *eac-cpf:resourceRelation* e *eac-cpf:cpfRelation*) e di alcune proprietà di altri domini (*bio:birth*, *bio:-death*). Gli altri elementi dello Schema che appartengono a *<description>* sono stati anch'essi formalizzati in proprietà⁸³, ma non concorrono qui a restringere la classe *eac-cpf:descriptionArea*.

Occorre fare delle precisazioni: definire una sottoclasse sulla base di una restrizione di proprietà non è una condizione sufficiente per identificare le istanze di questa classe. Infatti ci possono essere istanze di altre classi che sono dominio di quelle proprietà ma non sono istanze di *eac-cpf:descriptionArea*.

In questo caso, le proprietà che restringono la sottoclasse *eac-cpf:descriptionArea* hanno tutte come dominio dichiarato *eac-cpf:entity*: perciò la classe viene inferita come sottoclasse di *eac-cpf:entity* e non può essere definita univocamente sulle restrizioni delle proprietà.

Un apporto ad una migliore definizione degli elementi ontologici potrebbe essere ridefinire il dominio delle proprietà che vanno a restringere la classe, di

82 Cfr. EAC-CPF Tag Library, *ibid*.

83 Le proprietà in EAC-CPF Ontology che si richiamano all'elemento *<description>* di EAC-CPF Schema sono: *dc:identifier*, *dc:creator*, *eac-cpf:legalStatus*, *eac-cpf:function*, *eac-cpf:languageUsed*, *eac-cpf:occupation*, *eac-cpf:mandate*, *eac-cpf:structureOrGenealogy*, *eac-cpf:generalContext*

modo che le tre classi (*eac-cpf:entity*, *eac-cpf:descriptionArea*, *eac-cpf:controlArea*) possano essere allo stesso livello, senza legami di parentela. Uno spunto di riflessione ulteriore per future revisioni dell'ontologia può essere anche la trasformazione delle *data property* in questione in *object property* e la revisione della definizione delle classi, il che sarebbe un ulteriore passo per l'arricchimento semantico dell'ontologia, eventualità che verrà discussa in seguito in merito alla classe *eac-cpf:relation*.

Tornando alle proprietà che restringono *eac-cpf:descriptionArea*, troviamo la già citata *object property* *eac-cpf:nameEntry*, con dominio in *eac-cpf:entity* e codominio in *eac-cpf:nameArea*. La proprietà rappresenta l'elemento *<nameEntry>*⁸⁴ della Tag Library, sotto-elemento di *<identity>*, e la corrispondente area dell'identificazione di ISAAR(CPF) (5.1.2, 5.1.3, 5.1.4, 5.1.5): le congruenze tra le rappresentazioni sono precise, eccetto per la definizione della tipologia dell'entità (che come già visto, è specificata tramite una classe apposita, *eac-cpf:entity*). Questa proprietà concorre a definire l'area della descrizione; per quanto riguarda *nameArea*, di questa non è stata ancora fornita una definizione sulla base dello Schema, ma è identificata da una restrizione delle proprietà *foaf:name*, *foaf:givenName*, *foaf:depiction* e *foaf:page*.

Le proprietà appena citate, nel namespace originario ne specificano il dominio: per *foaf:name*, *foaf:depiction* e *foaf:page* il dominio è un generico *owl:Thing* (sono ammessi individui appartenenti a qualsiasi classe), mentre *foaf:givenName* è specificamente usato per le persone⁸⁵. Nell'ontologia non viene specificato il dominio di queste proprietà, che se fosse formalizzato sarebbe indiscutibilmente *eac-cpf:entity*. *Eac-cpf:nameArea* non è disgiunta da *eac-cpf:entity*: volendo specificare il dominio di *foaf:givenName*, si presenterebbe il problema di inferenza già discusso per le classi *eac-cpf:descriptionArea* e *eac-cpf:controlArea*. Questa è comunque una questione marginale: finché non si specifica il dominio di una proprietà è concesso utilizzarla nel modo ritenuto migliore e non contraddice nessuno *statement*.

.....
84 Cfr. EAC-CPF Tag Library, ibid.

85 Cfr. Dan Brickley, Libby Miller, "FOAF", "foaf:givenName", http://xmlns.com/foaf/spec/#term_givenName

4.2 LE RELAZIONI

L'ontologia prevede la possibilità di collegare risorse appartenenti a domini e ambiti diversi fra loro tramite l'ausilio di *object property*: queste ultime possono relazionare risorse all'interno della stessa ontologia o risorse interne con altre provenienti da altri aggregatori. Rispondendo alle regole già previste da ISAAR (CPF) e EAC (CPF) Schema, anche EAC-CPF Ontology prevede relazioni tra *authority* e risorse collegate, altre *authority* e funzioni.

Nella classe *eac-cpf:descriptionArea* è possibile trovare le relazioni con risorse e entità come vincolanti per la stessa definizione della classe, *eac-cpf:resourceRelation* ed *eac-cpf:cpfRelation*, ovvero le rappresentazioni semantiche dei rispettivi elementi dello Schema *<resourceRelation>*⁸⁶ e *<cpfRelation>*⁸⁷.

Queste due proprietà hanno dominio in *eac-cpf:entity* e codominio in *eac-cpf:relation*, la classe speculare all'elemento *<relations>*⁸⁸, che raggruppa le istanze (gli URI) delle risorse collegate. Ogni istanza di *eac-cpf:relation* è associata ad alcune *data property*: *dc:date*, *dc:description*, *dc:relation* e *eac-cpf:cpfRelationType*, ma non *eac-cpf:resourceRelationType*.

Facendo un confronto diretto tra standard, schema e ontologia si possono notare alcune scelte rappresentative effettuate in sede di modellizzazione.

In ISAAR (CPF) i tre tipi di relazione (*resource*, *function*, *cpf*) sono descritti separatamente: le funzioni sono riportate nell'area descrizione (5.2.5), le relazioni con persone, famiglie ed enti sono nell'apposita area delle relazioni (5.3) ed il collegamento con la documentazione archivistica e altre risorse sono in un'altra sezione dedicata (6).

In EAC (CPF) Schema gli elementi inerenti alle tre tipologie di relazione sono tutti inclusi in un'unico elemento *<relations>*, a sua volta sotto-elemento di *<cpf-Description>*.

In EAC-CPF Ontology si ripresentano i tre corrispondenti elementi (*function-Relation*, *resourceRelation*, *cpfRelation*) sotto forma di proprietà:

- troviamo le *object property* *eac-cpf:resourceRelation* e *eac-cpf:cpfRelation*

.....

86 Cfr. EAC-CPF Tag Library, *ibid.*

87 Cfr. EAC-CPF Tag Library, *ibid.*

88 Cfr. EAC-CPF Tag Library, *ibid.*

come restrizioni sulla classe *eac-cpf:descriptionArea*, rappresentate quindi specularmente allo Schema;

- troviamo solo *eac-cpf:cpfRelationType* come restrizione sulla classe *eac-cpf:relation*. Teoricamente le relazioni contemplate (anche se non vincolate) in questa classe sono solo quelle tra entità, e non quelle tra entità e risorse. Questa impostazione, la separazione del luogo di descrizione delle diverse tipologie di relazioni, era prevista da ISAAR (CPF) ma è stata modificata in EAC (CPF), dove in *<relations>* vengono incluse tutte e tre le relazioni possibili. Per una maggiore precisione formale dovrebbe essere inserita la specifica *eac-cpf:resourceRelationType* (per includere le relazioni con le risorse collegate) come restrizione sulla classe *eac-cpf:relation*;
- per le relazioni con le funzioni, si possono osservare nell'ontologia le stesse difficoltà già riscontrate nello schema XSD. Non esiste, ad oggi, una *object property* “*functionRelation*” che permetta di collegare entità a funzioni e queste a loro volta con risorse prodotte nell'espletamento di quelle funzioni. Esiste invece una *data property* *eac-cpf:function*, con dominio in *eac-cpf:entity*, che consente di descrivere (solo) testualmente (seguendo le linee guida di ISDF) le funzioni esercitate dall'entità, ma non collegate alle risorse inerenti.

In conclusione, le soluzioni adottate da EAC-CPF Ontology meritano un approfondimento ed un ulteriore momento di sperimentazione che consenta di affinare gli strumenti più espressivi a disposizione; le considerazioni fatte fino a qui vanno infatti nella direzione di mantenere aperto il dibattito sull'implementazione di strumenti finalizzati all'apertura dei dati archivistici e alla creazione di nuovi e utili servizi per la ricerca.

I punti “dolenti” che hanno mosso questo studio muovono infatti non tanto dalle difficoltà riscontrate nella definizione del modello ontologico, ma da una mancanza strutturale, che riguarda l'intero ambito di ricerca archivistico, sulla definizione di *access point* rispondenti alle stringenti richieste dei ricercatori: in particolare, verrà affrontata nel prossimo capitolo una riflessione nel merito della (mancata) formalizzazione a pieno potenziale dei ruoli (o funzioni) di entità collegati a risorse archivistiche e non, in particolare le fonti storiche.

CAPITOLO TERZO

Un percorso nella gestione delle relazioni del soggetto produttore: l'approccio al full text

“In complesso, lo storico s'impadronisce
del tipo di fatti che ha deciso di cercare.
La storia è essenzialmente interpretazione.”
Edward H. Carr, *Sei lezioni sulla storia*

PREMESSA

Lo studio del percorso avviato dalla disciplina archivistica verso le scienze dell'informazione ha richiesto l'analisi delle riflessioni (e delle relative problematiche) provenienti dai due differenti ambiti: il tentativo qui proposto di continuare a sperimentare metodologie interdisciplinari e provare a dare una qualche, seppur parziale, risposta ha poi inevitabilmente richiesto di andare alle fondamenta delle discipline, non solo archivistiche e informatiche, che hanno aperto il dibattito teorico. Partendo ovviamente da punti di vista individuali, e quindi parziali, permeati delle reciproche tradizioni.

In questo capitolo sono riportate considerazioni provenienti da campi di studio che ancora oggi dibattono sui punti di tangenza e su compromessi da adottare per fornire un contributo utile all'obiettivo comune, ovvero la comunicazione e condivisione della conoscenza.

Tanti sono infatti gli attori che partecipano a questo dibattito e che hanno voce in capitolo: gli storici, che per comprendere ed interpretare la memoria,

necessitano di una consapevolezza della sua organizzazione; gli archivisti, che sono custodi e primi valorizzatori delle necessità della memoria organizzata; gli informatici, che si offrono come mediatori tra memoria strutturata e nuovi mezzi espressivi; si può dire con certezza ormai, gli archivisti informatici, che contribuiscono concretamente nel decifrare i percorsi comuni possibili; infine, gli utenti, di qualsivoglia livello culturale, e il loro diritto ad usufruire della conoscenza.

Per dare un apporto alla discussione bisogna allora valutare le argomentazioni e le prospettive che si pone ciascuno di essi e darne una lettura aperta, dal punto di vista dell'archivistica informatica.

Innanzitutto va gettato uno sguardo sulle rappresentazioni informatiche progettate per le fonti storiche, sulle possibilità offerte agli studiosi dalla modellizzazione e sui punti di accesso alla ricerca, rispondenti sia alle esigenze pragmatiche degli utenti sia alla fedeltà nei confronti delle fonti stesse. La scelta dei punti di accesso è infatti fondamentale per consentire all'ampio spettro di utenti una riduzione dello scarto di comprensione tra la "loro idea" di ricerca e il "metodo storico" degli archivisti. Seguendo questa articolazione del problema si vuole mettere in evidenza come altri settori del *cultural heritage* abbiano già fornito un approccio alle fonti (intese nel senso ampio di testi rappresentati come oggetti digitali) che ha consentito di estrapolare informazioni utili dalla documentazione senza inficiare l'organizzazione materiale (o meglio, l'ordinamento) dei complessi archivistici/librari per andare a soddisfare esigenze di visibilità e comprensione.

Da qui la decisione di indagare alcuni aspetti che la standardizzazione archivistica ha tralasciato o a cui non si è mai avvicinata: la formalizzazione di punti di accesso per le fonti e i ruoli dei soggetti produttori, estrapolati dal testo pieno dei documenti. La descrizione archivistica infatti non giunge mai ad analizzare il livello del singolo documento, che invece, in ultima analisi interessa al fruitore di strumenti di corredo.

Un terreno che può fungere da anello di raccordo tra la tendenza archivistica nel fermarsi alla descrizione delle serie e quella delle altri settori delle *digital humanities*, che studiano la rappresentazione semantica dei testi in forma digitale, è l'approfondimento sui ruoli politici di persone che hanno depositato la

loro memoria in archivio, desumendo questa conoscenza dallo studio dei singoli documenti. Questa documentazione, ibrida e di difficile “collocazione” tra la raccolta disorganica e il fondo ordinato è un utile campo di sperimentazione per la collaborazione tra discipline, al fine di consolidare una metodologia che risponda agli scopi preposti delle descrizioni di oggetti culturali: reperire, identificare, selezionare e ottenere¹.

.....
1 M. Savoia, "Lo standard ISAAR come riferimento per la messa a punto di sistemi informativi archivistici",
<http://web.mclink.it/MD1431/sito/isaargrp/isaar.html>

1. IL DIBATTITO INTERDISCIPLINARE

1.1 «LE FONTI DELLO STORICO NELL'ERA DEL COMPUTER»²

Il passaggio delle fonti storiche d'archivio e delle loro descrizioni a oggetti informativi reperibili sul web ha dato vita a non poche diatribe sulla loro stabilità nel tempo (a causa dell'obsolescenza dei mezzi informatici) e sulla loro presunta originalità e attendibilità (sulla fiducia e l'autorevolezza che rivestono, in termini più vicini all'informatica).

Diversi orientamenti si sono poi confrontati su come modellizzare le difficili strutture informative delle fonti storiche, ricche di relazioni e plurilivellari: se dare maggiore rilevanza all'estrapolazione delle informazioni intrinseche dei testi (*model oriented*), o se puntare ad una approssimativa mimesi del documento in sé e tutelarne il contesto in cui è immerso (*source oriented*)³.

Una conclusione sempre più condivisa ormai, date le inevitabili modifiche apportate dall'informatica al rapporto tra storici e fonti, è che l'informazione reperibile dai sistemi informativi sul web

(...) non è più una copia, sia pure abbreviata e impoverita, della fonte originale, ma una fonte nuova, “costruita” dallo storico stesso⁴.

I sistemi informativi creati dagli operatori culturali di archivi, biblioteche, musei allora altro non fanno che seguire i nuovi indirizzi della scienza dell'informazione, decifrando le esigenze degli utenti colti e proponendo un sistema di reperimento che soddisfi al meglio sia le esigenze di catalogazione e ordinamento, sia le spasmodiche ricerche degli studiosi di “preziosità documentarie da loro predilette”⁵, sia le basilari necessità di un'organizzazione dell'informazione che ambisca a ricostruire su uno mezzo diverso e condizionante il contesto originario dei documenti che descrivono.

.....
² Cfr. S. Vitali, *Passato digitale. Le fonti dello storico nell'era del computer*, Milano, Bruno Mondadori, 2004

³ Cfr. S. Vitali, *ibidem*

⁴ O. Itzovich, *L'uso del calcolatore in storiografia*, cit. in S. Vitali, *ivi*, p. 42

⁵ C. Pavone, in *Intorno agli archivi e alle istituzioni*, *ivi*, p. 312

1.2 GLI ARCHIVI FUORI DAGLI ARCHIVI

Le integrazioni interdisciplinari hanno necessariamente allargato il contesto d'azione degli archivisti che immettono informazioni standardizzate in rete, ma la tensione verso l'uscita dall'autoreferenzialità dei propri sistemi informativi passa attraverso il confronto costante con metodi e approcci che appartengono ad altri domini culturali.

In altre parole, insomma, la filosofia sottesa ai processi di standardizzazione insegna soprattutto questo, gli archivisti devono continuare a perseguire i loro obiettivi secondo le proprie specifiche metodologie di lavoro, ma devono divenire capaci – laddove già non lo siano – di trasferire i risultati conseguiti anche al di fuori dello specifico contesto di riferimento. Descrizione, riordino, inventariazione, generazione di sistemi informativi non si esauriscono in sé, ma sono attività ed oggetti finalizzati alla comunicazione e tanto più sono capaci di integrarsi con altri strumenti di ricerca e comunicazione e di esser fruibili, tanto più perseguono il loro scopo⁶.

In primo luogo, è con la sfera degli storici che gli archivisti vanno costantemente a confrontarsi, essendo i loro primi e più colti interlocutori, nonché immediata cartina da tornasole dei passi compiuti verso un migliore metodo di condivisione del patrimonio culturale.

Ma quando entrambe queste sfere conoscitive si apprestano agli strumenti informatici, come rileva sempre Federico Valacchi

L'errore più facile da commettere, guardando al rapporto tra risorse telematiche e archivi storici, è quello di pensare che da tale rapporto possano scaturire opportunità di semplificare il quadro complessivo degli intrecci e delle complicate relazioni che da sempre condizionano e caratterizzano la possibilità degli archivisti di rispondere in maniera esaustiva alle richieste dei loro utenti⁷.

.....

6 Federico Valacchi, "I siti web come strumenti per la ricerca", p. 6, <http://eprints.rclis.org/11683/1/wearchivi.pdf>

7 F. Valacchi, *ivi*, p. 7

Difatti il mezzo comunicativo informatico impone non solo nuove metodologie di lavoro per descrivere la struttura delle informazioni complesse, ma anche nuove problematiche per l'accesso ai contenuti informativi.

Sul piano operativo ciò significa che prima di calare la rete sugli archivi o, se vogliamo, introdurre informazione archivistica ai diversi livelli dentro la rete, occorre sul versante archivistico un lavoro attento, orientato a valutare quali informazioni veicolare attraverso il web e in che modo organizzare la restituzione di queste informazioni, con la consapevolezza che non solo la rete non è uno strumento neutro, ma che, anzi, può generare derive difficilmente controllabili se la si utilizza per descrivere un microcosmo sostanzialmente molto delicato quale quello degli archivi⁸.

Queste giustificate remore da parte degli archivisti derivano dalla consapevolezza di essere i mediatori primari tra le fonti e gli utenti: l'ulteriore passaggio attraverso un medium informativo pone una maggiore attenzione preliminare sui requisiti che le risorse immesse in rete devono soddisfare, poiché la rete non può sopperire alle eventuali carenze informative a monte. Ciò non toglie che, se il mezzo informatico condiziona la struttura dell'informazione, offre anche la possibilità di far convivere approcci diversi all'interno della stessa comunità.

Non è certo una novità dell'informatica quello di offrire agli utenti percorsi di ricerca che mirano più al reperimento dei contenuti che alla chiara comprensione dell'organizzazione di un complesso documentale.

(...) è sommamente auspicabile, che l'inventario secondo la struttura sia integrato da un numero più ampio possibile di indici e guide tematiche, da compilare anche in tempi diversi. È su questo terreno che si recuperano tutti i buoni argomenti adducibili a favore dell'inventario "secondo la materia". È su questo terreno, inoltre, che l'apporto delle tecniche informatiche sembra più produttivo⁹.

Le guide tematiche, gli indici, i *thesauri* sono da sempre considerate utili appendici, ma spesso dispendiose in termini di tempo ed energie, ai già complessi

.....
8 F. Valacchi, *ivi*, pp. 7-8

9 C. Pavone, *ivi*, p. 92

inventari, strumenti di corredo per eccellenza: nell'ambito della gestione delle risorse in rete questo dispendio si riduce notevolmente.

Il dibattito sulla standardizzazione dei punti di accesso per la ricerca nasce anche da queste considerazioni e da qui si estende per identificare punti di raccordo tra le comunità culturali interessate a collegare e condividere le proprie risorse al fine di incrementare il livello qualitativo di queste e offrire un miglior servizio agli utenti.

1.2.1 I punti di accesso nelle descrizioni archivistiche

Riprendendo alcune riflessioni già affrontate qui e nel primo capitolo, l'organizzazione delle descrizioni archivistiche deve essere progettata per dare informazioni rappresentative dell'oggetto descritto e istituire adeguati punti di accesso alle descrizioni stesse: vengono innanzitutto scelti alcuni elementi caratterizzanti delle descrizioni, coerenti con l'ambito del sistema descrittivo e adatti a svolgere la funzione di chiave d'accesso. In seguito si creano forme controllate dei nomi, o liste d'autorità (*authority file*), per identificarli univocamente e indirizzare la ricerca su termini autorizzati, disambigui.

Il ruolo degli *authority file* in campo archivistico è ormai assodato e ha permesso una proficua contiguità con banche dati di altri settori culturali (in primis, quelle catalogafiche): come già ricordato però, i record d'autorità archivistici si impongono di controllare non solo la terminologia finalizzata alla non ambiguità, ma anche quella inerente altre preziose informazioni, che oltre ad identificare, descrivono le entità e le inseriscono nelle relazioni col tempo ed il contesto¹⁰. Dalla nascita dello standard per la descrizione separata dei soggetti produttori, la prospettiva si è molto ampliata al riguardo. Non solo vengono fornite notizie sui mutamenti dei nomi (dovuti a cambiamenti amministrativi o semplicemente a varianti dei nomi), con tipizzazioni sulla natura di queste relazioni, ma vengono formalizzate informazioni relative all'esistenza e alle attività dei soggetti necessarie a contestualizzarli.

In realtà, rispetto al quadro internazionale, il dibattito italiano sulla neces-

.....
¹⁰ Cfr. M. Savoja, "Descrizioni di archivi e soggetti produttori (ISAD-G E ISAAR-CPF)", <http://web.mclink.it/MD1431/sito/isaargrp/isaardef2.rtf>

sità di apertura di nuove chiavi d'accesso ai soggetti produttori e i loro fondi, dopo alcuni importanti passi in questa direzione, come la traduzione del già citato standard ISDF, ha subito una battuta d'arresto.

1.2.2 *Ruoli e funzioni: perché utili added access point*

Approfondire il dibattito sul controllo della terminologia inerente alle funzioni per gli enti, ed i ruoli, le occupazioni, le professioni per le persone fisiche può essere significativo se concepito come offerta di ulteriori chiavi di accesso ai soggetti produttori d'archivio e di conseguenza un arricchimento del percorso di ricerca.

In altri settori culturali sono state messe a punto liste di forme controllate per specificare questo elemento, legandolo alla tipologia documentale a cui inerisce (un esempio su tutti, l'*Art and Architecture Thesaurus (AAT)*¹¹). Ragionare su come estendere le chiavi di ricerca non è utile solo all'utente finale, ma anche a chi descrive le entità archivistiche, permettendo di mappare le relazioni che intercorrono tra esse e le risorse collegate tramite nuovi punti di vista: non basta dire che c'è un legame associativo o di paternità tra una persona e un documento, spesso è riduttivo, se si considera la complessità di informazioni ricavabili da una più attenta analisi dei testi.

Caso emblematico di questa multidimensionalità relazionale tra un soggetto e un documento si ha con le fonti che attestano informazioni sulla natura del soggetto produttore: si può sicuramente dire che un'entità è un "*subjectOf*" di quella risorsa, ma non si arricchisce la definizione né la descrizione dell'entità se non acquista rilevanza semantica nei confronti del legame con la risorsa.

Dire invece che un'entità ha un ruolo e che questo è attestato in una risorsa arricchisce entrambe: per l'entità avremo più punti d'accesso, per cui sarà possibile ricercare l'entità non solo tramite la forma controllata del suo nome, delle sue date significative e dei luoghi con cui è relazionato, ma grazie a termini inerenti uno specifico aspetto della sua esistenza si possono reperire le fonti che ne attestano informazioni specifiche; per le risorse si avrà

.....

11 Cfr. The Getty Research Institute, "Art and Architecture Thesaurus (AAT)", <http://www.getty.edu/research/tools/vocabularies/aat/?find=role&logic=AND¬e=&page=1>

una maggiore indicizzazione dei contenuti e una più precisa relazione con i soggetti di cui tratta.

Una maggiore articolazione delle descrizioni in questo versante descrittivo, ad opinione di chi scrive, non può far altro che contribuire ad una rappresentazione informativa più aderente alla complessità dei legami che intercorrono tra gli oggetti reali descritti dall'archivistica.

1.3 LE EDIZIONI DIGITALI: QUALI SPUNTI DAL FULL TEXT

Addentrarsi nel filone di ricerca che mira ad estrapolare nuove informazioni dal testo pieno di documenti, implica un ripensamento generale di attitudine della disciplina archivistica nei confronti della descrizione dell'unità documentale.

Per tradizione, l'unità archivistica, ovvero l'elemento minimo di descrizione di un fondo, è il fascicolo, o, a seconda della scuola di pensiero, addirittura la serie: l'item, il documento singolo, è raramente indagato e descritto. Questa prassi consolidata per la descrizione dei fondi, la cui mole e collocazione non facilita l'esame accurato di ogni singolo elemento, può essere però rivisitata quando si parla di singole risorse testuali collegate a persone, non per forza risorse prodotte dall'entità, e le si descrive con un modello formale ontologico. Come rileva Francesca Tomasi,

La progressiva estensione degli ambiti di intervento computazionale agli oggetti del patrimonio culturale ha determinato un'attenzione maggiore al documento inteso come dato la cui capacità espressiva va oltre la sola descrizione a livello paratestuale¹².

Estrapolare dal testo di un documento, sia esso archivistico o meno, informazioni sul suo contenuto e rendere queste informazioni partecipi del processo di descrizione di un'entità archivistica, altro non è che trasformare quella stessa fonte documentale in “una nuova fonte, costruita”¹³, ovvero in una sua edizione

.....
12 Francesca Tomasi, *Le edizioni digitali come nuovo modello per dati di autorità concettuali*, in «JLIS», 4, 2, 2013, p. 1; consultabile sul sito <http://leo.cilea.it/index.php/jlis/article/view/8808/8067>

13 O. Itzovich, op. cit.

digitale. Ciò presuppone la scelta di un approfondimento multilivellare sulle informazioni estraibili ed un'interpretazione soggettiva di quali estrapolare e come definirle formalmente: standardizzare queste scelte diviene un passaggio fondamentale, sia per fornire una metodologia univoca di lavoro, sia per dare vita a relazioni tra informazioni che hanno dignità di autonomia, in linea con l'approccio archivistico informatico per cui è sempre auspicabile ampliare le descrizioni ma rendendole autosufficienti e correlabili tramite legami intuitivi.

L'operazione di reperimento di metadati inerenti il contenuto di una fonte consente di aggiungere un nuovo livello di descrizione e integrarlo con uno preesistente per dedurne nuove conoscenze, ovvero

trasformare le entità e le loro relazioni in ontologie significa trasformare i testi in basi di conoscenza¹⁴.

Le nuove conoscenze inferite da questo procedimento arricchiscono la descrizione della natura di un'entità e il suo contesto, andando quindi a convalidare la tesi per cui un *authority record* archivistico offre potenzialità descrittive autonome ed è già una nuova fonte a suo modo.

Il contesto che viene così arricchito è poi intrinsecamente legato sia ai soggetti sia alle documentazioni collegate, anch'esse arricchite semanticamente da queste nuove relazioni: poiché in ogni documento viene attestato un ruolo, sia esso diverso o uguale, ma in diversi contesti, vengono create più relazioni entità-fonte, implementando infinite potenziali specifiche del generico *subjectOf*.

Nell'ottica di creazione di nuovi *access point* e di un nuovo modello di *authority record* archivistico, questa scelta metodologica, se intrapresa, permetterebbe di aprire al web di dati descrizioni archivistiche molto più efficienti dal punto di vista dei percorsi di ricerca e qualitativamente più ricche ed integrabili da altre comunità, come *Text Encoding Initiative (TEI)*¹⁵, che ad oggi affrontano ricerche sulla metadattazione dei testi digitali.

.....
14 F. Tomasi, *ivi*, p. 5

15 Cfr. Text Encoding Initiative (TEI), <http://www.tei-c.org/index.xml>

2. LA DESCRIZIONE DEI RUOLI POLITICI

La decisione di rappresentare con predicati ontologici i ruoli politici di soggetti produttori ed entità ad essi collegate deriva da una generale riflessione sullo scenario nazionale dei fondi afferenti a questa tipologia di soggetti produttori (legati alla sfera politica) ed agli archivi storici che ne detengono la custodia.

Dallo studio sul dibattito attorno alle problematiche che presentano l'ordinamento e la descrizione di questi complessi ho potuto dedurre che si tratta di un contesto proficuo in cui sviluppare i ragionamenti sopracitati.

Questo alla luce delle esigenze storiografiche di una migliore comprensione della composizione e distribuzione di complessi che costituiscono un fondamentale punto di riferimento per la ricostruzione storica degli avvenimenti politici del secolo scorso.

2.1 GLI ARCHIVI STORICI DELLE RELAZIONI POLITICHE

La scelta dell'Amministrazione archivistica di indirizzare risorse ed impegni verso archivi e carte di personaggi è il frutto di una cultura storiografica diffusa di ispirazione idealistica che affrontava la storia politica del nostro paese utilizzando categorie interpretative etico-politiche all'interno delle quali grande spazio ed importanza aveva l'azione di singole personalità e il dibattito delle idee¹⁶.

L'attitudine degli operatori culturali a tutelare e valorizzare le fonti documentarie delle relazioni politiche considerandole un lascito di singole personalità, e non bene comune da controllare e conservare all'interno del più ampio spettro di documentazione che ha finalità legali, sociali e storiche, rivela uno dei sintomi più diffusi della stessa natura dei rapporti politici in Italia nell'ultimo secolo e mezzo. Già Claudio Pavone constatava

il fatto che il carteggio schiettamente politico si trovi (...) negli archivi privati, è

.....

¹⁶ Linda Giuva (a cura di), *Gli archivi storici dei partiti politici europei, Atti del convegno, Roma 13-14 dicembre 1996*, Ministero per i beni e le attività culturali, 2001, p. 18, http://www.archivi.beniculturali.it/dga/uploads/documents/Quaderno_94.pdf

una sia pure indiretta prova, ci sembra, della ristrettezza del ceto politico che fece l'Italia, abituato a trattare le questioni decisive in contatti e in carteggi personali e ufficiosi¹⁷.

Fatte queste premesse, l'interesse per la "memoria della politica"¹⁸ nel corso degli ultimi 50 anni ha subito un rinnovamento e la rivisitazione delle fonti documentarie archivistiche ha aperto nuovi interrogativi storici.

Lo stato dell'arte di questi documenti rivela quell'amalgama sopracitato di storia individuale e storia di partito, per cui indagare diacronicamente le carte di un'istituzione politica o di un singolo militante presenta difficoltà sia di natura materiale di conservazione (a causa delle numerose scissioni dei partiti, con relative perdite di documentazioni, o dell'azione devastatrice del fascismo nei confronti degli archivi non di regime, o ancora dai sequestri delle magistrature e dall'abitudine dei dirigenti politici di considerare personali le carte prodotte in funzione della propria carica), sia dipendenti da scelte volontaristiche di singole personalità di conservare altrove dalle istituzioni la propria partecipazione politica. Questo panorama di frantumazione e dispersione richiede inevitabilmente uno sforzo da parte del sistema conservativo nazionale che

si sta sempre di più costruendo come policentrico accentuando così un carattere già presente nella storia italiana ma che ha subito un'accelerazione negli anni Sessanta grazie all'iniziativa di soggetti civili e politici che hanno salvaguardato la propria memoria dai rischi di perdita e dispersione con la costruzione di luoghi separati¹⁹.

2.2 DALLE CARTE AL SOGGETTO: IL FILO ROSSO DELLE RELAZIONI POLITICHE

Si potrebbe dire, parlando di archivi che custodiscono la memoria delle relazioni politiche di singole personalità e di partiti, che questi sono

.....

17 C. Pavone, *ivi*, p. 52

18 *La memoria della politica*, Roma, Archivio radicale, 1993

19 L. Giuva, *Gli archivi storici dei partiti politici in Italia*, *ivi*, pp. 137-8

(...) in realtà, un meta-archivio. (...) esso risulta un corpo documentale nei confronti del quale ci si approssima per difetto, ossia non si contempla, anche nel caso di una somma di singoli fondi archivistici collocati in sedi diverse, la possibilità del raggiungimento della sua completezza²⁰.

Allora creare un filo rosso che congiunga i vari “frammenti” che descrivono un'entità può essere proficuo per ricostruire il percorso delle carte di un'istituzione, persona o famiglia e per concorrere ad un maggiore livello informativo che descriva questi soggetti.

Questa prospettiva, è inevitabilmente in controtendenza rispetto alla visione accentrante dell'archivio storico: ma per come si presentano le sedimentazioni documentali dei rapporti politici, amalgama di elementi di provenienza individuale e amministrativa, se si vuole offrire una descrizione completa (se non esaustiva) di un'entità politica nell'ambito della gestione di descrizioni archivistiche di soggetti produttori, bisogna far ricorso ad un espediente che consenta di far convivere una metodologia di lavoro che rispetti la tradizione di ordinamento ed una che si apra alle relazioni con documenti, anche singoli, ubicati altrove (in altri fondi o conservati da altri enti) e non direttamente prodotti da quell'entità, che però concorrono ad un suo arricchimento informativo.

Per usare il linguaggio di uno dei più diffusi strumenti di modellizzazione della realtà adottato dagli informatici, il modello entità-relazioni, una situazione (...) non implica affatto una relazione uno a uno fra fondi e soggetti produttori, quanto una relazione molti a molti²¹.

Parafrasando questo concetto basilare dell'archivistica informatica si potrebbe dire che il rapporto *molti a molti* è anche quello tra soggetti e fonti che attestano qualche forma tipizzata di informazione: informazione che viene esplicitata partendo dalle solide tradizioni “oggettive” di conservazione e organizzazione dei complessi documentali, fornendone però una scelta “soggettiva” di va-

.....
20 David Bidussi, *Archivi dei partiti, archivi politici e fondi di documentazione*, in L. Giuva, *ivi*, p. 142

21 S. Vitali, "La seconda edizione di ISAAR (CPF) e il controllo d'autorità nei sistemi di descrizione archivistica", p. 3, http://www.sba.unifi.it/ac/relazioni/vitali_ita.pdf

lorizzazione. Se si potesse fare un paragone “analogico” di un simile percorso tra fonti, ruoli politici e soggetti politici, allora il più rappresentativo potrebbe essere l'inclusione di una forma standardizzata di guida tematica all'interno dell'introduzione all'inventario.

Ritornando al concetto di *reference file*, provare a fornire uno scenario che dia conto delle fonti utilizzate da un redattore di una scheda di soggetto produttore secondo le ISAAR-CPF, confluisce nel processo di definizione autonoma della descrizione stessa in quanto “nuova fonte” di informazioni.

3. UNA PROPOSTA ECLETTICA: UN MODELLO ONTOLOGICO PER FONTI, RUOLI ED ENTITÀ ARCHIVISTICHE

La formalizzazione di relazioni tra metadati che rappresentano entità, ruoli e fonti è interpretabile in un'ottica plurilivellare:

- innanzitutto vi è una relazione tra un'entità ed un ruolo che questa esercita;
- viene caratterizzata la tipologia di ruolo tramite forme controllate del nome;
- il ruolo è indicizzato temporalmente e calato in un contesto, che può essere l'esercizio di quel ruolo semplicemente attestato in un documento, o svolto nei confronti di un'altra entità, o ancora incluso in un evento (e quindi in un luogo e in un tempo);
- questo ruolo indicizzato e contestualizzato è inserito a sua volta in un meta-contesto, ovvero il documento che attesta il ruolo e l'*editore* che ha estratto questa asserzione, o meglio tripla rdf, dal testo;
- la descrizione dell'entità restituita è più strutturata e arricchita dalla serie di relazioni che ne documentano il contesto;

Questo modello relazionale può essere inteso come uno strumento autonomo per la rappresentazione dei soggetti di un ristretto dominio di conoscenza oppure, più auspicabile, divenire un contributo per arricchire *authority system* che già approfondiscono i legami tra *authority*, descrizioni e risorse collegate, dando ulteriori chiavi di lettura delle relazioni e degli *access point* generabili.

Di seguito vengono esposte alcune premesse teoriche/terminologiche al modello implementato: per i riferimenti alle classi e le proprietà citate, si rimanda al capitolo successivo di presentazione dell'ontologia.

3.1 LE FONTI

Le fonti qua intese non sono riconducibili ad un'unica tipologia, né possono riferirsi ovviamente a fonti primarie, ma solo a rappresentazioni digitali di esse o di altre forme di attestazione di informazioni a vario livello. Infatti, per mantenere ampio lo spettro di possibili “fonti”, il concetto va considerato nella sua accezione più dinamica di documento contenente informazioni inerenti al soggetto descritto, interpretabili queste da un osservatore non imparziale: in tal senso, la categoria include le fonti d'archivio prodotte dal soggetto descritto, risorse collegate mediante rapporti di paternità non esclusivamente culturale, risorse che trattano del soggetto e ne descrivono aspetti parziali della sua esistenza, o anche risorse in cui il soggetto è solo citato come contesto di un'altro soggetto descritto. Da questa ultima osservazione, l'interpretazione soggettiva del testo, deriva il concetto qui proposto di “editore”, che decide sulla base di regole condivise ed il più possibile standardizzate, di estrapolare determinate informazioni da un contesto testuale, interpretato alla luce del metadato che se ne vuole estrapolare.

Le soluzioni adottate nel modello ontologico, il meta-contesto delle fonti e l'uso di specifiche *object property*, consentono di dichiarare per ogni asserzione un editore responsabile di quell'affermazione e, a sua volta, di essere contestata da altri sviluppatori di asserzioni, al fine di far convivere nel modello anche affermazioni che si contraddicono, descrivendone le motivazioni.

3.2 I RUOLI INDICIZZATI

Si è già detto della funzione del “ruolo” come mediatore per l'accesso, il riacquisto e la descrizione delle entità archivistiche; descrivere qualsiasi entità, o “attributo” di un'entità, richiede preliminarmente di interrogarsi su quali aspetti far emergere e come definirli formalmente. Nel caso specifico, la rappresentazione dei ruoli di persone, si può tentare un generico confronto con le regole dello standard ISDF per descrivere le funzioni di enti, in modo da ricavarne linee guida sui punti fermi da sviluppare e poi riflettere sui predicati.

L'identificazione del ruolo (speculare a 5.1 ISDF) si ha con la definizione

della classe *pro:Role* e le sue sottoclassi per i ruoli editoriali e politici, *pro:Publishing role* e *proles:Political role*, e le forme autorizzate del nome possono considerarsi le istanze di queste classi: non ci sono vincoli rigidi di nomenclatura, possono essere sempre aggiunte nuove istanze a seconda delle necessità di rappresentazione di un ruolo particolare. Per ora non ci si è posti il problema di definire univocamente quali istanze siano accettate, per poter continuare la discussione sulla base delle esigenze specifiche di un numero maggiore di descrizioni e dedurre degli assunti il più possibile generalizzabili. Le istanze presentate in questa sede infatti sono solo un tentativo di richiamare gli elementi potenzialmente più interessanti, per mappare le relazioni tra entità mediate dai ruoli, senza pretese di esaustività.

La contestualizzazione storica del ruolo (5.2 ISDF) avviene grazie all'ausilio di annotazioni per la descrizione testuale della relazione (*dc:description*) e ad un *pattern* ontologico per le indicizzazioni temporali (*ti:Time Interval*). Le altre informazioni standardizzate da ISDF (gli aspetti storici e legislativi) non sono al momento previste, ma è possibile riusare proprietà di altri domini (*DC* ad esempio) per includerle.

Non è stata pensata una relazione diretta tra ruoli (5.3 ISDF), che rimangono autonomi l'uno dall'altro, senza legami associativi o gerarchici, ma è possibile in fase di interrogazione reperire ruoli aventi elementi comuni, come le entità che esercitano il ruolo o sono relazionate, gli eventi a cui partecipano, le risorse che attestano la loro "esistenza" e gli intervalli di tempo.

Il controllo della *provenance* della descrizione del ruolo (5.4 ISDF) viene esplicitato riusando due proprietà dell'apposita ontologia *PROV-O*, *prov:wasAttributedTo* per identificare chi ha creato l'asserzione sui ruoli e *prov:hadPrimarySource* per specificare la fonte da cui è stata estratta, e due proprietà di *DC*, *date* e *language*. Per ciò che concerne le relazioni tra ruoli e altre entità o risorse (6. ISDF), vale quanto detto per la relazionalità tra ruoli.

3.3 I RUOLI POLITICI

I ruoli politici, vanno in questa sede interpretati come uno dei possibili espedienti per dimostrare le possibilità relazionali tra risorse e entità descritte.

L'obiettivo principale che ha mosso questo lavoro non è stata la ricerca e la definizione delle migliori forme controllate possibili per i ruoli politici, ma la sperimentazione sulle relazioni tra un soggetto e un oggetto mediate da un arricchimento semantico della relazione stessa.

Inoltre, i ruoli politici non sono stati ristretti all'esercizio di una carica istituzionale, livello che volutamente non viene approfondito perché più confacente alla categoria di funzione amministrativa e perciò ristretta in una gerarchia di situazioni a loro volta vincolate al contesto dell'istituzione di riferimento. Per ruolo politico si intende più una relazione politica, quando non semplicemente sociale, tra personalità e organizzazioni che hanno intrattenuto legami politici in contesti determinati, al di fuori delle sole realtà istituzionali: perciò, ad esempio, anche i legami di amicizia intrattenuti da militanti di uno stesso partito o la rivalità tra due personaggi che hanno partecipato allo stesso movimento politico sono stati inclusi tra i possibili scenari da contestualizzare, ampliando lo spettro delle relazioni associative.

Si è pensato di riusare a tal fine (e per una maggiore interoperabilità) alcuni modelli già in uso, come l'ontologia *Pro*, nata per descrivere i ruoli editoriali in un dominio più ampio, o il pattern *Nary participation*, per descrivere la partecipazione di oggetti ad un evento, che si sono prestati ad estensioni con nuove (o già esistenti) proprietà in grado di rappresentare uno scenario più complesso: le estensioni identificano gli elementi necessari a descrivere le relazioni politiche, quindi una classe per fornire forme controllate dei ruoli, *proles:Political role*, una per gestire le relazioni politiche tra entità e restringerle in un arco temporale, *proles:Political role in time*, una classe per contestualizzare le relazioni in un evento, *proles:Participation with political role*, e una serie di sottoproprietà per restringere i casi di relazionalità.

CAPITOLO QUARTO

Una proposta di ontologia: Political Roles (PROles)

PREMESSA

L'ontologia Political Roles (PROles)¹ descrive le relazioni politiche intrattenu-
te da personalità e/o organizzazioni in un determinato contesto e le circoscri-
ve temporalmente. Ogni relazione è documentata da una fonte che ne attesta
un'occorrenza ed è posta sotto la responsabilità culturale della persona/agen-
zia che ha estratto l'informazione dal testo.

Il modello si basa fundamentalmente su *Publishing Roles Ontology (PRO)*²
per la descrizione dei ruoli e su *Nary participation pattern*³ per la contestualiz-
zazione delle relazioni politiche in un evento; vengono riusate alcune proprietà
di *Provenance Ontology (PROV-O)*⁴, per stabilire la provenienza e la paternità
delle relazioni, ed alcune classi e proprietà appartenenti ai più comuni data
model, come *DCTerms*⁵ e *FOAF*⁶.

.....
1 Cfr. Marilena Daquino, Silvio Peroni, Francesca Tomasi, "PROles Ontology",
<http://www.essepuntato.it/2013/10/politicalroles>

2 Cfr. Silvio Peroni, David Shotton, Fabio Vitali, *Scholarly publishing and the Linked Data: describing roles, statuses, temporal and contextual extents*, in H. Sack, T. Pellegrini, *Proceedings of the 8th International Conference on Semantic Systems*, ACM, New York, 2012; Publishing Roles Ontology (PRO), <http://purl.org/spar/pro>

3 Cfr. Aldo Gangemi, "Nary Participation",
http://ontologydesignpatterns.org/wiki/Submissions:Nary_Participation

4 Cfr. Timoty Lebo, Satya Sahoo, Deborah McGuinness, "PROV-O: The PROV Ontology. W3C Recommendation, 30 April 2013, World Wide Web Consortium"; PROV-O, <http://www.w3.org/TR/prov-o/>

5 Cfr. Dublin Core Metadata Initiative, "DCTerms", <http://dublincore.org/documents/dcmi-terms/>

6 Cfr. Dan Brickley, Libby Miller, "FOAF", 2010, <http://xmlns.com/foaf/spec/>

PRoles è stata ideata per essere una soluzione autonoma di descrizione e relazione tra ruoli, entità e documenti ma è predisposta per essere riusata ed importata in altre ontologie che descrivono entità di cui si vuole fornire un contesto informativo più ampio.

Le istanze che rappresentano i ruoli, editoriali e politici, possono essere estese semplicemente aggiungendone di nuove senza modificare in alcun modo la parte terminologica (TBoX) dell'ontologia, in base alle esigenze di descrizione.

1. INTRODUZIONE

L'approccio alla base dell'ontologia prevede l'attribuzione di un ruolo ad un "agente" (*pro:holdsRoleInTime*) in uno specifico contesto d'azione (*pro:relatesToEntity*) che può essere l'esercizio del ruolo nei confronti di una persona, l'espletamento di una funzione all'interno di un'organizzazione, l'attestazione generica di quel ruolo in un documento o lo svolgimento di quel ruolo in un luogo: ogni attribuzione di ruolo è indicizzata temporalmente (*tvc:atTime*), di modo da poter ricostruire l'attività degli agenti, ovvero le istanze di *foaf:Agent*, in una prospettiva diacronica che renda conto anche delle relazioni che essi intrattengono contemporaneamente, creando una rete di rapporti tra persone, organizzazioni, luoghi e attività che restituisca una maggiore complessità della realtà in cui agiscono.

Per questa ricostruzione di "fatti", è stato sufficiente l'uso di PRO: il fulcro della modellizzazione è la classe *pro:RoleInTime* e la sua estensione con la sottoclasse *proles:PoliticalRoleInTime*, minimo comune denominatore delle altre relazioni approfondite nell'ontologia.

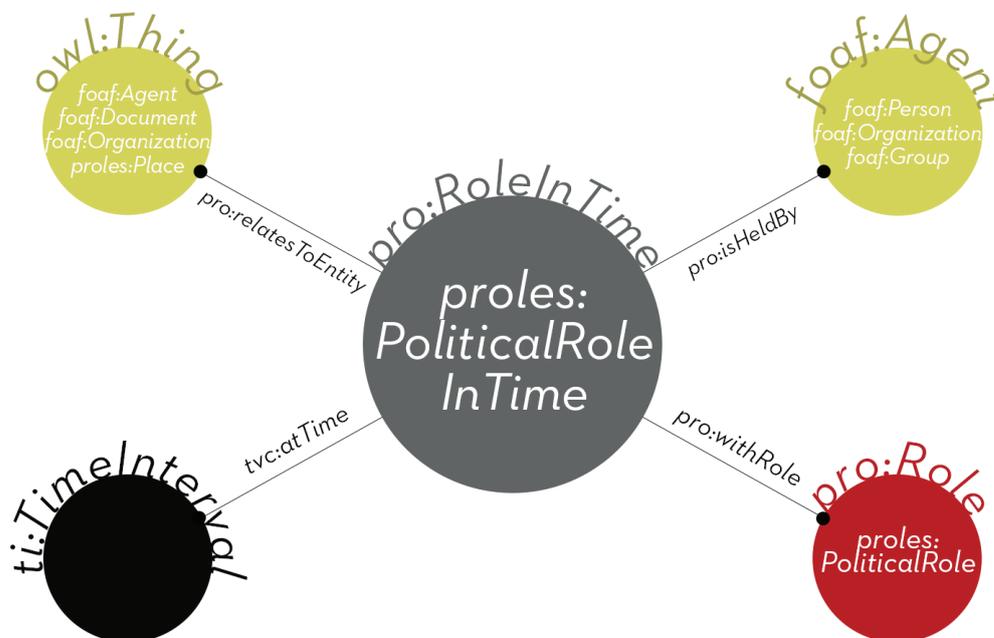


Fig. 4.1 Rappresentazione grafica dell'estensione di PRO in PRoles.

In una prima fase di sperimentazione del modello le relazioni vengono applicate al solo contesto dei ruoli politici, e non ai ruoli in generale, perciò gli ulteriori livelli di approfondimento che seguono vanno intesi unicamente per *proles:PoliticalRoleInTime*. Partendo dalle istanze di questa classe, il modello si estende in due direzioni.

Le circostanze in cui si espleta un ruolo politico di un agente vengono ampliate all'interno di un evento: l'inclusione di un'istanza di *proles:PoliticalRoleInTime* in un evento è stata resa grazie all'uso del pattern ontologico⁷ *Nary Participation*. Grazie a questo espediente è possibile definire dei “partecipanti” ad un evento, sulla cui definizione si è voluti rimanere molto generici al fine di un uso più elastico del pattern (approccio comune a molti data model, vedi CIDOC-CRM⁸), e attribuire ad essi una partecipazione mediata da un ruolo politico indicizzato.

La classe *nary:NaryParticipation* è stata perciò estesa con una sottoclasse, *proles:ParticipationWithPoliticalRole*, che descrive la situazione particolare di un *foaf:Agent* che partecipa ad un evento con un determinato ruolo politico.

Al tempo stesso, questa generale visione (un'entità avente un ruolo, immersa in un contesto d'azione descritto da una situazione) è una restrizione del ruolo stesso: infatti maggiore è il livello analitico di descrizione del contesto, minore è la generalizzazione di attribuzione di un ruolo, che al di fuori della situazione descritta potrebbe non essere valido.

Un agente può esercitare un ruolo politico in un preciso contesto, ma in un contesto diverso (o in un diverso arco temporale) può non essere un'attribuzione corretta.

Ad esempio, un militante di partito può partecipare ad un movimento politico allineandosi ad un preciso schieramento ideologico interno; terminata l'esperienza/partecipazione a quel movimento la sua appartenenza alle posizioni del gruppo d'appartenenza possono mutare, perciò è corretto dire che quel militante ha avuto un ruolo politico in un'organizzazione nello specifico contesto del movimento politico, ma al di fuori di esso non è plausibile.

.....

⁷ Un *pattern ontologico* è una soluzione generale offerta per descrivere situazioni comuni nella modellizzazione di un'ontologia.

⁸ Martin Doerr, *The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata*, in «AI Magazine», vol.24, n°3; CIDOC-CRM, <http://www.cidoc-crm.org/>

Viene inoltre fornito un meta-contesto alle asserzioni fatte sui ruoli, dichiarando la fonte da cui sono state estratte, grazie alla proprietà *prov:hadPrimarySource*, e determinando una figura responsabile di aver dedotto quelle asserzioni da un testo, mediante la proprietà *prov:wasAttributedTo*.

Questo espediente consente di fornire informazioni sulla provenienza delle asserzioni e contemporaneamente definire un (meta-)contesto ancora più specifico per le relazioni politiche.

Una relazione esiste ed è attestata in una fonte: al di fuori di quella fonte può esservi un'altro documento che ne attesta una variante o una contraddizione, perciò quella prima asserzione sul ruolo ha una validità circoscritta alla fonte che l'attesta e può convivere in separata sede con una sua contraddizione; anche all'interno della stessa fonte può esservi una contraddizione, dovuta ai diversi punti di vista di più "editori" di asserzioni ontologiche che possono contestare le affermazioni pregresse in merito ad una relazione e definirne una nuova che sostanzialmente nega la precedente, consentendo un dibattito sulla critica delle fonti anche in sede di metadattazione. Di conseguenza, le informazioni di provenienza sono riferite sia alle istanze di *proles:PoliticalRoleInTime* sia alle istanze di *proles:ParticipationWithPoliticalRole*.

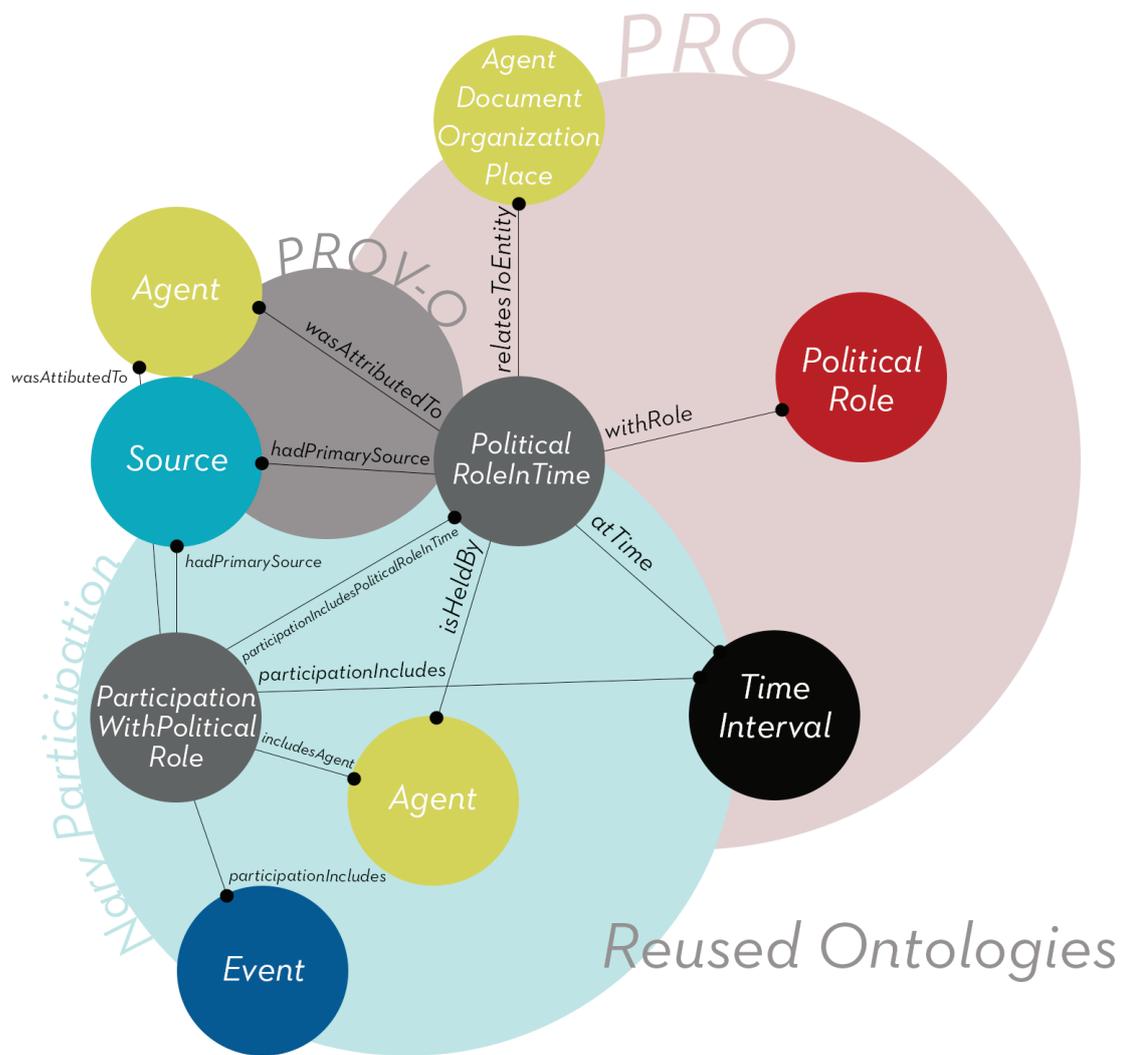


Fig. 4.2 Rappresentazione grafica dei modelli riusati in PRoles e delle loro estensioni.

È possibile, in definitiva, adottare tre livelli di contestualizzazione per il ruolo di un agente, assimilabili alla volontà di usare uno o più modelli importati in PRoles.

Contesto di paternità. Si può attribuire ad un individuo di *foaf:Agent* un ruolo indicizzato temporalmente avente come contesto un documento, utilizzando solo il modello PRO.

In realtà, in PRO la relazione è intesa come una forma di paternità o di esercizio diretto di un ruolo (editoriale) di *foaf:Agent* su *foaf:Document* (un agente può essere l'editore, il revisore... del documento), mentre in PRoles il legame viene considerato una generica relazione, che non esplicita un'attività precisa di un agente sul documento e può essere quindi inteso come una forma di

attestazione di quel ruolo. Essendo questa generalizzazione una forzatura del significato originario della proprietà *pro:relatesToEntity*, è comunque preferibile utilizzare i livelli successivi per attribuire la *provenance*.

Contesto di provenienza. Si attribuisce ad un individuo di *foaf:Agent* un ruolo indicizzato temporalmente e relazionato con altre entità diverse da un documento (*Person, Organization, Place*) come previsto dal modello, contestualizzando questa asserzione in una fonte (*proles:Source*), ovvero estendendo PRO con le proprietà di PROV-O.

Contesto di partecipazione. Si attribuisce ad un individuo di *foaf:Agent* una situazione (*nary:Nary Participation*) in cui questo vi interviene con un ruolo indicizzato temporalmente e già contestualizzato in una fonte (PRO + PROV-O); a sua volta la situazione generata (*ParticipationWithPoliticalRole*) è ri-contestualizzata in una fonte, che può essere la medesima dell'asserzione sul ruolo nel tempo oppure un'altra.

Colui che attesta la partecipazione di un *foaf:Agent* in un evento, può riutilizzare una precedente asserzione inerente un'altra fonte che descrive un ruolo di *foaf:Agent* e dichiarare che l'agente in questione è intervenuto nella nuova situazione indicizzata con quel preciso ruolo attestato altrove.

È questo il livello più particolareggiato di specificità, che richiede l'uso di PRO, PROV-O e Nary Participation.

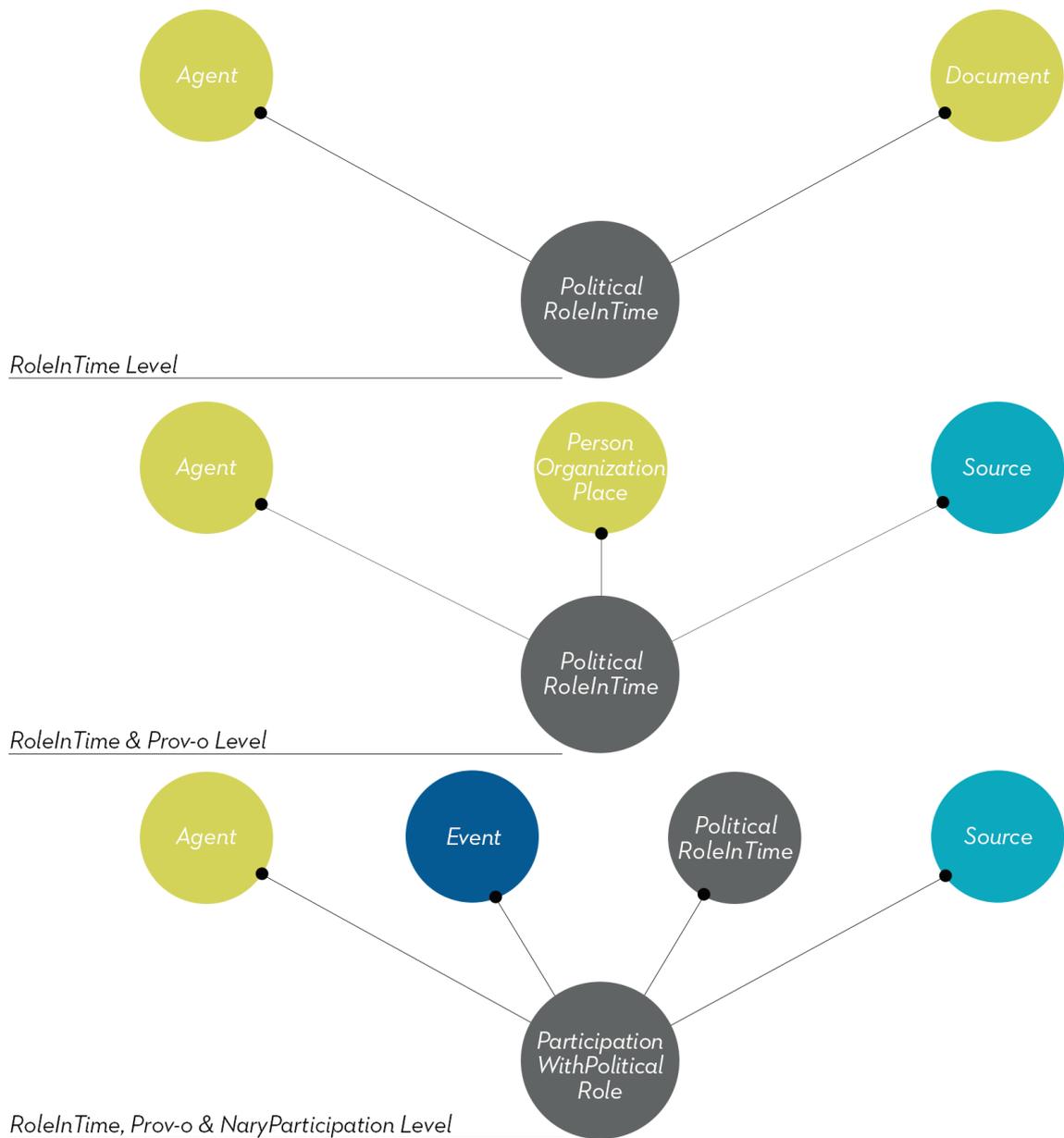


Fig. 4.3 Livelli rappresentativi delle relazioni di contesto in PRoles.

Come già detto, l'articolazione delle relazioni di contesto è inversamente proporzionale alla loro generalizzazione, perciò i nodi che intercorrono e mediano tra un *foaf:Agent*, un *proles:PoliticalRole* e una *proles:Source* aumentano.

Per favorire la convivenza di più livelli di analiticità e l'allineamento con altre ontologie sono state create delle proprietà generali che fungono da “ponte tra le classi”:

- *proles:describesAgent;*
- *proles:describesSituation;*

- *proles:hasRelatedAgent*;
- *proles:hasRelatedSource*.

Le relazioni sopra descritte tra classi sulla base dei ruoli indicizzati temporalmente e calati nel contesto di un evento possono essere interpretate come una ulteriore specializzazione di queste quattro proprietà, non in senso gerarchico, da adottare per descrivere situazioni complesse.

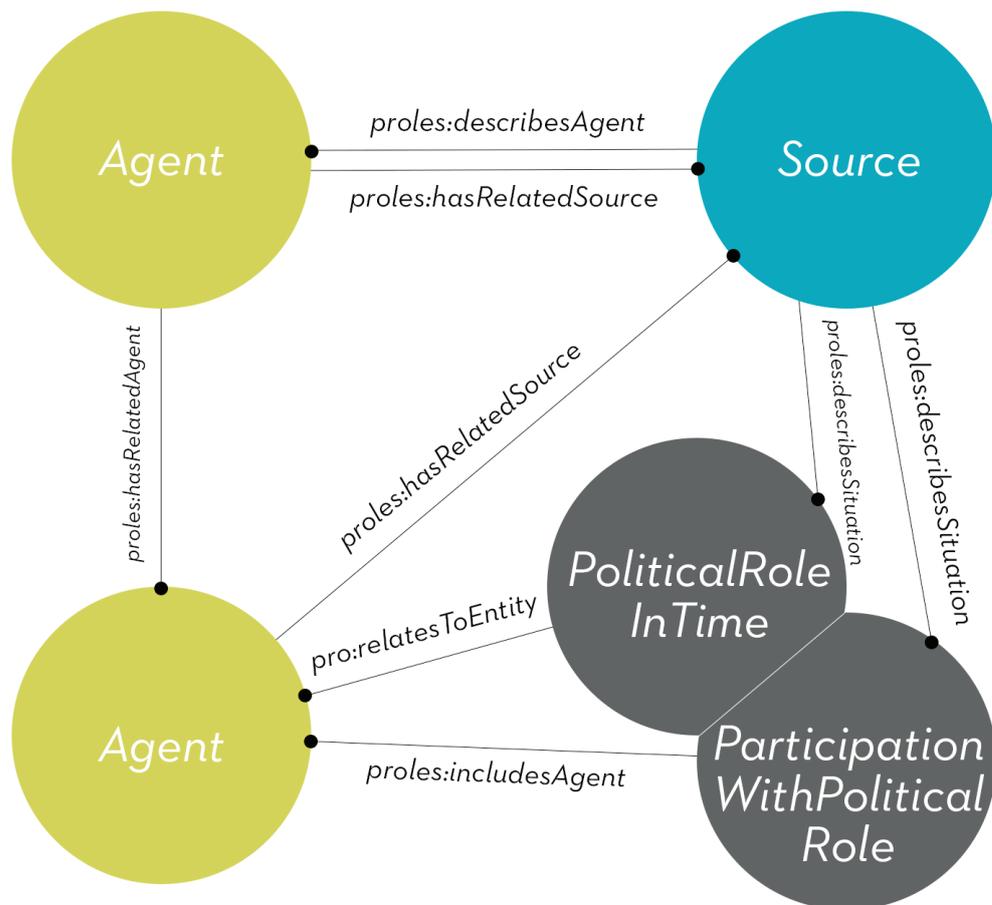


Fig. 4.4 Rappresentazione grafica delle proprietà di PRoles per l'allineamento con altre ontologie.

2. LE CLASSI PRINCIPALI

Vengono qui illustrate le caratteristiche delle classi fondamentali dell'ontologia di cui si è detto sopra.

Per una visualizzazione completa dell'element set di PRoles in formato human-readable è stato utilizzato un tool di rendering in formato HTML per le classi e le proprietà estratte dall'ontologia, *Live Owl Documentation Environment (LODE)*⁹.

2.1 AGENT (FOAF:AGENT)¹⁰

È la classe di FOAF per entità che compiono o sono in grado di compiere azioni: ha come sottoclassi *foaf:Person*, *foaf:Group* (ovvero un insieme di individui di *foaf:Person*) e *foaf:Organization*. In PRoles *foaf:Agent* è uno dei nodi centrali della rappresentazione, a cui vengono attribuiti ruoli, sono coinvolti in situazioni e sono in relazione con altre risorse che attestano asserzioni su di essa.

È il dominio di una serie di proprietà che la relazionano un agente:

- con altri agenti (individui della classe *foaf:Agent*) mediante la proprietà *proles:hasRelatedAgent*;
- con le risorse in cui è attestata un'informazione su un ruolo politico che l'agente esercita in un contesto specifico, mediante la proprietà *proles:hasRelatedSource*;
- con la rappresentazione ontologica dei ruoli politici indicizzati temporalmente e contestualizzati (*proles:holdsPoliticalRoleInTime*);
- con una situazione in cui è coinvolto in veste del suo ruolo politico (*proles:isIncludedInParticipationWithPoliticalRole*);

Gli individui della classe *foaf:Agent* possono essere, conseguentemente, descritti mediante l'uso di molteplici relazioni che, per usare un concetto figlio dell'archivistica informatica, affinano la descrizione della sua natura.

.....
⁹ Cfr. Silvio Peroni, David Shotton, Fabio Vitali, *The Live OWL Documentation Environment: A Tool for the Automatic Generation of Ontology Documentation*, in A. Teije, J. Völker, S. Handschuh, H. Stuckenschmidt, M. d' Aquin, A. Nikolov, ... N. Hernandez, *Proceedings of the 18th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2012), Lecture Notes in Computer Science*, Berlin, Heidelberg, Germany, Springer, 2012; LODE, www.essepuntato.it/lode

¹⁰ Cfr. FOAF, "foaf:Agent", http://xmlns.com/foaf/spec/#term_Agent

2.2 POLITICAL ROLE (PROLES:POLITICALROLE)¹¹

È la classe dei ruoli politici che un *foaf:Agent* può avere nell'ambito delle relazioni che intrattiene con altri agenti o in un luogo che lo contestualizza.

Gli individui di questa classe sono relazionati con istanze della classe *proles:PoliticalRoleInTime* mediante una restrizione sull'object property *pro:withRole*.

Allo stato corrente di sviluppo, la scelta dei ruoli politici introdotti nell'ontologia si è basata sullo studio di alcune biografie di personaggi politici, vissuti a cavallo del '900, estratte dalle schede EAC-CPF di IBC-xDams, ciò anche in vista di una tentativo di uso combinato di PROLES e EAC-CPF Ontology.

Ad ogni modo, come già anticipato, è possibile sopperire alle lacune di definizione degli individui aggiungendone di nuovi, senza modificare la base terminologica (TBox) dell'ontologia stessa, in base ad uno studio più accurato della loro occorrenza nelle fonti. Gli individui della classe *proles:PoliticalRole* sono:

- *proles:enemy* per relazionare personaggi politici che hanno militato in schieramenti opposti ideologicamente;
- *proles:friend* per attestare legami tra persone vicine sia politicamente che a livello umano;
- *proles:influential-person* per indicare un ruolo gerarchico tra persone sulla base delle loro idee comuni;
- *proles:institutional-role* per fornire un generico ambito dell'attività amministrativa svolta dalla persona all'interno di un organismo istituzionale;
- *proles:leader* per relazionare un personaggio ad un movimento politico, un'organizzazione, un evento di cui è stato promotore o vi ha esercitato quella funzione;
- *proles:militant* per correlare l'attività di un individuo ad un'organizzazione di riferimento;
- *proles:prisoner* per relazionare una persona ad un luogo in cui è stato prigioniero;
- *proles:refugee* per indicare i luoghi in cui il personaggio ha viaggiato clandestinamente;

.....
¹¹ Cfr. PRoles, <http://www.essepuntato.it/2013/10/politicalroles/PoliticalRole>

- *proles:sympathizer* per attestare un legame associativo tra persone sulla base di una simpatia politica verso un'idea comune.

2.3 SOURCE (PROLES:SOURCE)¹²

È la classe delle fonti che attestano asserzioni sui ruoli politici di una agente.

Insieme alle classi *foaf:Agent* e *proles:PoliticalRole* è il fulcro del modello per la rappresentazione dei legami tra gli agenti, i loro ruoli e risorse collegate.

Per favorire il mapping con altre ontologie, *proles:Source* è dichiarato come dominio di alcune proprietà che mantengono un livello generale di relazioni tra le risorse, le situazioni (*proles:describesSituation*), qui intese come istanze di *proles:PoliticalRoleInTime* e *proles:ParticipationWithPoliticalRole*, e gli agenti descritti (*proles:describesAgent*).

Non sono previste ulteriori specifiche per le tipologie di fonti contemplate, lasciando ampi margini di inclusione e riconoscendo validi testi provenienti da altri domini, aventi supporti diversi e diversi livelli di descrizione metadateale: sono ammesse anche risorse che non hanno alcuna rappresentazione sul web, per le quali è possibile fornire una descrizione testuale (*rdfs:comment* o *dc:description*) per la loro identificazione e collocazione.

2.4 TIME INTERVAL (TI:TIMEINTERVAL)¹³

Circoscrive gli archi temporali trattandoli come oggetti a cui attribuire un intervallo di tempo, formato da una data di inizio e una data di termine.

Le istanze di *ti:TimeInterval*, in quanto codominio della proprietà *tvc:atTime*, concorrono a restringere le classi *pro:RoleInTime*, *proles:PoliticalRoleInTime*, *nary:NaryParticipation* e *proles:ParticipationWithPoliticalRole*.

La possibilità di esplicitare un intervallo di tempo per un ruolo, separatamente dall'intervallo di tempo della partecipazione in un evento, consente di fare affermazioni più complesse. Un'entità può esercitare un ruolo per un arco di tempo maggiore o inferiore rispetto alla sua partecipazione in un evento.

.....

¹² Cfr. PRoles, <http://www.essepuntato.it/2013/10/politicalroles/Source>

¹³ Cfr. Valentina Presutti, "TimeInterval", <http://ontologydesignpatterns.org/wiki/Submissions:TimeInterval>

Ad esempio, un iscritto ad un partito per dieci anni può partecipare ad un movimento politico esercitando il determinato ruolo di leader; il movimento può svilupparsi nell'arco di due anni ma il militante vi partecipa con il ruolo di leader solo per un anno.

2.5 POLITICAL ROLE IN TIME (PROLES:POLITICALROLEINTIME)¹⁴

La classe descrive una specifica situazione in cui viene attribuito un ruolo politico ad un agente per un periodo di tempo limitato ed in un determinato contesto.

È sottoclasse di *pro:RoleInTime*, a sua volta sottoclasse di *tvc:ValueInTime*¹⁵, *tisit:TimeIndexedSituation*¹⁶ e *situation:Situation*¹⁷, pattern utilizzati per descrivere situazioni indicizzate temporalmente.

Rispetto a queste classi, *proles:PoliticalRoleInTime* prevede la dichiarazione esplicita della provenienza (ovvero la specificazione della fonte da cui è stata tratta la descrizione ontologica e l'agente responsabile di questa descrizione) delle asserzioni fatte sui ruoli politici. Le proprietà che caratterizzano la classe sono:

- *tvc:atTime* per definire l'intervallo di tempo (un individuo di *ti:TimeInterval*) in cui è esercitato il ruolo;
- *pro:isHeldBy* per attribuire il ruolo indicizzato ad un agente;
- *pro:relatesToEntity* per relazionare il ruolo ad un documento (*foaf:Document*), un agente (*foaf:Person* o *foaf:Organization*) o un luogo (*proles:Place*);
- *proles:withRole* per identificare il ruolo politico (definito dagli individui di *proles:PoliticalRole*);
- *prov:hadPrimarySource* per stabilire la fonte (*proles:Source*) da cui si sono definite le asserzioni ontologiche relative al ruolo;
- *prov:wasAttributedTo* per identificare chi (*foaf:Agent*) ha estrapolato quest'asserzione dalla fonte.

14 Cfr. PRoles, <http://www.essepuntato.it/2013/10/politicalroles/PoliticalRoleInTime>

15 Cfr. S. Peroni, F. Vitali, D. Shotton, "The Time-Indexed Value in Context (TVC)", <http://www.essepuntato.it/2012/04/tvc>

16 Cfr. A. Gangemi, "Time Indexed Situation", <http://ontologydesignpatterns.org/wiki/Submissions:TimeIndexedSituation>

17 Cfr. A. Gangemi, "Situation", <http://ontologydesignpatterns.org/wiki/Submissions:Situation>

2.6 EVENT (PARTICIPATION:EVENT)¹⁸

La classe rappresenta ogni evento, fisico, sociale, materiale o situazione in senso lato. Ha una definizione molto ampia, affinché il maggior numero di scenari per le relazioni possano essere considerate eventi e collegati ad altre entità.

Nella classe *nary:NaryParticipation*, *participation:Event* viene usata per relazionare la partecipazione di agenti e altri oggetti, vincolata in un preciso intervallo di tempo, ad un determinato evento.

In PRoles viene utilizzata per offrire un contesto ancora più specifico di quello proposto da PRO nelle relazioni tra ruoli politici e agenti.

2.7 PARTICIPATION WITH POLITICAL ROLE (PROLES:PARTICIPATIONWITHPOLITICALROLE)¹⁹

La classe descrive una situazione dove un un'entità partecipa, per un determinato periodo di tempo, ad un evento con un determinato ruolo (indicizzato temporalmente).

È un'estensione della classe *nary:NaryParticipation* e permette l'inclusione di un individuo della classe *proles:PoliticalRoleInTime* e delle informazioni aggiuntive per stabilire la provenienza delle asserzioni ontologiche (ovvero la specificazione della fonte da cui è stata tratta la descrizione ontologica e l'agente responsabile di questa descrizione). Le proprietà che definiscono la classe sono:

- *nary:participationIncludes* per l'inclusione di almeno un evento (*participation:Event*), un intervallo di tempo (*ti:TimeInterval*) e un altro oggetto (*participation:Object*), inteso come qualsiasi rappresentazione di un agente, entità mentale o fisica;
- *proles:participationIncludesPoliticalRoleInTime* per indicare il ruolo (*proles:PoliticalRoleInTime*) con cui un agente partecipa all'evento;
- *proles:includesAgent* per identificare gli agenti che partecipano alla situazione che include un evento;

18 Cfr. A. Gangemi, "Participation", <http://ontologydesignpatterns.org/wiki/Submissions:Participation>

19 Cfr. PRoles, <http://www.essepuntato.it/2013/10/politicalroles/ParticipationWithPoliticalRole>

- *prov:hadPrimarySource* per dichiarare la fonte (un individuo della classe *proles:Source*) da cui si è desunta la partecipazione all'evento;
- *prov:wasAttributedTo* per identificare chi (*foaf:Agent*) ha estrapolato questa asserzione alla fonte.

3. LE OBJECT PROPERTY

PROles riassume le object property dei modelli sopracitati, PROV-O, PRO e Nary Participation (incluse le proprietà che questi importano a loro volta) e ne implementa di nuove per le specifiche dei ruoli politici.

Verranno qui illustrate solo le proprietà essenziali alla descrizione dei ruoli politici esercitati in un determinato contesto, partendo dalle relazioni generiche tra entità fino alle proprietà propedeutiche alla definizione delle situazioni in cui sono coinvolti ruoli, persone e fonti. Per una visualizzazione completa di tutte le entità ontologiche, incluse le proprietà importate ma non espressamente richieste per la rappresentazione dei ruoli, si rimanda alla relativa pagina HTML restituita da LODE²⁰.

Per ogni proprietà vengono forniti: un'etichetta (o *rdfs:label*) e un nome locale, un IRI in nota, un commento, il dominio e codominio della proprietà (ove non siano genericamente *owl:Thing* o non specificati), le sue caratteristiche ed i rapporti di super-/sotto-proprietà.

Di alcune proprietà, la cui definizione ha richiesto la creazione di una catena di proprietà²¹, verranno illustrati i passaggi intermedi.

3.1 DESCRIBES AGENT (PROLES:DESCRIBESAGENT)²²

Dominio: *proles:Source*

Codominio: *foaf:Agent*

La proprietà relaziona una fonte ad un agente di cui fornisce informazioni.

.....

²⁰ Cfr. PROles, <http://www.essepuntato.it/lode/http://www.essepuntato.it/tmp/1384372764-ontology>

²¹ Cfr. Pascal Hitzler, Markus Krötzsch, Bijan Parsia, Peter F. Patel-Schneider, Sebastian Rudolph, "OWL 2 Primer, Property chains", http://www.w3.org/TR/owl2-primer/#Property_Chains

²² Cfr. PROles, <http://www.essepuntato.it/2013/10/politicalroles/describesAgent>

È una proprietà generale, utile a legare direttamente un'entità ad una risorsa ad essa collegata senza ulteriori specifiche relazionali, ma dal nome significativo che rimanda al tipo di relazione che intercorre tra gli elementi.

Per definire univocamente quali agenti debbano essere considerati codominio della proprietà è stato necessario creare una catena di proprietà che specificasse i passaggi intermedi²³: una fonte descrive solo *foaf:Agent* che partecipano ad una situazione (*proles:ParticipationWithPoliticalRole*) o esercitano un ruolo politico (*proles:PoliticalRoleInTime*).

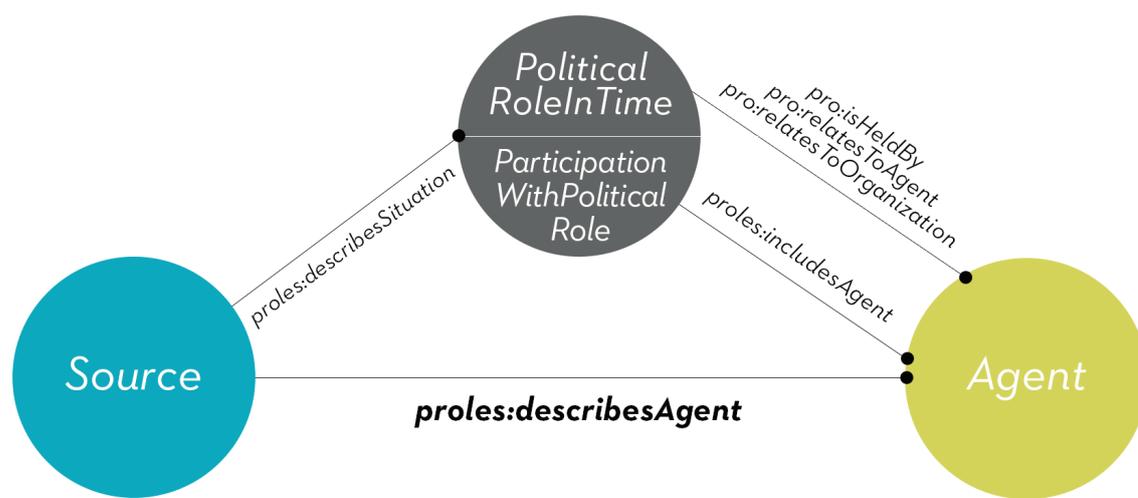


Fig.4.5 Rappresentazione grafica delle catene di proprietà di *proles:describesAgent*.

3.2 DESCRIBES SITUATION (PROLES:DESCRIBESSITUATION)²⁴

Dominio: *proles:Source*

Codominio: *situation:Situation*

Proprietà inversa: *prov:hadPrimarySource*

La proprietà relaziona una fonte ad una situazione descritta in essa.

Situation è la superclasse che include gli scenari possibili di contesto per una relazione politica, tra cui *proles:PoliticalRoleInTime* e *proles:ParticipationWithPoliticalRole*. La proprietà è espressamente generica per lasciare che vi sia un

```
23 rdfs:domain proles:Source; rdfs:range foaf:Agent; owl:propertyChainAxiom
( proles:describesSituation <http://purl.org/spar/pro/isHeldBy> ) ,
( proles:describesSituation proles:includesAgent ) ,
( proles:describesSituation <http://purl.org/spar/pro/relatesToOrganization> ) ,
( proles:describesSituation <http://purl.org/spar/pro/relatesToPerson> ) .
```

24 Cfr. PRoles, <http://www.essepuntato.it/2013/10/politicalroles/describesSituation>

ampio range di contesti attestati in una fonte, nell'ottica di arricchire le relazioni di *proles:Source*.

3.3 HAS RELATED AGENT (PROLES:HASRELATEDAGENT)²⁵

Dominio: *foaf:Agent*

Codominio: *foaf:Agent*

La proprietà relaziona direttamente due agenti coinvolti in una relazione politica (*pro:RoleInTime*). Come le precedenti, è una proprietà generica che non tiene conto, se non nella sua definizione tramite una catena di proprietà, delle specifiche relazioni intrattenute dalle entità coinvolte, ma vuole solo fornire una generale relazione associativa²⁶.

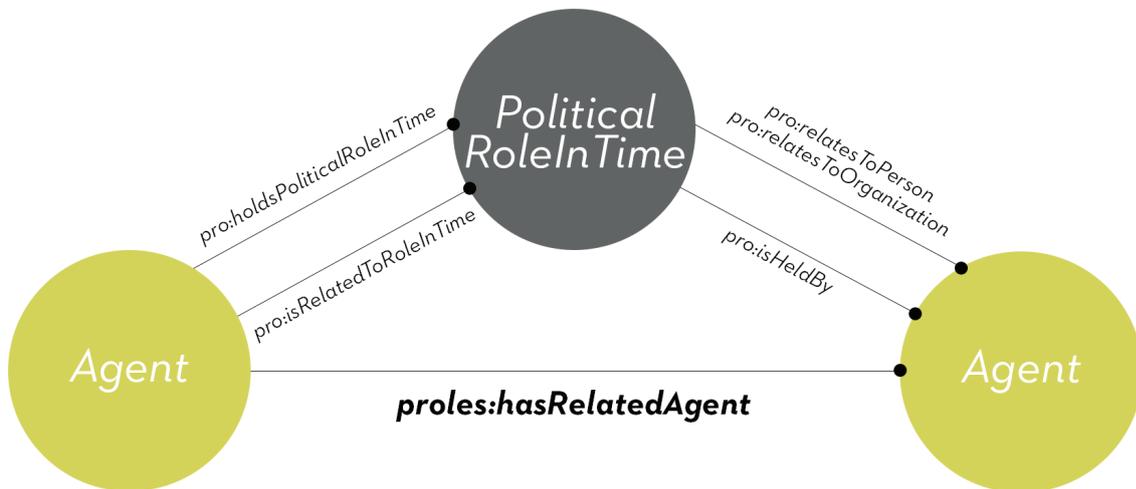


Fig. 4.6 Rappresentazione grafica delle catene di proprietà di *proles:hasRelatedAgent*.

3.4 HAS RELATED SOURCE (PROLES:HASRELATEDSOURCE)²⁷

Dominio: *foaf:Agent* (or) *proles:Place*

²⁵ Cfr. PRoles, <http://www.essepuntato.it/2013/10/politicalroles/hasRelatedAgent>

²⁶ `owl:propertyChainAxiom (<http://purl.org/spar/pro/isRelatedToRoleInTime> <http://purl.org/spar/pro/isHeldBy>) , (proles:holdsPoliticalRoleInTime <http://purl.org/spar/pro/relatesToOrganization>) ;`
`rdfs:range [NS5:type owl:Class ; owl:unionOf (foaf:Organization foaf:Person)] ;`
`owl:propertyChainAxiom (proles:holdsPoliticalRoleInTime <http://purl.org/spar/pro/relatesToPerson>) ;`
`rdfs:domain [NS5:type owl:Class ; owl:unionOf (proles:Place foaf:Agent)] .`

²⁷ Cfr. PRoles, <http://www.essepuntato.it/2013/10/politicalroles/hasRelatedSource>

Codominio: *proles:Source*

La proprietà relaziona agenti e fonti che attestano informazioni inerenti alle loro relazioni politiche. Tramite una catena di proprietà²⁸, viene creato un legame diretto tra le due entità, che consente di reperire più agilmente la “bibliografia” di fonti che forniscono descrizioni di un determinato *foaf:Agent*.

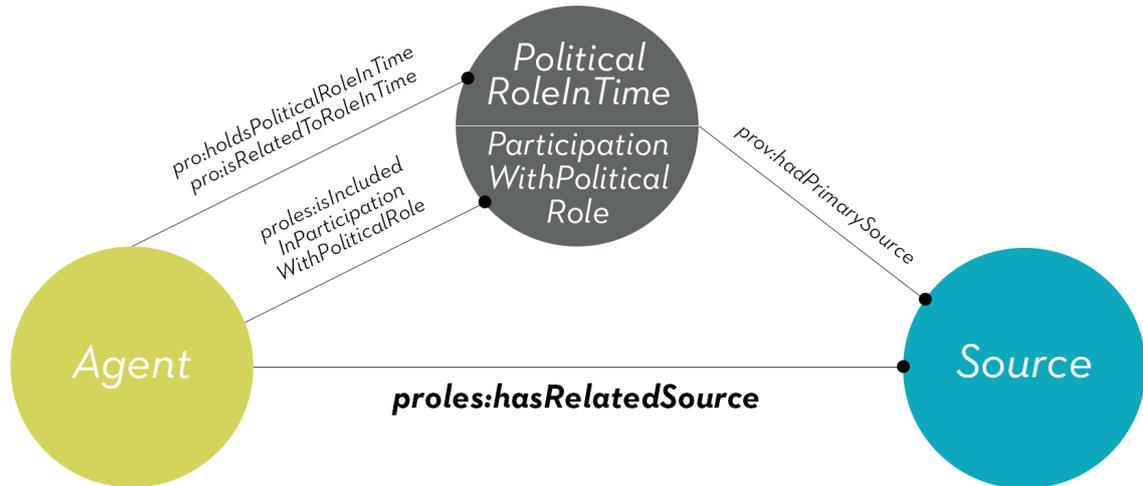


Fig. 4.7 Rappresentazione grafica delle catene di proprietà di *proles:hasRelatedSource*.

3.5 HAD PRIMARY SOURCE (PROV:HADPRIMARYSOURCE)²⁹

Dominio: *proles:PoliticalRoleInTime* (or) *proles:ParticipationWithPoliticalRole*

Codominio: *proles:Source*

Proprietà inversa: *proles:describesSituation*

La proprietà relaziona una particolare situazione ad una fonte che attesta il suo svolgimento. Ciò consente, per ogni affermazione fatta in merito a entità e rispettive relazioni, di fornire un contesto informativo da cui è stata estrapolata l'asserzione.

.....
28 owl:propertyChainAxiom
(proles:isIncludedInParticipationWithPoliticalRole prov:hadPrimarySource) ,
(<http://purl.org/spar/pro/isRelatedToRoleInTime> prov:hadPrimarySource) ,
(proles:holdsPoliticalRoleInTime prov:hadPrimarySource);
rdfs:domain [NS5:type owl:Class ; owl:unionOf (proles:Place foaf:Agent)] .

29 Cfr. PROV-O, <http://www.w3.org/TR/prov-o/hadPrimarySource>

3.6 WAS ATTRIBUTED TO (PROV:WASATTRIBUTEDTO)³⁰

Dominio: *proles:PoliticalRoleInTime* (or) *proles:PoliticalRoleInContext*

Codominio: *foaf:Agent*

La proprietà definisce insieme a *prov:hadPrimarySource* la provenance di un'affermazione, relazionando una situazione ad un responsabile delle asserzioni fatte in merito alle relazioni politiche incluse nella situazione.

3.7 HAS SETTING (SITUATION:HASSETTING)³¹

Dominio: *owl:Thing*

Codominio: *situation:Situation*

La proprietà relaziona entità e situazioni. Le sue sottoproprietà sono specializzazioni delle relazioni intercorribili tra un soggetto qualsiasi (*owl:Thing*) e un contesto in cui ha altre relazioni.

Le tre principali relazioni incluse in *situation:hasSetting* sono:

- *tisit:hasTimeIndexedSetting* da cui discendono le proprietà per definire *proles:PoliticalRoleInTime*;
- *naryparticipation:isIncludedInParticipation* da cui derivano le sottoproprietà per definire *proles:ParticipationWithPoliticalRole*;
- *tisit:isTimeIndexFor* per indicizzare l'arco temporale della situazione.

PROles estende queste proprietà con ulteriori sottoproprietà, al fine di dichiarare i rapporti politici come relazioni tipizzate.

3.7.1 holds political role in time (*proles:holdsPoliticalRoleInTime*)³²

Dominio: *foaf:Agent*

Codominio: *proles:PoliticalRoleInTime*

Proprietà inversa: *pro:holdsRoleInTime*

Sotto-proprietà di: *tisit:hasTimeIndexedSetting*

.....

³⁰ Cfr. PROV-O, <http://www.w3.org/TR/prov-o/wasAttributedTo>

³¹ Cfr. Situation, <http://www.ontologydesignpatterns.org/cp/owl/situation.owl#hasSetting>

³² Cfr. PROles, <http://www.essepuntato.it/2013/10/politicalroles/holdsPoliticalRoleInTime>

La proprietà relaziona un agente ad un ruolo politico indicizzato temporalmente e calato in un contesto semplice. Un'entità può detenere più ruoli in uno stesso arco temporale o uno stesso ruolo in contesti diversi: ciò consente una maggiore articolazione diacronica dell'attività di un soggetto descritto nei vari contesti in cui interviene.

3.7.2 *is related to role in time (pro:isRelatedToRoleInTime)*³³

Dominio: *owl:Thing*

Codominio: *pro:RoleInTime*

Proprietà inversa: *pro:relatesToEntity*

Sotto-proprietà di: *tisit:hasTimeIndexedSetting*

La proprietà relaziona una generica entità ad un ruolo indicizzato nel tempo. Questa relazione attribuisce un contesto all'esercizio del ruolo, sia esso politico o editoriale: sono previsti quattro tipologie di contesto semplice in cui un ruolo si esplica, rappresentate da quattro sottoproprietà di *pro:isRelatedToRoleInTime*.

3.7.2.1 *is document context for (pro:isDocumentContextFor)*³⁴

Dominio: *foaf:Document*

Proprietà inversa: *pro:relatesToDocument*

Sotto-proprietà di: *tisit:hasTimeIndexedSetting, pro:isRelatedToRoleInTime*

La proprietà relaziona un documento ad un ruolo indicizzato temporalmente. Nell'accezione originaria di PRO, questa relazione comporta un rapporto di paternità nei confronti del documento, rivestita dal soggetto che detiene un ruolo editoriale sul documento in questione; in PROles è possibile utilizzare la relazione anche per rappresentare una attestazione semplice del ruolo in una fonte, nel caso di una minore esigenza di articolazione del contesto per i ruoli, senza dover ricorrere al meta-contesto ideato per specificare la *provenance* delle asserzioni.

.....
33 Cfr. PRO, <http://purl.org/spar/pro/isRelatedToRoleInTime>

34 Cfr. PRO, <http://purl.org/spar/pro/isDocumentContextFor>

3.7.2.2 *is organization context for (pro:isOrganizationContextFor)*³⁵

Dominio: *foaf:Organization*

Proprietà inversa: *pro:relatesToOrganization*

Sotto-proprietà di: *tisit:hasTimeIndexedSetting*, *pro:isRelatedToRoleInTime*

La proprietà relaziona un'organizzazione ad un ruolo indicizzato temporalmente: per organizzazione si intende il contesto istituzionale in cui viene esercitato il ruolo descritto.

3.7.2.3 *is personal context for (pro:isPersonalContextFor)*³⁶

Sotto-proprietà di: *tisit:hasTimeIndexedSetting*, *pro:isRelatedToRoleInTime*

La proprietà relaziona un agente ad un ruolo indicizzato temporalmente: il ruolo viene esercitato nei confronti di un agente che fornisce un contesto alla relazione. Nel caso delle relazioni politiche, una persona può rivestire un ruolo politico nei confronti di un'altra, in un rapporto gerarchico, oppure può intrattenervi una relazione politica, mediante un legame associativo.

3.7.2.4 *is place context for (proles:isPlaceContextFor)*³⁷

Dominio: *proles:Place*

Proprietà inversa: *pro:relatesToPlace*

Sottoproprietà di: *tisit:hasTimeIndexedSetting*, *pro:isRelatedToRoleInTime*

La proprietà relaziona un luogo ad un ruolo indicizzato temporalmente: il luogo è l'ambito, volutamente generico, in cui viene esercitato un ruolo o intrattenuta una relazione politica.

.....

35 Cfr. PRO, <http://purl.org/spar/pro/isOrganizationContextFor>

36 Cfr. PRO, <http://purl.org/spar/pro/isPersonalContextFor>

37 Cfr. PROles, <http://www.essepuntato.it/2013/10/politicalroles/isPlaceContextFor>

3.7.3 *is political role in (proles:isPoliticalRoleIn)*³⁸

Dominio: *proles:PoliticalRole*

Codominio: *proles:PoliticalRoleInTime*

Proprietà inversa: *proles:withPoliticalRole*

Sotto-proprietà di: *tisit:hasTimeIndexedSetting, pro:isRoleIn*

La proprietà relaziona uno specifico ruolo politico alla situazione in cui questo viene esercitato da un soggetto, calato in un contesto e indicizzato temporalmente.

3.7.4 *is included in participation with political role (proles:isIncludedInParticipationWithPoliticalRole)*³⁹

Dominio: *foaf:Agent*

Codominio: *proles:ParticipationWithPoliticalRole*

Proprietà inversa: *proles:includesAgent*

Sotto-proprietà di: *nary:isIncludedInParticipation*

La proprietà relaziona un agente ad una situazione complessa in cui viene esercitato un ruolo politico indicizzato temporalmente all'interno di un evento, a sua volta definito nel tempo. Un agente partecipa ad un evento in veste del (ed insieme al) suo ruolo politico indicizzato ed è inserito in un contesto più ampio che fornisce maggiori informazioni circostanziali sull'espletamento del ruolo.

3.8 IS SETTING FOR (SITUATION:ISSETTINGFOR)⁴⁰

Dominio: *situation:Situation*

Codominio: *owl:Thing*

Proprietà inversa: *situation:hasSetting*

La proprietà relaziona situazioni ad altre entità. Le sue sotto-proprietà sono specializzazioni delle relazioni intercorribili tra un contesto d'azione e i soggetti

.....

38 Cfr. PRoles, <http://www.essepuntato.it/2013/10/politicalroles/isPoliticalRoleIn>

39 Cfr. PRoles, <http://www.essepuntato.it/2013/10/politicalroles/isIncludedInParticipationWithPoliticalRole>

40 Cfr. Situation, <http://www.ontologydesignpatterns.org/cp/owl/situation.owl#isSettingFor>

(*owl:Thing*) che in esso si relazionano. Le tre principali relazioni incluse in *situation:isSettingFor* sono:

- *tvc:atTime* per l'indicizzazione temporale di una situazione;
- *tisit:forEntity* che relaziona la situazione ad agenti, contesti e ruoli;
- *naryparticipation:participationIncludes* per dichiarare le entità coinvolte in un evento ed in situazioni complesse.

3.8.1 *at time* (*tvc:atTime*⁴¹, *tisit:atTime*⁴²)

Dominio: *tisit:TimeIndexedSituation*, *tvc:ValueInTime*

Codominio: *time:TemporalEntity*, *ti:TimeInterval*

Proprietà inversa: *situation:hasSetting*

Sotto-proprietà di: *situation:isSettingFor*

Le proprietà relazionano un'entità temporale ad una specifica situazione. Anche se non espressamente dichiarato nell'allineamento dei due modelli, sono l'una (*tvc:atTime*) sotto-proprietà dell'altra (*tisit:atTime*), per cui il loro utilizzo è simile.

3.8.2 *is held by* (*pro:isHeldBy*)⁴³

Dominio: *pro:RoleInTime*

Codominio: *foaf:Agent*

Proprietà inversa: *pro:holdsRoleInTime*

Sotto-proprietà di: *situation:isSettingFor*, *tisit:forEntity*

La proprietà relaziona un ruolo indicizzato temporalmente ad un agente che esercita un preciso ruolo, politico o editoriale, circostanziandolo.

.....

41 Cfr. TVC, <http://www.essepuntato.it/2012/04/tvc/atTime>

42 Cfr. Time Indexed Situation, <http://www.ontologydesignpatterns.org/cp/owl/timeindexedsituation.owl#atTime>

43 Cfr. PRO, <http://purl.org/spar/pro/isHeldBy>

3.8.3 *relates to entity (pro:relatesToEntity)*⁴⁴

Dominio: *pro:RoleInTime*

Proprietà inversa: *pro:isRelatedToRoleInTime*

Sotto-proprietà di: *situation:isSettingFor, tv:withinContext*

La proprietà relaziona un ruolo indicizzato temporalmente ad altre entità coinvolte o utili a definire il contesto di espletamento del ruolo. Sono previste quattro tipologie di entità (documenti, organizzazioni, persone e luoghi) e rispettivamente quattro sotto-proprietà di *pro:relatesToEntity* per specificarle, equivalenti alle proprietà inverse delle sotto-proprietà di *pro:isRelatedToRoleInTime*⁴⁵.

3.8.4 *with role (proles:withRole)*⁴⁶

Dominio: *pro:RoleinTime*

Codominio: *pro:Role*

Proprietà inversa: *pro:isRoleIn*

Sotto-proprietà di: *situation:isSettingfor*

La proprietà relaziona una specifica situazione indicizzata temporalmente al ruolo esercitato in essa. Nel caso delle relazioni politiche viene esplicitata una restrizione universale sulla proprietà affinché abbia solo ruoli politici come codominio.

3.8.5 *includes agent (proles:includesAgent)*⁴⁷

Dominio: *proles:ParticipationWithPoliticalRole*

Codominio: *foaf:Agent*

Proprietà inversa: *proles:isIncludedInParticipationWithPoliticalRole*

Sotto-proprietà di: *situation:isSettingFor*

.....

44 Cfr. PRO, <http://purl.org/spar/pro/relatesToEntity>

45 Cfr. note 19,20,21,22,23

46 Cfr. PRO, <http://www.essepuntato.it/2013/10/politicalroles/withRole>

47 Cfr. PROles, <http://www.essepuntato.it/2013/10/politicalroles/includesAgent>

La proprietà relaziona una situazione complessa ad un agente: in questa un agente esercita un ruolo politico indicizzato temporalmente a sua volta calato all'interno di un evento circoscritto nel tempo.

3.8.6 *participation includes political role in time*
(*proles:participationIncludesPoliticalRoleInTime*)⁴⁸

Dominio: *proles:ParticipationWithPoliticalRole*

Codominio: *proles:PoliticalRoleInTime*

Sotto-proprietà di: *situation:isSettingFor, naryparticipation:participationIncludes*

La proprietà relaziona un ruolo politico contestualizzato ad un ruolo politico indicizzato temporalmente e afferente ad un contesto “minore”. Ad un agente viene attribuito un ruolo politico indicizzato nel tempo (*proles:PoliticalRoleInTime*), definito da soggetto – ruolo – arco di tempo – contesto, il quale è “partecipe” insieme all'agente di una situazione più ampia all'interno di un evento (*proles:ParticipationWithPoliticalRole*), unitamente ad altri fattori che descrivono uno scenario complesso per contestualizzare ruoli o relazioni politiche.

4. DATAPROPERTY E ANNOTATION PROPERTY

PRoles, come già detto, riusa le proprietà descrittive degli element set che importa, e nel caso di data property e annotation property non necessita per le sue esigenze rappresentative di ulteriori proprietà. Nello specifico, le data property fondamentali riusate appartengono a:

- FOAF per fornire la descrizione testuale, o forme controllate degli elementi, rappresentativi degli agenti;
- Time e TimeInterval per una corretta definizione degli intervalli di tempo in cui sono circoscritte le situazioni.

Le annotation property consigliate sono invece quelle fornite da di owl, rdfs, DC e DCterms.

.....

48 Cfr. PRoles, <http://www.essepuntato.it/2013/10/politicalroles/participationIncludesPoliticalRoleInTime>

5. APPLICAZIONI ED ESEMPI⁴⁹

PROles può essere utilizzato come modello per arricchire semanticamente la rappresentazione di entità che sono coinvolte nella descrizione dei metadati di testi presenti sul web.

Il campo d'indagine da cui è partita la ricerca di casi di studio e declinazioni possibili del modello riferisce all'ambito archivistico informatico, dove però solo di rado vengono considerate singole entità documentali nella descrizione dei complessi archivistici e dei soggetti produttori d'archivio: perciò il modello è stato incrementato prevedendo la possibilità di più livelli descrittivi dei contesti documentali, a seconda della provenienza della fonte interrogata, della sua "mole" (sia essa un singolo documento, un fascicolo, un libro...) e in base al suo livello di formalizzazione (una risorsa metadatata o non). Di fatto, è possibile definire una "fonte" come una qualsiasi entità, presente o meno sul web, che fornisce qualche informazione sul soggetto descritto: ciò fa sì che non vi sia la necessità di avere descrizioni formali per le risorse che attestano notizie sui ruoli politici o che esse siano esclusivamente materiale d'archivio piuttosto che di altra natura e supporto. Questo approccio permette potenzialmente di estendere a 360° il tipo di sorgenti collegabili ed ampliare il raggio di relazionalità delle entità descritte nell'ambito del *cultural heritage*.

Cambiando prospettiva, ovvero partendo dalla descrizione degli agenti, in particolar modo delle persone, PROles restituisce una maggiore circoscrizione informativa delle loro "azioni": fornendo un approccio che si focalizza sul contesto dell'espletamento dei ruoli e dei rapporti politici e sul meta-contesto della bibliografia che fornisce queste informazioni, viene arricchita la descrizione di un soggetto produttore (o non) d'archivio, e ampliata la possibilità di aprire ulteriormente i dati alla rete dei Linked Open Data. Da questa angolazione, PROles può essere visto come uno strumento ausiliare per modelli che già approfondiscono le descrizioni dei soggetti culturali, a cui aggregare relazioni aggiuntive, ai fini dell'integrazione dei dati di un determinato dominio di conoscenza (qui i dati archivistici) con altri dataset.

I casi d'uso qui forniti sono tratti da uno studio, al momento solo teorico, sui

⁴⁹ Cfr. PROles, "Use case", <http://www.semanticweb.org/marilena/ontologies/2013/9/usecase>

dati riversati in x-Dams da IBC Emilia-Romagna e Regesta.exe inerenti le carte di Andrea Costa. Per semplicità le fonti considerate sono native sul web e per una maggiore linearità ed intuitività sono ristrette a sole pagine di Wikipedia.

5.1 RELAZIONI POLITICHE TRA PERSONE

L'utilizzo della classe *proles:PoliticalRoleInTime* permette di collegare più entità mediante una relazione indicizzata nel tempo. Seguendo questa modalità è possibile dare una rappresentazione formale della seguente asserzione: "Anna Kulisoff è stata una persona influente nei confronti di Andrea Costa tra il 1878 ed il 1881." Questa affermazione, di cui si vuole fornire un contesto di provenienza mediante il riuso delle proprietà di PROV-O, è tratta dalla pagina di Wikipedia dedicata ad Andrea Costa ed è stata estratta da Marilena Daquino.

Il risultato che restituisce PRoles formalizzando queste asserzioni mediante le sue proprietà è:

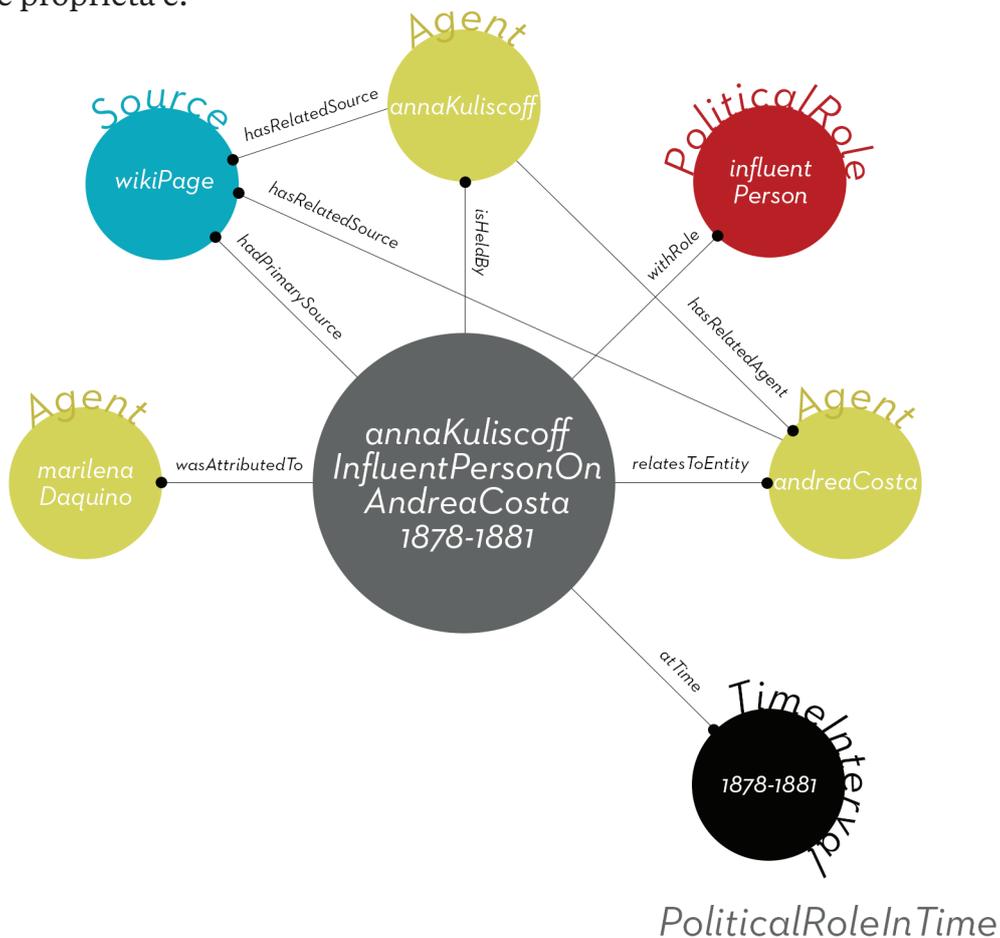


Fig. 4.8 Rappresentazione grafica della classe *proles:PoliticalRoleInTime* applicata alle relazioni politiche.

Da notare, nonostante vi sia un nodo intermedio tra gli individui di *foaf:Agent* e tra gli agenti e le fonti (ossia il nodo fornito dalla classe *PoliticalRoleInTime*), il modello restituisce come inferenze le proprietà generiche *proles:hasRelatedAgent* e *proles:hasRelatedSource*; queste non sono altro che gli speculari legami associativi adottati dalla maggior parte dei modelli per relazionare le descrizioni di persone e risorse collegate.

Una maggiore definizione delle relazioni tramite nodi intermedi che le contestualizzano non comporta, come si potrebbe pensare, uno sforzo ulteriore di implementazione rispetto alle generali (o generiche) relazioni normalmente indagate dai più comuni modelli; anzi, PRoles ottiene gli stessi risultati sfruttando i meccanismi dell'inferenza, richiedendo minor sforzo interpretativo da parte dell'utente e fornendo al contempo maggiori informazioni.

5.2 RUOLI POLITICI INDICIZZATI

Allo stesso modo delle relazioni politiche tra persone, è possibile definire un ruolo politico inerente un singolo *foaf:Agent*, fornendo come contesto un luogo, un documento o un'organizzazione di riferimento. Ad esempio: "Andrea Costa è stato imprigionato ad Imola nel 1874." Questa asserzione è stata estratta dallo stesso contesto della precedente.

Come si può vedere PRoles restituisce per la pagina di Wikipedia tutte e due le situazioni relazionate. È possibile quindi, implementando ulteriormente le relazioni, arrivare a metadattare molteplici informazioni ricavate dall'analisi *full text* della fonte.

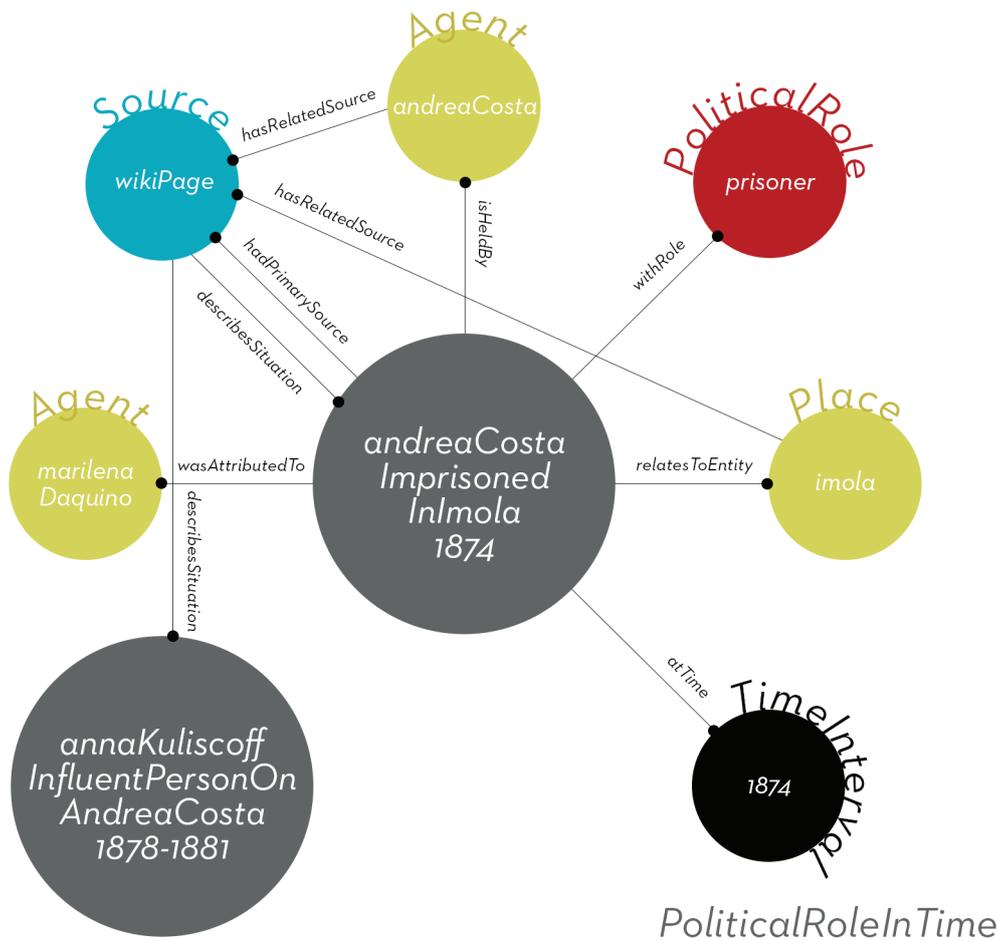
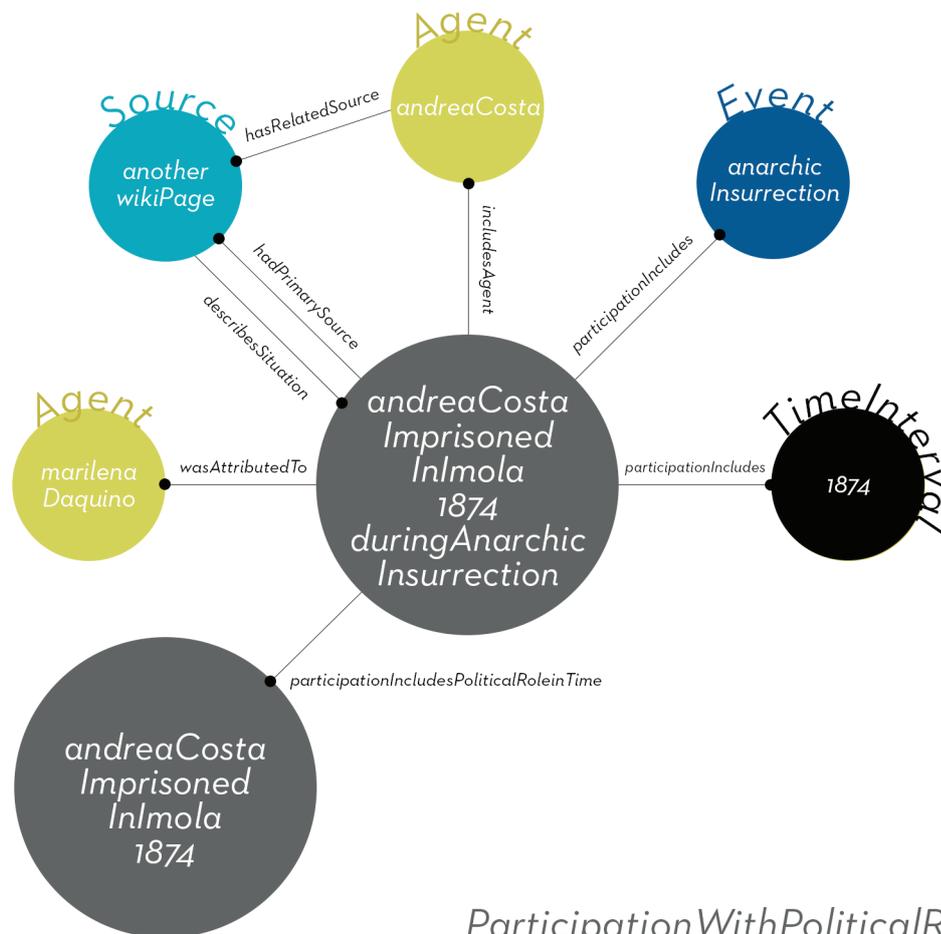


Fig. 4.9 Rappresentazione grafica della classe *proles:PoliticalRoleInTime* applicata ai ruoli politici.

5.3 RELAZIONI E RUOLI POLITICI IN UN CONTESTO COMPLESSO

Spesso accade che il contesto in cui si avvengono le relazioni e i ruoli politici sia più articolato di un rapporto con una singola persona o dello svolgimento di un ruolo in un luogo. Perciò è sorta l'esigenza di ampliare gli scenari di inclusione dei ruoli politici indicizzati utilizzando la classe *proles:ParticipationWithPoliticalRole*.

Riprendendo l'asserzione precedente in un contesto più ampio si può ottenere: "Andrea Costa è stato imprigionato nel 1874 a Imola durante un'insurrezione anarchica a cui ha partecipato nel corso dello stesso anno." La fonte da cui è stata estratta l'informazione è un'altra pagina Wikipedia, e l'affermazione è fornita dallo stesso responsabile dell'asserzione precedente.



ParticipationWithPoliticalRole

Fig. 4.10 Rappresentazione grafica della classe *proles:ParticipationWithPoliticalRole*.

Il meccanismo per la definizione della classe *proles:ParticipationWithPoliticalRole* prevede l'attestazione della *provenance*, come per *proles:PoliticalRoleInTime*, ma non vi è alcun vincolo che la fonte o il responsabile siano gli stessi già inerenti l'asserzione sul "ruolo nel tempo" coinvolto nell'evento.

È possibile così effettuare precisazioni su informazioni già esistenti (*political role in time*) ampliandone lo scenario di svolgimento: è un'informazione complessa incrementabile da parte di più estrattori di informazioni ed è consentito integrare più fonti per definire una situazione articolata.

Un simile approccio permette la convivenza di costrutti che possono confluire in un "dibattito" sulla formalizzazione di attestazioni plurime di informazioni, consentendo la compresenza di asserzioni che possono contraddirsi e comunque arricchire la descrizione della risorsa.

6. PROLES ED EAC-CPF ONTOLOGY

La formalizzazione di PROles muove da una riflessione sulle capacità descrittive offerte dall'archivistica: l'ontologia EAC-CPF è stata di conseguenza un metro di paragone per lo sviluppo di nuove relazioni sinora trascurate o solo parzialmente indagate dalla comunità archivistica, cercando di dare un contributo alla discussione sulla creazione di un modello concettuale archivistico.

Partendo dalla mancanza strutturale nel dibattito italiano di una descrizione formale delle funzioni per gli enti produttori di complessi archivistici e soffermandosi sul ruolo delle persone a cui afferiscono risorse collegate di varia natura, l'idea che ha mosso il progetto di PROles è quella di dare vita ad una rete di relazioni che colmi le lacune dell'assenza di uno schema formale per ISDF e che integri queste specifiche direttamente nella descrizione di un soggetto produttore, così come era previsto nello schema XSD EAC-CPF.

PROles non vuole sostituirsi ad un'ipotetica formalizzazione di ISDF, ma vuole far emergere nel dibattito la possibilità di valutare altre strade per l'aggregazione delle informazioni partendo dalla centralità dei soggetti produttori, dalle loro descrizioni, e proponendo un loro arricchimento tramite relazioni tipizzate che aprano alla rete di dati ulteriori informazioni e forniscano *access points* utili all'utente finale.

Il confronto con EAC-CPF Ontology è stato propedeutico allo studio di fattibilità e di puntualizzazione degli obiettivi minimi necessari all'approfondimento delle necessità degli *open archival data*.

PROles propone innanzitutto un approccio basato sul riuso di modelli, da integrare e allineare per fornire una rappresentazione efficace del suo dominio di conoscenza, quindi un'analisi dell'esistente e degli espedienti utilizzati dagli altri modelli per il *cultural heritage*, al fine di non reiterare formalizzazioni simili e ridondanti e condividere un punto di vista sulle risorse.

Inoltre fa uso di metodi inferenziali, affinché le potenzialità espressive del modello vengano sfruttate il più possibile e perché questo possa aderire maggiormente ad uno preesistente, con un minore sforzo interpretativo da parte di chi lo implementa e di chi ne usufruisce.

6.1 L'ALLINEAMENTO TRA PROLES E EAC-CPF ONTOLOGY⁵⁰

Nell'allineamento tra PRoles e EAC-CPF Ontology si è optato per una mappatura di pochi ma fondamentali elementi, essendo la prima ontologia autonoma ed interpretabile come un'integrazione più che un modello speculare del secondo in cui stabilire relazioni gerarchiche o di equivalenza.

La classe *eac-cpf:entity* è stata resa equivalente alla classe *foaf:Agent* utilizzata in PRoles, così da inferire le relazioni tra soggetti produttori (e non) all'interno di un'unica classe che aggrega tutte quelle entità che hanno agito o sono state “chiamate in causa” all'interno di una fonte che le cita.

Le object properties *eac-cpf:resourceRelation* e *eac-cpf:cpfRelation* sono state dichiarate super-proprietà delle rispettive *proles:hasRelatedSource* e *proles:hasRelatedAgent*, al fine di inferire le relazioni indagate da EAC come imparentate con quelle di PRoles e stabilire dei punti di contatto tra i due modelli.

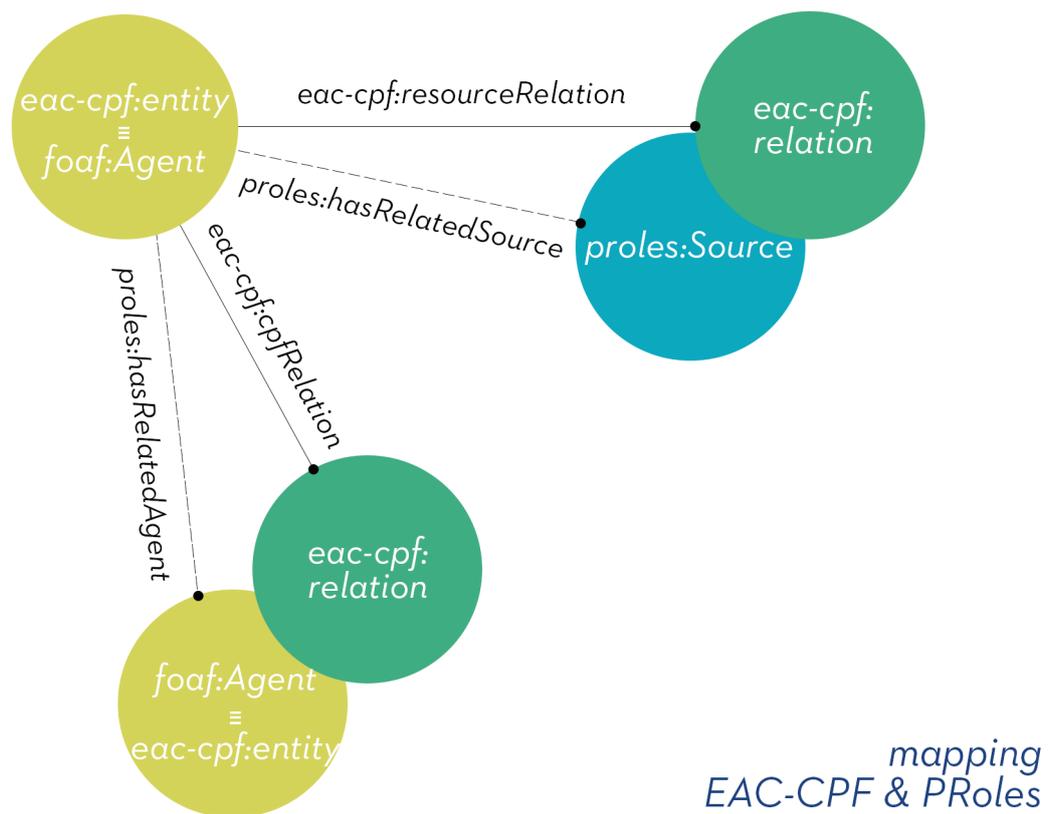


Fig. 4.11 Rappresentazione grafica del mapping tra EAC-CPF e PRoles.

.....
⁵⁰ Cfr. M. Daquino, S. Peroni, F. Tomasi, "Mapping EAC-CPF and PRoles",
<http://www.semanticweb.org/marilena/ontologies/2013/10/mappingEAC-PROLES/>;

Non è stata, come ci si potrebbe aspettare, formalizzata in PRoles una proprietà speculare alla mancante “*eac-cpf:functionRelation*”: in EAC-CPF Ontology questa relazione esiste sotto forma di data property, *eac-cpf:function*, mentre in PRoles la scelta progettuale ha portato alla definizione delle classi *proles:PoliticalRoleInTime* e *proles:ParticipationWithPoliticalRole* come luogo più proficuo dove sviluppare il ragionamento sulle funzioni/ruoli delle entità come medium tra esse e le risorse collegate.

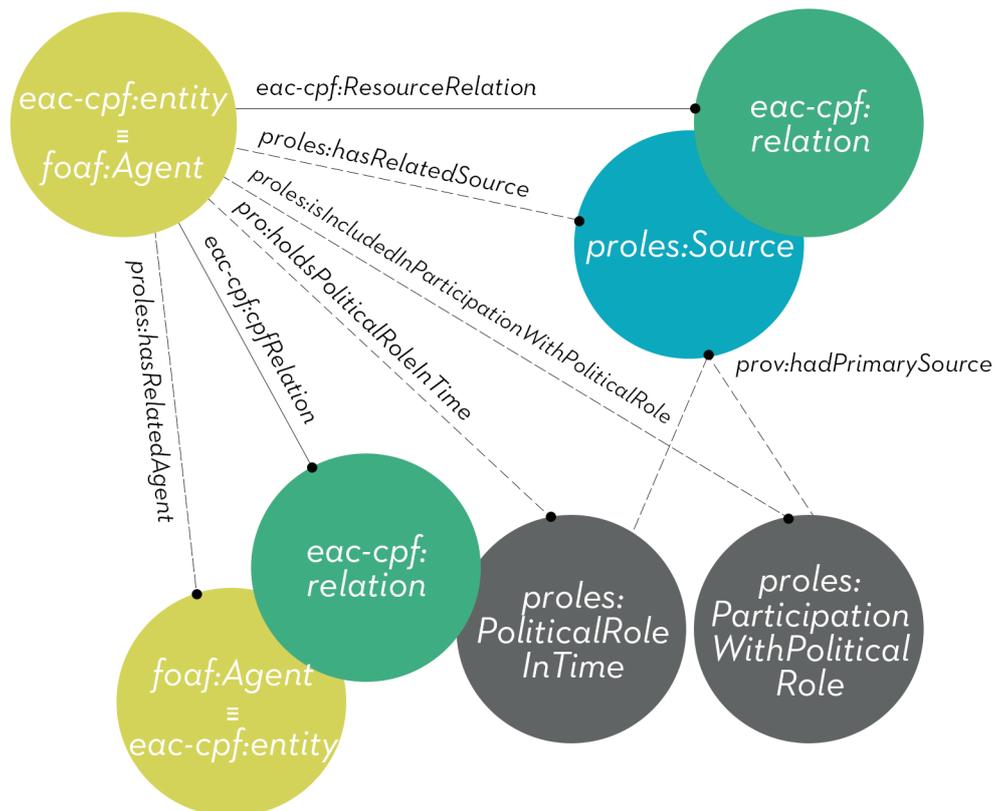


Fig. 4.12 Rappresentazione grafica dei ruoli politici come possibili sottoclassi di *eac-cpf:relation*.

Le classi principali di PRoles potenzialmente sono sottoclassi di *eac-cpf:relation*, così come lo sono *eac-cpf:entity* e le risorse collegate nel momento in cui esiste una object property per definire le relazioni che intercorrono tra loro e una data entità. Le proprietà *pro:holdsRoleInTime* e *proles:isIncludedInParticipationWithPoliticalRole* potrebbero essere interpretate come sotto-proprietà di una generica relazione tra entità e funzioni, nel momento in cui l'ontologia di EAC-CPF prevedesse una revisione delle sue proprietà e definisse delle object properties per rappresentare le funzioni.

La soluzione adottata in PRoles prevede, come vuole lo standard per la descrizione delle funzioni ISDF, che il legame tra un soggetto produttore e una sua risorsa collegata sia mediato dalla funzione/ruolo che ha indotto l'entità a produrre quella documentazione; in PRoles l'accento viene spostato su una considerazione più storiografica che archivistica, poiché il ruolo in questione non è il mordente che ha mosso il soggetto a produrre il documento, ma un'attestazione di un'informazione sulla natura del soggetto desunta da una fonte.

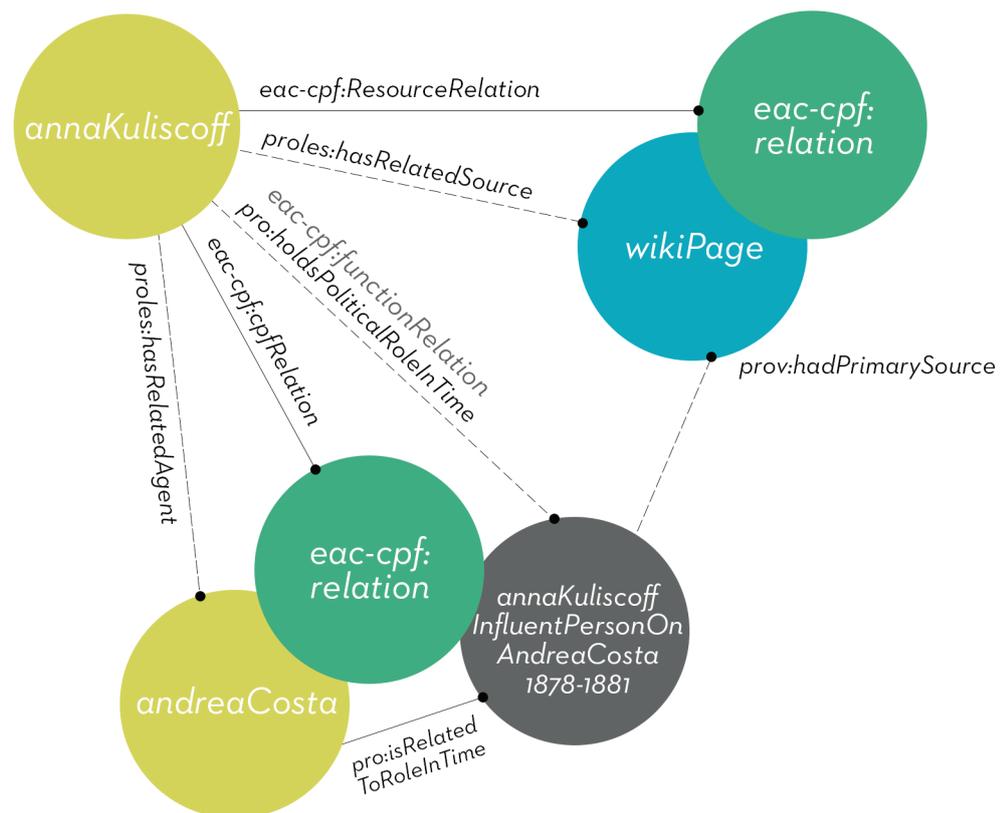


Fig. 4.13 Rappresentazione grafica di un'applicazione combinata di PRoles e EAC-CPF Ontology.

Anche se l'applicazione diverge, il principio alla base della formalizzazione è sostanzialmente riproponibile per EAC-CPF e la metodologia, ciò che contava far emergere in questo progetto, è che si può utilizzare in maniera notevolmente espressiva le capacità della modellizzazione ontologica per arrivare a descrivere ogni asserzione rilevante: così come si possono descrivere i ruoli di una persona attestati in una fonte si possono descrivere, con i medesimi predicati ontologici, le funzioni/attività/ruoli/occupazioni dei soggetti produttori che hanno generato una documentazione archivistica, semplicemente implemen-

tando nuove istanze (e quindi forme controllate dei nomi) per definire la tipologia di ruolo che si vuole attribuire ad un soggetto.

Infine, ciò a cui si vuole dare risalto in questa sede, è la possibilità di descrivere situazioni complesse, come l'esercizio di ruoli in un contesto, rappresentandole non solo come proprietà unilaterali che collegano genericamente un'entità ad un ruolo/funzione, ma istanziando delle classi che aggregano molte più informazioni al riguardo e ne forniscono uno scenario descrittivo completo.

LEGENDA

| | | |
|---|--|---|
|  | <i>owl:Class</i> | Rappresentazione grafica di una classe; all'interno del cerchio è riportato il nome della classe. |
|  | <i>rdfs:subClassOf</i> | Rappresentazione grafica di una sottoclasse; all'interno del cerchio è inscritto il nome della sottoclasse appartenente alla classe riportata sopra. |
|  | <i>rdf:resource</i> | Rappresentazione grafica di un individuo; negli esempi applicativi, il nome interno al cerchio rappresenta una istanza della classe contrassegnata precedentemente dallo stesso colore. |
|  | <i>owl:ObjectProperty</i> | Rappresentazione grafica di una object property; il codominio della proprietà è indicato dalla direzione della retta. |
|  | <i>rdfs:subPropertyOf</i> | Rappresentazione grafica di una sottoproprietà; è utilizzata negli esempi di mapping tra ontologie per indicare quale proprietà è stata allineata ad una di un altro dominio con un rapporto gerarchico. |
| <hr/> | | |
|  | <i>foaf:Agent; proles:Place</i> | |
|  | <i>proles:PoliticalRole</i> | |
|  | <i>ti:TimeInterval</i> | |
|  | <i>proles:PoliticalRoleInTime; proles:ParticipationWithPoliticalRole</i> | |
|  | <i>participation:Event</i> | |
|  | <i>proles:Source</i> | |
|  | <i>eac-cpf:relation</i> | |

BIBLIOGRAFIA

Agosti M., Esposito E., Ferilli S., Ferro N., *Digital Libraries and Archives, 8th Italian Research Conference, IRCDL 2012. Bari, Italy, February 2012. Revised Selected Papers*, Berlin-Heidelberg, Springer, 2013

Argelli B., *Una regione in rete: gli archivi storici dell'Emilia-Romagna e il progetto IBC*, in «DigItalia, Rivista del digitale nei beni culturali», III, 2, 2008, <http://digitalia.sbn.it/article/view/449>

Barbanti I., *La codifica informatica dei dati di archivio: EAD ed EAC*, in "Scrinia", 2-3/2005, http://media.regesta.com/dm_o/regesta/regestacms/generic/documentazione/Barbanti-EADEAC2005.pdf

Bearman D., *Authority Control Issues and Prospects*, in «American Archivist», Vol. 52, 1989, <http://archivists.metapress.com/content/g562600um1063123/fulltext.pdf>

Bountouri L., Gergatsoulis M., "Mapping Encode Archival Description in CIDOC-CRM", <http://eprints.rclis.org/15844/1/02.Bountouri.pdf>

Brenneke F., *Archivistica. Contributo alla teoria ed alla storia archivistica europea*, Milano, Giuffrè, 1968, http://www.icar.beniculturali.it/biblio/_view_volume.asp?ID_VOLUME=62

Carucci P., Guercio M., *Manuale di archivistica*, Carocci, 2008

Crupi G., *Oltre le colonne d'Ercole: Linked Data e cultural heritage*, in «JLIS», 4, n.1, 2013, p. 9, <http://leo.cilea.it/index.php/jlis/article/view/8587/7888>

Davies J., Studer R., Varren P., *Semantic Web Technologies: Trends and Research in Ontology-based Systems*, Wiley Ed., 2006

Fazio (di) S., “Web semantico e descrizione archivistica:un applicativo semantico per gli archivi”,
http://www.san.beniculturali.it/web/san/dettaglio-notizia-san?p_p_id=56_INSTANCE_X7Qi&articleId=2237879&p_p_lifecycle=1&p_p_state=normal&groupId=10704&viewMode=normal

Guercio M., *Archivistica informatica. I documenti in ambiente digitale*, Carocci, 2010

Guerrini M., Tillett B. B. (a cura di), *Authority Control: definizione ed esperienze internazionali. Atti del convegno internazionale, Firenze, 10-12 Febbraio 2003*, Firenze, University Press, 2003

Guerrini M., Possemato T., “Linked data: un nuovo alfabeto del web semantico”,
<http://www.bibliotecheoggi.it/content/201200300701.pdf>

Giuva L. (a cura di), *Gli archivi storici dei partiti politici europei, Atti del convegno, Roma 13-14 dicembre 1996*, Ministero per i beni e le attività culturali, 2001,
http://www.archivi.beniculturali.it/dga/uploads/documents/Quaderno_94.pdf

Giuva L., Vitali S., Zanni Rosiello Isabella, *Il potere degli archivi. Usi del passato e difesa dei diritti nella società contemporanea*, Mondadori Bruno, 2007

International Council on Archives, CBPS - Sub-Committee on Descriptive Standards, “International Standard for Describing Functions (ISDF)”, 2011,
<http://www.ica.org/10208/standards/isdf-international-standard-for-describing-functions.html>; “ISAD (G), General International Standard Archival Description, Second Edition”, [http://www.icacds.org.uk/eng/ISAD\(G\).pdf](http://www.icacds.org.uk/eng/ISAD(G).pdf); “ISDIAH, International Standard for Describing Institutions with Archival Holdings”, <http://www.ica.org/?lid=10198>;

Mazzini S., Ricci F., *Linked archival authority data. Una sperimentazione sui linked open data del sistema informativo regionale IBC Archivi*, in «DigItalia, Rivista del digitale nei beni culturali», VI, 2, 2011,
<http://digitalia.sbn.it/article/view/484/334>

Mazzini S., Ricci F., *EAC-CPF Ontology and Linked Archival Data*, in *Proceedings of the 1st International Workshop on Semantic Digital Archives, Berlin 29/9/2011 (SDA 2011)*, <http://ceur-ws.org/Vol-801/paper6.pdf>

Ore C., "Social and family relations and the CIDOC-CRM – some thoughts", 2006, http://www.cidoc-crm.org/docs/frbr_oo/frbr_docs/meeting_presentations/10th_meeting_presentations/Fam_relation_CRM.rtf

Peroni S., Shotton D., Vitali F., *Scholarly publishing and the Linked Data: describing roles, statuses, temporal and contextual extents*, in H. Sack, T. Pellegrini, *Proceedings of the 8th International Conference on Semantic Systems*, ACM, New York, 2012, <http://speroni.web.cs.unibo.it/publications/peroni-2012-scholarly-publishing-linked.pdf>

Ricci F., *Gli standard internazionali di descrizione archivistica: dalle origini alla seconda versione di ISAD (G) e Le linee guida per l'elaborazione e la presentazione di strumenti di ricerca*, in «Rassegna degli Archivi di Stato», anno LXIII – n. 1, 2003, pp. 11-59

Sanacore M., *Documenti "seriali" politici, letteratura non convenzionale e conservazione archivistica*, in «Rassegna degli Archivi di Stato», anno LX – n. 2, 2000, pp. 362-372

Savoja M., *Descrizione e liste d'autorità: alcune proposte nordamericane*, in *Standard, vocabolari controllati, liste d'autorità. Atti del seminario, Milano, 25 maggio 1994*, Milano: Regione Lombardia, 1995

Savoja M., "Lo standard ISAAR come riferimento per la messa a punto di sistemi informativi archivistici" e "Descrizioni di archivi e soggetti produttori (ISAD-G E ISAAR-CPF)", <http://web.mclink.it/MD1431/sito/isaargrp/isaar.html>

Solodovnik I., "Comunicazione e ricerca semantica di contenuti informativi: tra metadati, Linked open Data e Ontologie", <http://eprints.rclis.org/15966/1/Metad-LOD-Ontologie.pdf>

Tomasi F., *Le edizioni digitali come nuovo modello per dati di autorità concettuali*, in «JLIS», 4, 2, 2013, p. 1, <http://leo.cilea.it/index.php/jlis/article/view/8808/8067>

Valacchi F., "I siti web come strumenti per la ricerca", <http://eprints.rclis.org/11683/1/wearchivi.pdf>

Vassallo S., *Descrizioni archivistiche e web semantico: un connubio possibile?*, in «JLIS», 1, n.1, 2010, <http://leo.cilea.it/index.php/jlis/article/view/27/37>

Vassallo S., *Frammenti semantici. Riflessioni su descrizioni archivistiche e web semantico: il caso dell'archivio Giovanni Testori*, Tesi di dottorato in Scienze bibliografiche, archivistiche e documentarie e per la conservazione e restauro dei beni librari ed archivistici, Università di Udine, A.A. 2009/2010

Vitali S. (trad. italiana di), ISAAR(CPF): Standard internazionale per i record d'autorità archivistici di enti, persone, famiglie, traduzione italiana di International Standard Archival Authority Records for Corporate Bodies, Persons and Families, seconda edizione, 2004, in «Rassegna degli Archivi di Stato», 2003, pp. 191-333, <http://www.icar.beniculturali.it/biblio/pdf/standard/isaar%202.pdf>

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

Zanni Rosiello I. (a cura di), *Intorno agli archivi e alle istituzioni. Scritti di Claudio Pavone*, Ministero per i beni e le attività culturali, 2004

SITOGRAFIA

APEnet, Archives Portal Europe,

<http://www.archivesportaleurope.net/web/guest;jsessionid=A4BEE3118A1158AE6BAF8027C67B7544;>

BIO: a vocabulary for biographical information,

[http://vocab.org/bio/0.1/.html;](http://vocab.org/bio/0.1/.html)

CIDOC-CRM,

[http://www.cidoc-crm.org/;](http://www.cidoc-crm.org/)

DBpedia,

[http://dbpedia.org/About;](http://dbpedia.org/About)

DC, Dublin Core,

[http://dublincore.org/;](http://dublincore.org/)

DCAM, Dublin Core Abstract Model,

[http://dublincore.org/documents/abstract-model/;](http://dublincore.org/documents/abstract-model/)

DCTerms, Dublin Core Terms,

[http://dublincore.org/documents/dcmi-terms/;](http://dublincore.org/documents/dcmi-terms/)

DGA, Direzione Generale per gli Archivi,

[http://www.archivi.beniculturali.it/;](http://www.archivi.beniculturali.it/)

EAC-CPF, Encoded Archival Context – Corporate bodies, Persons, Families,

[http://eac.staatsbibliothek-berlin.de/;](http://eac.staatsbibliothek-berlin.de/)

EAC-CPF Tag Library,

[http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html;](http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html)

EAC-CPF Vocabulary Specification 1.1,
[http://archivi.ibc.regione.emilia-romagna.it/ontology/
reference_document/referencedocument.html](http://archivi.ibc.regione.emilia-romagna.it/ontology/reference_document/referencedocument.html);

EAD, Encoded Archival Description,
<http://www.loc.gov/ead/eaddev.html>;

EDM, Europeana Data Model,
<http://pro.europeana.eu/>;

FOAF Vocabulary Specification,
<http://xmlns.com/foaf/spec/>;

FRBR, Functional Requirements for Bibliographic Records,
http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf;

GeoNames ontology,
<http://www.geonames.org/ontology/documentation.html>;

International Council on Archives,
<http://www.icacds.org.uk/icacds.htm>;

ICA-AtoM,
<https://www.ica-atom.org/>;

ISAD (G), General International Standard Archival Description,
Second Edition, [http://www.icacds.org.uk/eng/ISAD\(G\).pdf](http://www.icacds.org.uk/eng/ISAD(G).pdf);

ISAAR (CPF), International Standard Archival Authority Record For
Corporate Bodies, Persons and Families, Second Edition,
http://www.icacds.org.uk/eng/isaar2ndedn-e_3_1.pdf;

ISAAR (CPF), Standard internazionale per i record d'autorità
archivistici di enti, persone, famiglie,
<http://www.icar.beniculturali.it/biblio/pdf/standard/isaar%202.pdf>

ISBD, International Standard Bibliographic Description,
http://www.ifla.org/files/assets/cataloguing/isbd/isbd-cons_2007-it.pdf;

ISDF, International Standard for Describing Functions,
<http://www.ica.org/10208/standards/isdf-international-standard-for-describing-functions.html>;

ISDIAH, International Standard for Describing Institutions with
Archival Holdings, <http://www.ica.org/?lid=10198>;

LCC, Library of Congress Classification
<http://www.loc.gov/catdir/cpsol/lcco/>;

LIDO, Lightweight Information Describing Objects,
www.lido-schema.org/schema/v1.0/lido-v1.0-specification.pdf;

Linked Data,
<http://www.w3.org/standards/semanticweb/data>;

METS, Metadata Encoding & Transmission Standard,
<http://www.loc.gov/standards/mets/>;

OAI-ORE, Open Archive Initiative – Object Reuse and Exchange,
<http://www.openarchives.org/ore/1.0/vocabulary>;

Ontology Design Pattern,
<http://ontologydesignpatterns.org/>;

OWL, Web Ontology Language,

<http://www.w3.org/OWL/>;

PRO, Publishing Role Ontology,

<http://purl.org/spar/pro>;

Protégé,

<http://protege.stanford.edu/>

PROV-O, Provenance Ontology,

<http://www.w3.org/TR/prov-o/>;

Relationship in archival descriptive systems,

<http://www.ica.org/13149/standards/cbps-relationship-in-archival-descriptive-systems.html>;

ReLoad, Repository for Linked Open Archival Data,

<http://labs.regesta.com/progettoReload/>;

RDA, Resource Description and Access,

<http://www.rda-jsc.org/rda.html>;

RDF, Resource Description Framework,

<http://www.w3.org/RDF/>;

RDFS, Resource Description Framework Schema,

<http://www.w3.org/2001/sw/wiki/RDFS>;

SAA, Society of American Archivists,

<http://www2.archivists.org/>;

Semantic Web,

<http://www.w3.org/standards/semanticweb/>;

SGML, Standard Generalized Markup Language,

[http://www.w3.org/MarkUp/SGML/;](http://www.w3.org/MarkUp/SGML/)

SKOS, Symple Knowledge Organiation System,

[http://www.w3.org/TR/2009/REC-skos-reference-20090818/;](http://www.w3.org/TR/2009/REC-skos-reference-20090818/)

SNAC, The Social Networks and Archival Context Project,

[http://socialarchive.iath.virginia.edu/;](http://socialarchive.iath.virginia.edu/)

SPAR, Semantic Publishing And Referencing Ontologies,

<http://purl.org/spar;>

SPARQL, Protocol and RDF Query Language,

[http://www.w3.org/TR/rdf-sparql-query/;](http://www.w3.org/TR/rdf-sparql-query/)

TEI, Text Encoding Initiative,

[http://www.tei-c.org/index.xml;](http://www.tei-c.org/index.xml)

URI, Universal Resource Identifiers,

[http://www.w3.org/TR/uri-clarification/;](http://www.w3.org/TR/uri-clarification/)

VIAF, Virtual International Authority File,

[http://viaf.org/;](http://viaf.org/)

x-Dams,

[http://www.xdams.org/;](http://www.xdams.org/)

XML, eXtensible Markup Language,

[http://www.w3.org/XML/;](http://www.w3.org/XML/)

XML Schema,

<http://www.w3.org/standards/xml/schema;>

W3C, World Wide Web Consortium,

<http://www.w3.org/>;

Ultima consultazione degli URL 16/11/2013