



Università degli Studi di Pisa
Corso di laurea in Informatica Umanistica
A.A. 2017/18

Linked Open Data per il Cultural Heritage italiano: conversione e pubblicazione nel Semantic Web del Fondo antico del Comune di Novi Ligure

Candidato: Diletta Goglia

Relatore: Andrea Marchetti

9 maggio 2019

Il Cultural Heritage

- Tanti dati
- Di diversa natura
- Nuovi utenti: le macchine
- Descriverli in maniera formale (*machine readable*)
- Aprirli nel web
- Collegarli fra loro

Stop hugging your data!
Tim Berners Lee, 2009



Le istituzioni culturali hanno avvertito la necessità di rendere aperto l'accesso al proprio patrimonio: ***gallerie, biblioteche, archivi digitali e musei*** oggi si preoccupano di convertire i propri dati secondo le nuove tecnologie superando l'approccio limitante e ormai obsoleto dei "silos di informazione".

Il patrimonio italiano di beni culturali è tra i più rilevanti al mondo.

(Paolini P., Di Blas N., Alonzo F., *ICT per i beni culturali: esempi di applicazione in Mondo Digitale*, n°3, 2005)



Galleries Libraries Archives Museums

Paradigma innovativo per la tutela e la diffusione del patrimonio culturale.

Catalogazione e metadati

Metadati = dati per descrivere altri dati.

BIBO	ICCD	METS
CAT-SAN	ICCU	MIX
CIDOC-CRM	ISAAR (CPF)	MODS
Dublin Core	ISAD (G)	NIERA (EPF)
EAC-CPF	ISBD	PICO
EAD	ISNI	SCONS
EDM	MAG	RDA
FRBR	MARC	VRA



risorsa (bene culturale)

descrizione della risorsa (metadati)

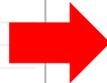


```
<?xml version="1.0" encoding="utf-8"?>
<icar:import xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:icar:import="http://www.
<header>
  <systemId>Memora</systemId>
  <systemTitle>La piattaforma digitale per i beni culturali piemontesi</systemTitle>
  <contact>
    <name>Csi Piemonte - Assistenza Cultura Gruppo</name>
    <url>http://www.csipiemonte.it/web/it</url>
    <mail>assistenza.cultura@csi.it</mail>
    <phone>telefono gruppo assistenza cultura</phone>
  </contact>
  <event eventType="other" eventDate="13 Apr 2019">
    <agent id="export" agentType="software">Memora</agent>
  </event>
  <fileDesc>
    <title>Esportazione dati per tracciato icar</title>
    <abstract>Questo file ml contiene i dati per il tracciato icar</abstract>
    <date>13 Apr 2019</date>
  </fileDesc>
</header>
<listRecords>
  <record>
    <recordHeader action="insert" cascade="true" groupEad="single" type="ead3">
      <id>ICAR-EAD-001</id>
    </recordHeader>
  </record>
  <record action="insert" cascade="true" type="eac-cpf">
    <recordHeader/>
    <id>ICAR-EAC-CPF-001</id>
  </recordHeader/>
  </record>
  <record action="insert" cascade="true" type="scons">
    <recordHeader/>
    <id>ICAR-SCONS-001</id>
  </recordHeader/>
  </record>
</listRecords>
</icar:import>
```

Dal bene (materiale) alle relazioni (immateriali)



SEGNATURA DEFINITIVA	A.S.N.L.3
UNITÀ DI CONSERVAZIONE	Numero: 2.2
IDENTIFICATIVO D'ORIGINE	221
SOGGETTO CONSERVATORE	Novi Ligure. Archivio storico
	Definizione cronologica: 1702 Datazione: 1702 Note alla cronologia: Estremi cronologici della documentazione
CRONOLOGIA E ALTRE DATAZIONI	
DATA TOPICA	Genova
DESCRIZIONE	gennaio 1135 estratto dall'Archivio Segreto di Genova il 20 marzo 1702)
DESCRIZIONE ESTRINSECA	1 fascicolo, lingua Latino, stato di conservazione buono
CONSISTENZA FISICA	Consistenza: 1 Tipologia: fascicolo
LINGUA	LAT Latino
STATO DI CONSERVAZIONE	Stato di conservazione: buono
SITO WEB	
SOGGETTI (ENTI, PERSONE, FAMIGLIE) COLLEGATE	soggetto produttore: Comune di Novi Ligure soggetto conservatore: Comune di Novi Ligure
CONDIZIONI DI ACCESSO E RESTRIZIONI	Pubblico
PROFILO DI ACCESSO	non visibile
IDENTIFICATIVO D'ORIGINE	221
IDENTIFICATIVO DELLA SCHEDA	437195
	Azione: inserimento Data compilazione: 8 March 2006 Compilatore: Dimitri Brunetti - Daniela Cabella - Simona Ballestrero
COMPILAZIONE E CONTROLLO	Azione: modifica Data compilazione: 27 November 2006



Lo sforzo di catalogazione da solo NON E' SUFFICIENTE per l'attività di comunicazione e valorizzazione del patrimonio culturale.

(Paolini P., Di Blas N., Alonzo F., *ICT per i beni culturali: esempi di applicazione in Mondo Digitale*, n°3, 2005)

Museums



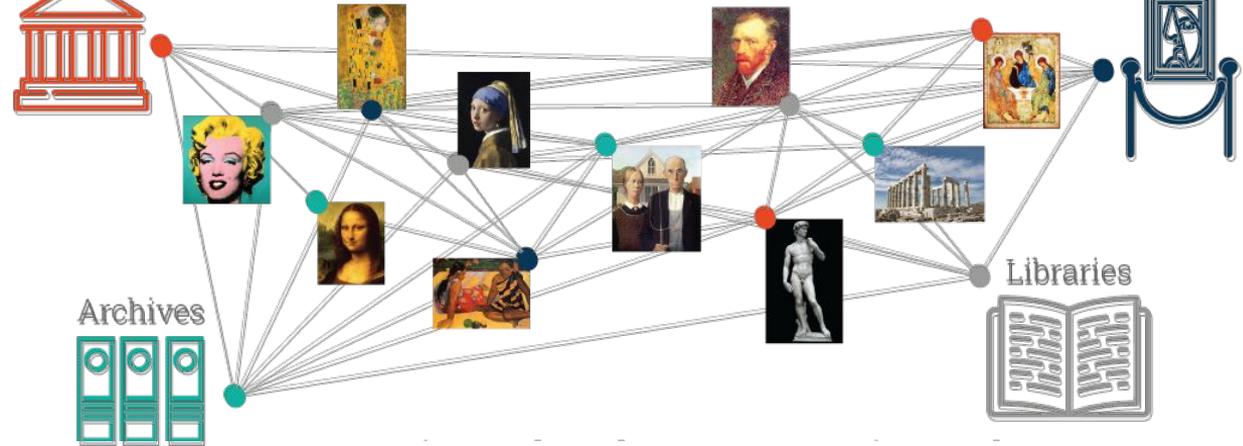
Galleries



Archives

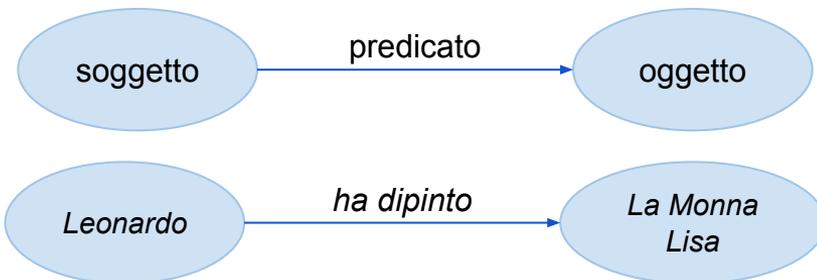


Libraries



Lo standard RDF

RDF è un modello proposto dal Consorzio W3 per l'interoperabilità semantica tra applicazioni che condividono le informazioni sul Web.
Serve a dare una rappresentazione del mondo in modo da renderlo conoscibile e comprensibile alle macchine.



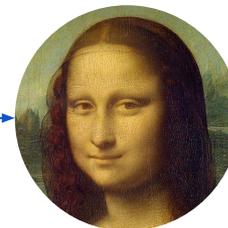
http://dbpedia.org/resource/Leonardo_da_Vinci

<http://wikidata.org/entity/Q12418>



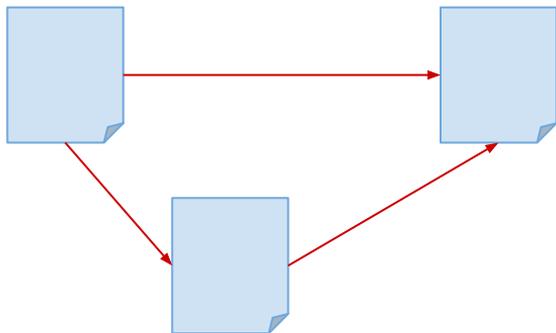
<http://purl.org/dc/terms/creator>

`dcterms:creator`

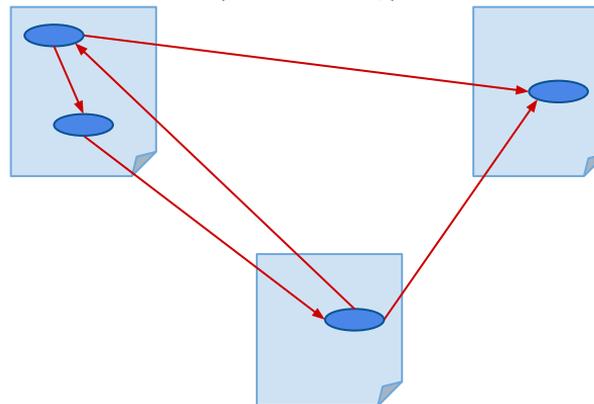


Il Semantic Web

Web of documents
(hyperlinks)



Web of data
(RDF links)



Perché sia giustificato l'aggettivo "semantico", occorre che i dati siano interoperabili non solo in quanto stringhe di bit ma anche in quanto significati.

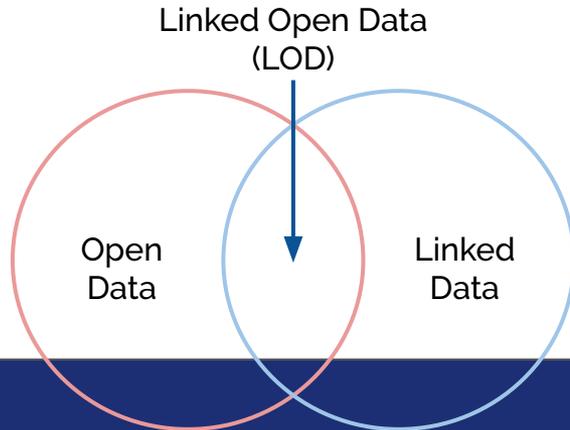
(Gnoli C., Connettere la conoscenza: quattro livelli di interoperabilità, in Biblioteche oggi, vol.32, n°5, 2014, pp. 9-10)

L'idea alla base del web semantico, che lo distingue dal web classico, è l'evoluzione dal concetto di *machine readable* a quello di *machine understandable*.

I Linked Data (o “dati collegati”)

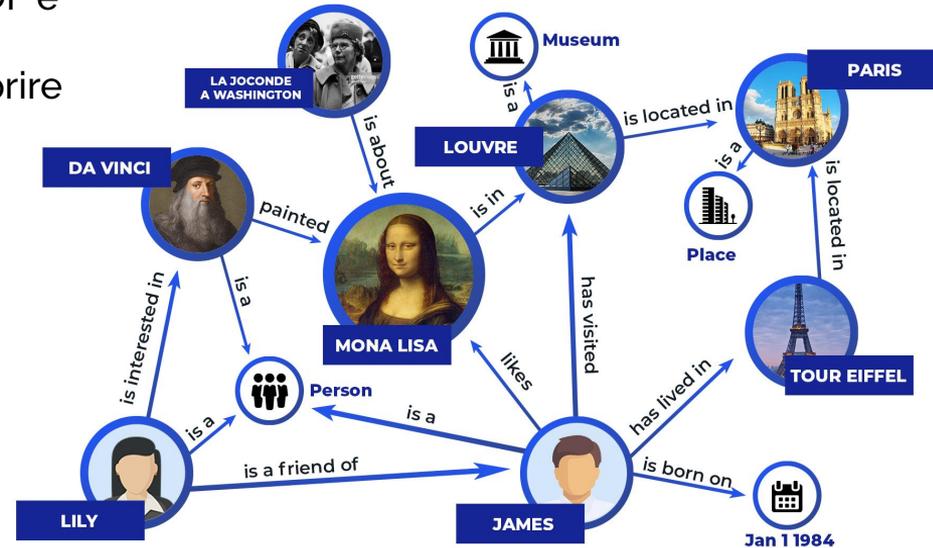
I principi dei Linked Data:

1. Usare URI per dare nomi alle cose
2. Usare URI HTTP in modo che gli utenti possano cercarli
3. Fornire informazioni utili utilizzando standard RDF e SPARQL
4. Includere link ad altri URI in modo da poter scoprire più cose



Linked Data becomes more powerful the more of it there is.

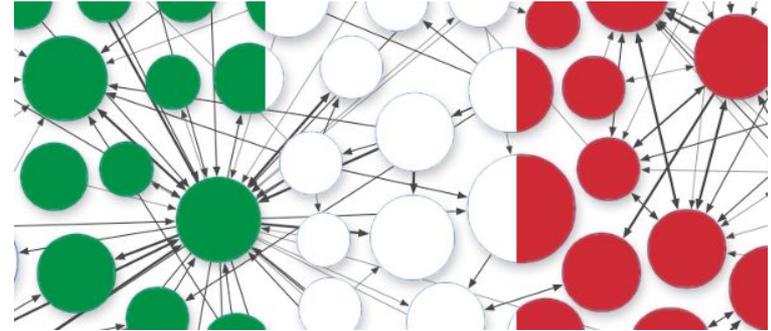
(Byrne G., Goddard L., *The Strongest Link: Libraries and Linked Data*, in *D-LibMagazine*, vol. 16, n°11-12, 2010)



Lo stato dell'arte: Linked Data

[...] la tendenza prevalente sembra concentrarsi più sull'apertura di dati pubblici sotto forma di open data piuttosto che sul ricorso ai LOD.

(Agenzia per l'Italia Digitale (AGID), Linee guida per l'interoperabilità semantica attraverso i Linked Open Data, 2012, pag. 46)



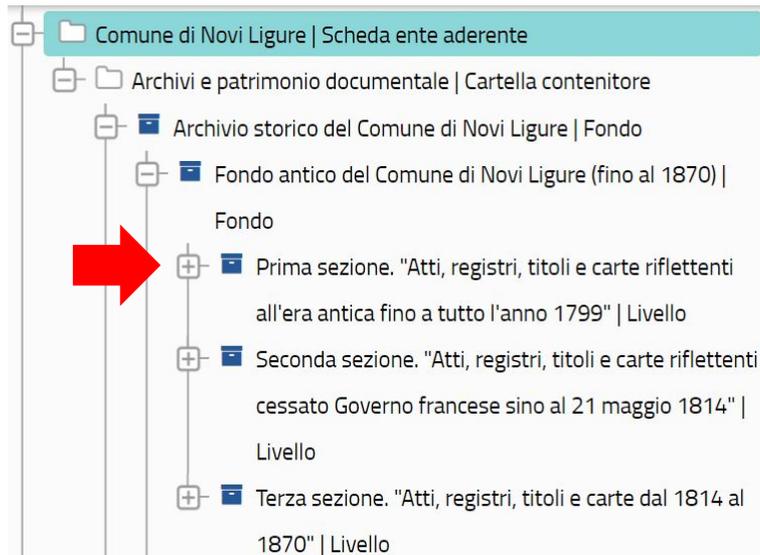
Laura Moro, direttore dell'ICCD, durante la conferenza “*Patrimoni materiali e immateriali: innovazioni e resilienze degli standard italiani di catalogazione*”, tenutasi ad aprile 2019 a Pisa, afferma che *pochissimi* dei dati aperti italiani disponibili sul web sono effettivamente dati collegati.

Sempre Laura Moro, aggiunge che il LOD relativo ai beni culturali in Italia “è un campo ancora da conquistare”, e incentiva le istituzioni culturali e la ricerca all'adozione di questo paradigma, definendolo la *prateria delle nuove generazioni*.

1. Individuazione dei dati

Il Fondo Antico del Comune di Novi Ligure contiene atti, registri, documenti e carte che coprono un periodo che va dagli inizi del 1500 fino al 1870.

Il lavoro di conversione dei dati in LOD è stato eseguito sulla prima sezione del Fondo che copre il periodo 1495 - 1799.



2. Esportazione e pulizia dei dati

- Individuazione delle entità d'interesse da rappresentare
 - Oggetto
 - Consistenza fisica
 - Lingua
 - Persone
 - Date
 - Luoghi



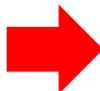
SEGNATURA DEFINITIVA	A.S.N.L.3
UNITÀ DI CONSERVAZIONE	Numero: 2.2
IDENTIFICATIVO D'ORIGINE	221
SOGGETTO CONSERVATORE	Novi Ligure. Archivio storico
CRONOLOGIA E ALTRE DATAZIONI	Definizione cronologica: 1702 Datazione: 1702 Note alla cronologia: Estremi cronologici della documentazione
DATA TOPICA	Genova
DESCRIZIONE	gennaio 1135 estratto dall'Archivio Segreto di Genova il 20 marzo 1702
DESCRIZIONE ESTRINSECA	1 fascicolo, lingua Latino, stato di conservazione buono
CONSISTENZA FISICA	Consistenza: 1 Tipologia: fascicolo
LINGUA	LAT Latino
STATO DI CONSERVAZIONE	Stato di conservazione: buono
SITO WEB	
SOGGETTI (ENTI, PERSONE, FAMIGLIE) COLLEGATE	oggetto produttore: Comune di Novi Ligure oggetto conservatore: Comune di Novi Ligure
CONDIZIONI DI ACCESSO E RESTRIZIONI	Pubblico
PROFILO DI ACCESSO	non visibile
IDENTIFICATIVO D'ORIGINE	221
IDENTIFICATIVO DELLA SCHEDA	437195
COMPILAZIONE E CONTROLLO	Azione: inserimento Data compilazione: 8 March 2006 Compilatore: Dimisi Brunetti - Daniela Cabella - Simona Ballestrero Azione: modifica Data compilazione: 27 November 2006

- Scelta del formato per l'esportazione:
 - MINI-icar-san, che mappa al suo interno altri tre standard per le descrizioni di beni archivistici: EAD, EAC-CPF e SCONS
 - scheda OA
- Esportazione delle schede documentarie in XML

3. Conversione dei dati

XML → CSV

```
<?xml version="1.0" encoding="utf-8"?>
<icar-import xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:icar-import="http://www.
<header>
  <systemId>Memora</systemId>
  <systemTitle>La piattaforma digitale per i beni culturali piemontesi</systemTitle>
<contact>
  <name>Csi Piemonte - Assistenza Cultura Gruppo</name>
  <url>http://www.csipiemonte.it/web/it</url>
  <mail>assistenza.cultura@csi.it</mail>
  <phone>telefono gruppo assistenza cultura</phone>
</contact>
<event eventType="other" eventDate="13 Apr 2019">
  <agent id="export" agentType="software">Memora</agent>
</event>
<fileDesc>
  <title>Esportazione dati per tracciato icar</title>
  <abstract>Questo file xml contiene i dati per il tracciato icar</abstract>
  <date>13 Apr 2019</date>
</fileDesc>
</header>
<listRecords>
  <record>
    <recordHeader action="insert" cascade="true" groupEad="single" type="ead3">
      <id>ICAR-EAD-001</id>
      <lastUpdate/>
      <recordBody/>
    </recordHeader>
  </record>
  <record action="insert" cascade="true" type="eac-cpf">
    <recordHeader/>
    <id>ICAR-EAC-CPF-001</id>
    <lastUpdate/>
    <recordBody/>
  </record>
  <record action="insert" cascade="true" type="scons">
    <recordHeader/>
    <id>ICAR-SCONS-001</id>
    <lastUpdate/>
    <recordBody/>
  </record>
</listRecords>
```



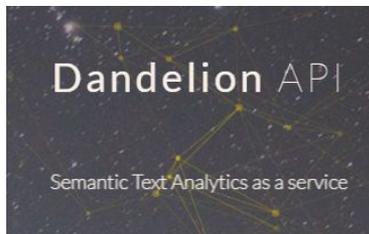
OpenRefine

id_data	tipo	anno	mese	giorno	anno_iniziale	mese_iniziale	id_pers	persona	no_finale
1.	1	giorno	1555	luglio	20		1.	1	Lorenzo Capellono
2.	2	giorno	1559	settembre	30		2.	2	Ercole Spinola
3.	3	giorno	1561	aprile	20		3.	3	duchi di Milano
4.	4	giorno	1569	gennaio	24		4.	4	M. Agost o Rosso
5.	5	giorno	1604	giugno	6		5.	5	Nicolao Bovone
6.	6	giorno	1640	agosto	30		6.	6	Petro Giorgio
7.	7	giorno	1649	maggio			7.	7	Alberto Cavana
8.	8	giorno	1656	maggio			8.	8	Jacomo Gianelli
9.	9	giorno	1756	giugno			9.	9	Ottavio Cattaneo
10.	10	giorno	1756	settembre			10.	10	Clemente Vaccari
11.	11	giorno	1797	marzo			11.	11	Pietro Strozzi
12.	12	giorno	1797	luglio			12.	12	Devesio Capurri
13.	13	giorno	1798	settembre					
14.	14	giorno	1799	febbraio					
15.	15	giorno	1799	febbraio					
16.	16	giorno	1799	aprile					
17.	17	anno	1527						
18.	18	anno	1544						
19.	19	anno	1549						
--	--	--	--	--	--	--			

id_ogg	serie	titolo	consistenza_fisica	lingua
1	1	Statutorum civium insignis oppidi Novarum domini genuensis nu volume		Latino, Italiano
2	1	Statutorum civium insignis oppidi Novarum domini genuensis nu volume		Latino, Italiano
3	2	Convenzione con la città di Genova del 1135	fascicolo	Latino
4	2	Convenzione con la città di Genova del 1157	fascicolo	Latino
5	2	Convenzione con la città di Genova del 1447	fascicolo	Latino
6	2	Molino di Busseto	fascicolo	Latino, Italiano
7	2	Disposizioni genovesi per l'amministrazione cittadina	registro	
8	2	Oratione di M. Lorenzo Capellono al Serenissimo Principe di Spai	quaderno	
9	2	Raccolta di atti antichi	quaderno	Latino, Italiano
10	2	Vertenza della Terra di Novi riguardante le gabelle di macina, carr	registro	
11	2	Cabella della macina e della carne	opuscolo	
12	2	Saggio storico della città di Nove del dottore Ercole Spinola	registro	
13	2	Diritti e privilegi del Comune di Novi, e cambiamenti di dominio. F	fascicolo	
14	2	Convenzione con la città di Genova del 1756	fascicolo	Latino
15	2	Acque	fascicolo	
16	2	Acque. Mulini del Capitolo del Comune	fascicolo	
17	2	Catastro. Memoria senza data	fascicolo	
18	2	Storia dei molini per lo Scrivia dal 1356 al 1779	fascicolo	
19	3	Libro dei decreti	registro	
20	3	Libro de decreti. Trascrizione di atti dal 1388 al 1796	registro	Latino, Italiano
21	3	Sentenze decreti dei duchi di Milano e dei supremi magistrati. Tras	registro	Latino, Italiano
22	4	Registro degli atti consolari dall'8 dicembre 1495 a tutto il 1518. V	registro	Latino
23	4	Registro degli atti consolari dal 24 ottobre 1524 al 1530. Volume I	registro	Latino

4. Estrazione informazioni

- Tool online di *Named Entity Recognition*
- Riconoscimento ed estrazione delle entità a partire dai titoli dei documenti dell'archivio.
- Aggiunta alle tabelle CSV.



Carteggio dei rappresentanti la Città e il Comune di Novi presso la Amministrazione centrale in Genova

0 persons 0 works 0 organisations 2 places 0 events 3 concepts

CONCEPT Epistolario

CONCEPT Città

CONCEPT Comune medievale

PLACE Novi Ligure

PLACE Genova

Sommario della rotta di Pietro Strozzi

1 person 0 works 0 organisations 0 places 0 events 0 concepts

PERSON Pietro Strozzi

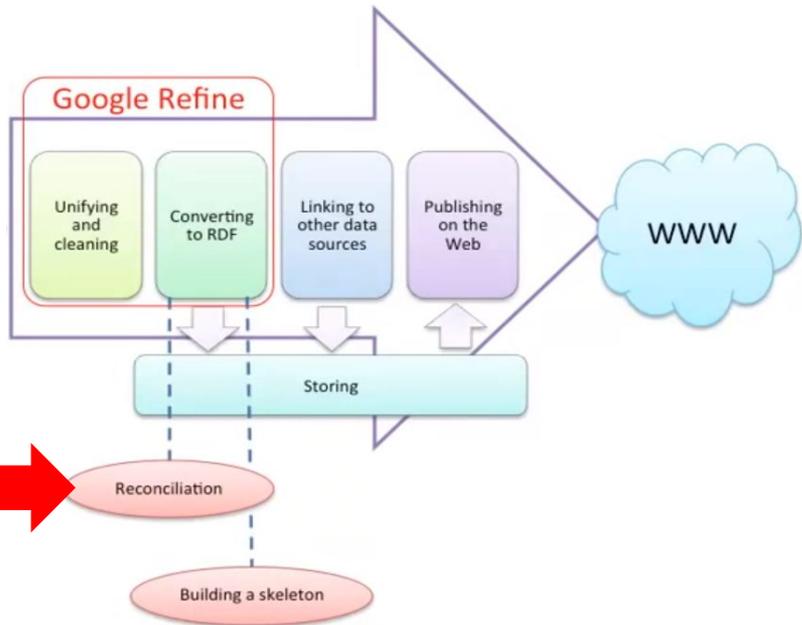
The image displays a user interface for a Named Entity Recognition tool. It shows a search result for a document titled "Carteggio dei rappresentanti la Città e il Comune di Novi presso la Amministrazione centrale in Genova". Below the title, there is a horizontal bar with colored segments representing the count of entities: 0 persons (red), 0 works (purple), 0 organisations (orange), 2 places (blue), 0 events (brown), and 3 concepts (green). Below this bar, there are four entity cards. The first is a green card labeled "CONCEPT" with a magnifying glass icon and the title "Epistolario". The second is a purple card labeled "CONCEPT" with a cityscape image and the title "Città". The third is a blue card labeled "CONCEPT" with a historical painting and the title "Comune medievale". The fourth is a blue card labeled "PLACE" with a night cityscape image and the title "Novi Ligure". Below these, there is a red card labeled "PLACE" with a cityscape image and the title "Genova". To the right, there is a section titled "Sommario della rotta di Pietro Strozzi" with a horizontal bar showing 1 person (red), 0 works (purple), 0 organisations (orange), 0 places (blue), 0 events (brown), and 0 concepts (green). Below this bar is a red card labeled "PERSON" with a portrait of a man and the title "Pietro Strozzi".

More Tags More Precision

A horizontal slider control with a white handle in the center. The text "More Tags" is on the left and "More Precision" is on the right.

5. Link a dataset esterni

- Scelta dei dataset esterni
- *Reconciliation services* (Refine)



```
name: "GND reconciliation for OpenRefine",
identifierSpace: "https://lobid.org/gnd",
schemaSpace: "https://lobid.org/gnd",
view: {
  url: "https://lobid.org/gnd/{id}"
},
preview: {
  height: 100,
  width: 320,
  url: "https://lobid.org/gnd/{id}.preview"
},
extend: {
  property_settings: [{
    name: "limit",
    label: "Limit",
    type: "number",
    default: 0,
    help_text: "Maximum number of values to return per row (0 for no limit)"
  }, {
    name: "content",
    label: "Content",
    type: "select",
    default: "literal",
    help_text: "Content type: ID or literal",
    choices: [{
      value: "id",
      name: "ID"
    }, {
      value: "literal",
      name: "Literal"
    }
  ]
}]
}
```



6. Disambiguazione

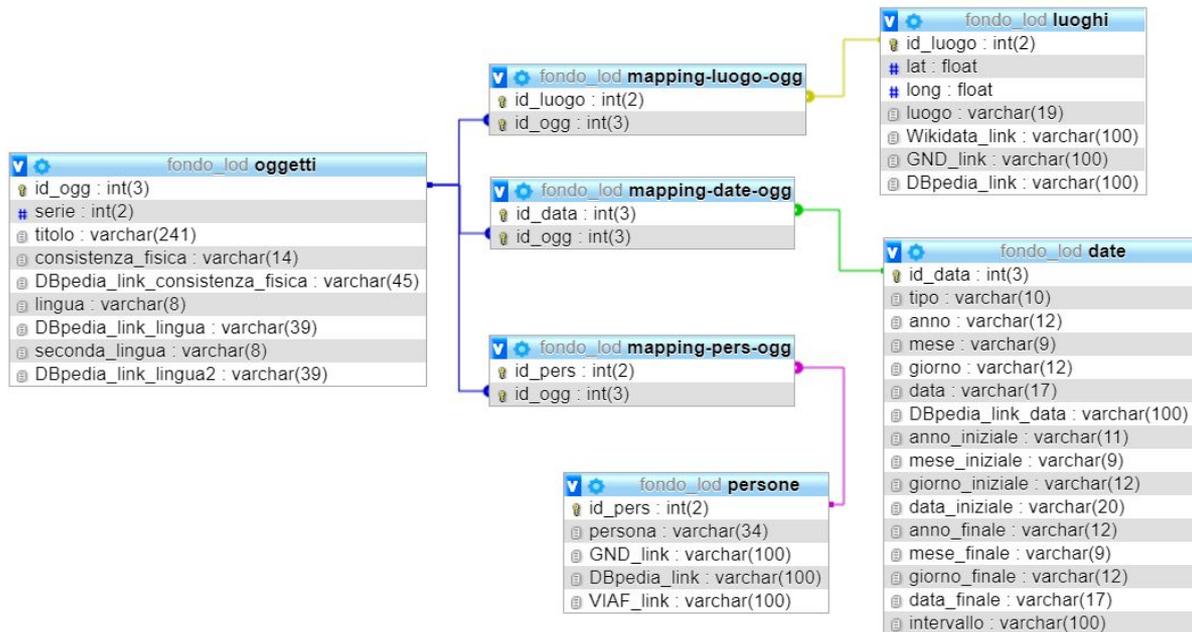
Il mapping ambiguo (riferito a collegamenti con più termini) viene separato da quello non ambiguo e viene corretto.

Il sistema viene “aiutato” a *matchare* correttamente le entità con le rispettive risorse recuperate da fonti esterne.

luogo	Wikidata_link	GND_link	DBpedia_link
Genova <input checked="" type="checkbox"/> Genova (100) <input checked="" type="checkbox"/> Genova C.F.C. (100) <input checked="" type="checkbox"/> Province of Genoa (100) <input checked="" type="checkbox"/> Create new topic Search for match	https://www.wikidata.org/wiki/Q1449	https://lobid.org/gnd/5099507-8	http://it.dbpedia.org/resource/Genova
Milano <input checked="" type="checkbox"/> Milan (100) <input checked="" type="checkbox"/> province of Milan (100) <input checked="" type="checkbox"/> Alyssa Milano (100) <input checked="" type="checkbox"/> Create new topic Search for match	https://www.wikidata.org/wiki/Q490		http://it.dbpedia.org/resource/Milano
Gavi <input checked="" type="checkbox"/> Gavi, Piedmont (100) <input checked="" type="checkbox"/> Gavi, the Vaccine Alliance (100) <input checked="" type="checkbox"/> Cortese di Gavi (100) <input checked="" type="checkbox"/> Create new topic Search for match	https://www.wikidata.org/wiki/Q17368	https://lobid.org/gnd/1600218-0	http://it.dbpedia.org/resource/Gavi
Busseto <input checked="" type="checkbox"/> Busseto (100) <input checked="" type="checkbox"/> Busseto (100) <input checked="" type="checkbox"/> Busseto (100) <input checked="" type="checkbox"/> Create new topic Search for match	https://www.wikidata.org/wiki/Q100772	https://lobid.org/gnd/4528029-0	http://it.dbpedia.org/resource/Busseto
Spagna <input checked="" type="checkbox"/> Spain (100) <input checked="" type="checkbox"/> Spagna (100) <input checked="" type="checkbox"/> Spagna metro station (100) <input checked="" type="checkbox"/> Create new topic Search for match	https://www.wikidata.org/wiki/Q29		http://it.dbpedia.org/resource/Spagna

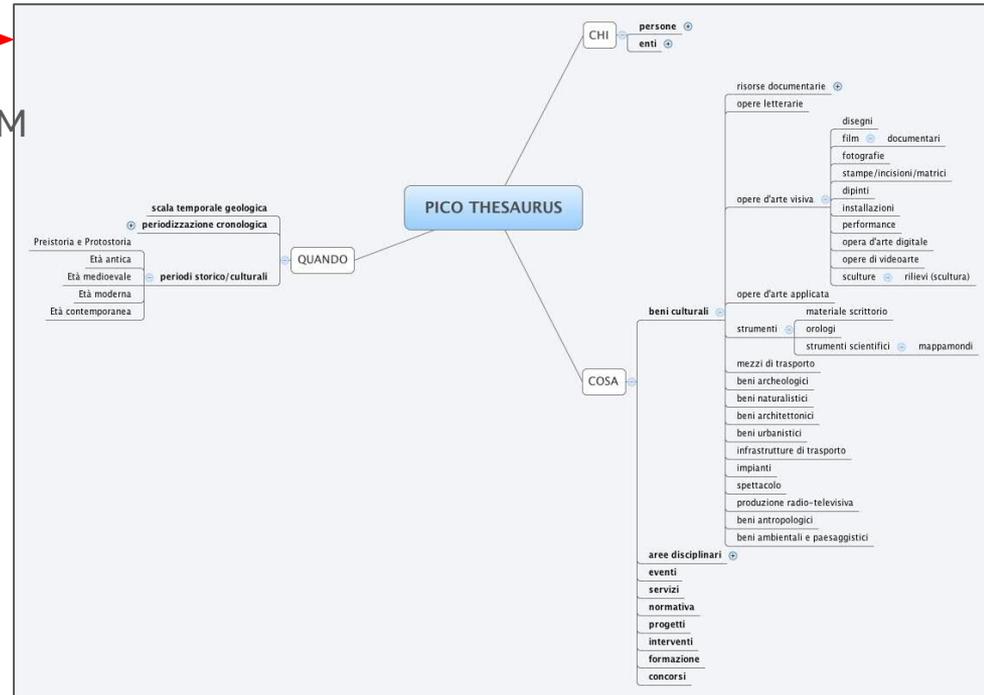
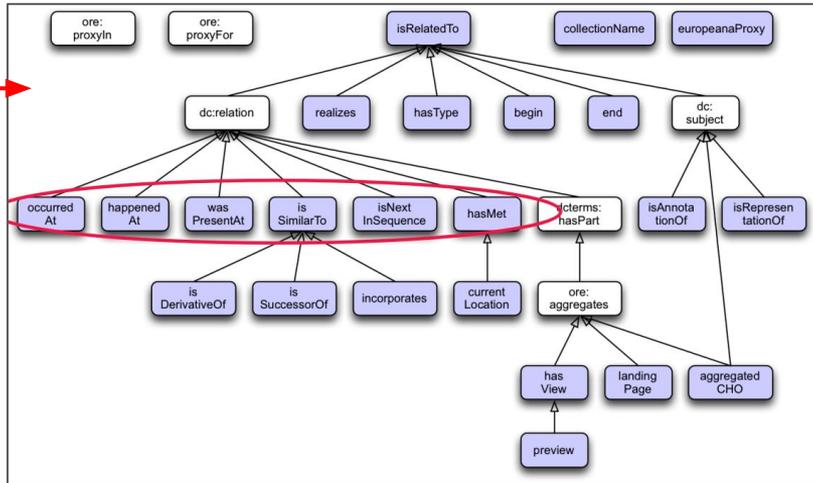
7. Database relazionale

- Import delle tabelle in un database MySQL
- Definizione delle relazioni M-M tra tabelle (*foreign key*)
- Tabelle di mapping



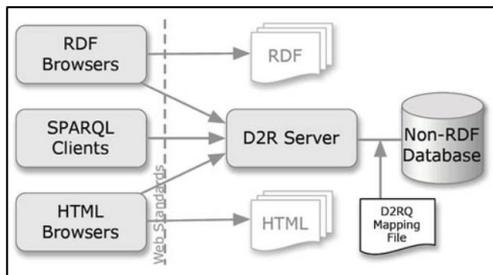
8. Scelta dell'ontologia

- PICO Thesaurus
- Europeana Data Model
- Basic Geo Vocabulary, raccomandato dall'EDM
- Vocabolari condivisi e ontologie del web semantico (SKOS, OWL, DC, RDFS, FOAF, ...)



9. Modellazione ontologica

- Primo mapping "di default" tramite D2RQ
generate-mapping -o mapping.ttl -u root
jdbc:mysql://localhost/fondo_lod



- Personalizzazione del mapping in linguaggio D2RQ un **linguaggio dichiarativo per la mappatura** di schemi di database relazionali in vocabolari RDF e ontologie OWL



```
# Mapping-luogo-ogg (n:m)
map:mapping-luogo-ogg__link a d2rq:PropertyBridge;
    d2rq:belongsToClassMap map:oggetti;
    d2rq:property edm:happenedAt;
    d2rq:refersToClassMap map:luoghi;
    d2rq:join "mapping-luogo-ogg.id_luogo => luoghi.id_luogo";
    d2rq:join "mapping-luogo-ogg.id_ogg => oggetti.id_ogg";
.

# Mapping-pers-ogg (n:m)
map:mapping-pers-ogg__link a d2rq:PropertyBridge;
    d2rq:belongsToClassMap map:oggetti;
    d2rq:property skos:related;
    d2rq:refersToClassMap map:persone;
    d2rq:join "mapping-pers-ogg.id_ogg => oggetti.id_ogg";
    d2rq:join "mapping-pers-ogg.id_pers => persone.id_pers";
.

# Oggetti
map:oggetti a d2rq:ClassMap;
    d2rq:dataStorage map:database;
    d2rq:uriPattern "oggetti/@oggetti.id_ogg@";
    d2rq:class pico:beni_archivistici;
    d2rq:classDefinitionLabel "oggetti";
.

map:oggetti__label a d2rq:PropertyBridge;
    d2rq:belongsToClassMap map:oggetti;
    d2rq:property skos:prefLabel;
    d2rq:pattern "oggetti #@@oggetti.id_ogg@";
.

map:oggetti_id_ogg a d2rq:PropertyBridge;
    d2rq:belongsToClassMap map:oggetti;
    d2rq:property dc:identifier;
    d2rq:propertyDefinitionLabel "oggetti id_ogg";
    d2rq:column "oggetti.id_ogg";
    d2rq:datatype xsd:integer;
```

10. Creazione degli URI

Gli URI fanno riferimento al nome associato all'indirizzo della macchina locale dove è stato implementato il progetto, ovvero `http://localhost:2020/`.

Tutti gli URI che identificano le risorse (entità) sono basati sulle chiavi primarie del database per renderli **dereferenziabili**.

Si ottengono tramite l'utilizzo di un prefisso del tipo `http://localhost:2020/resource/` al quale segue l'indicazione specifica della tipologia di risorsa (entità) e l'identificatore, ottenendo quindi un suffisso del tipo `entità/id`.

Esempi:

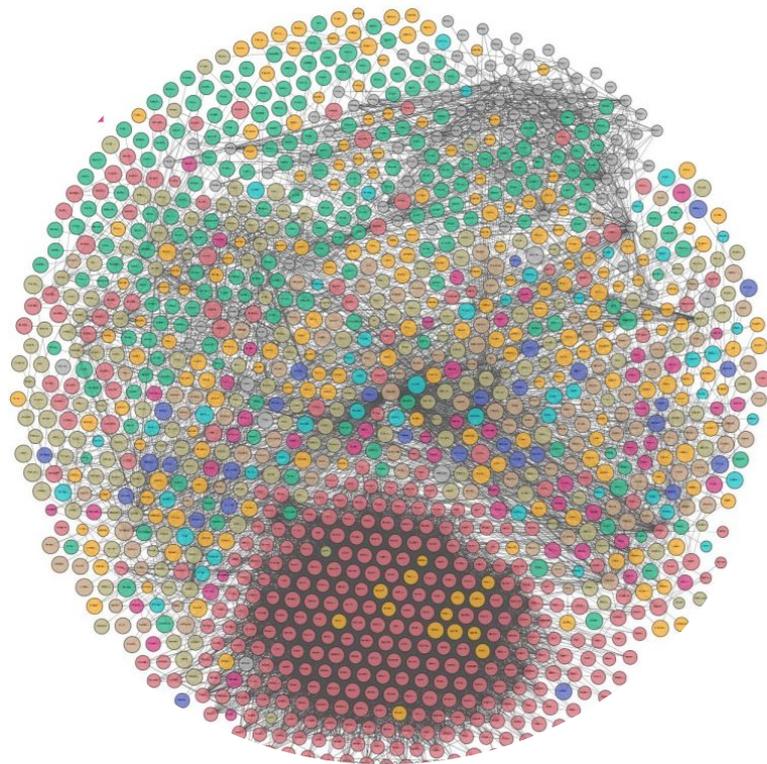
<http://localhost:2020/resource/persona/13>

<http://localhost:2020/resource/date/3>

11. Esportazione e accesso

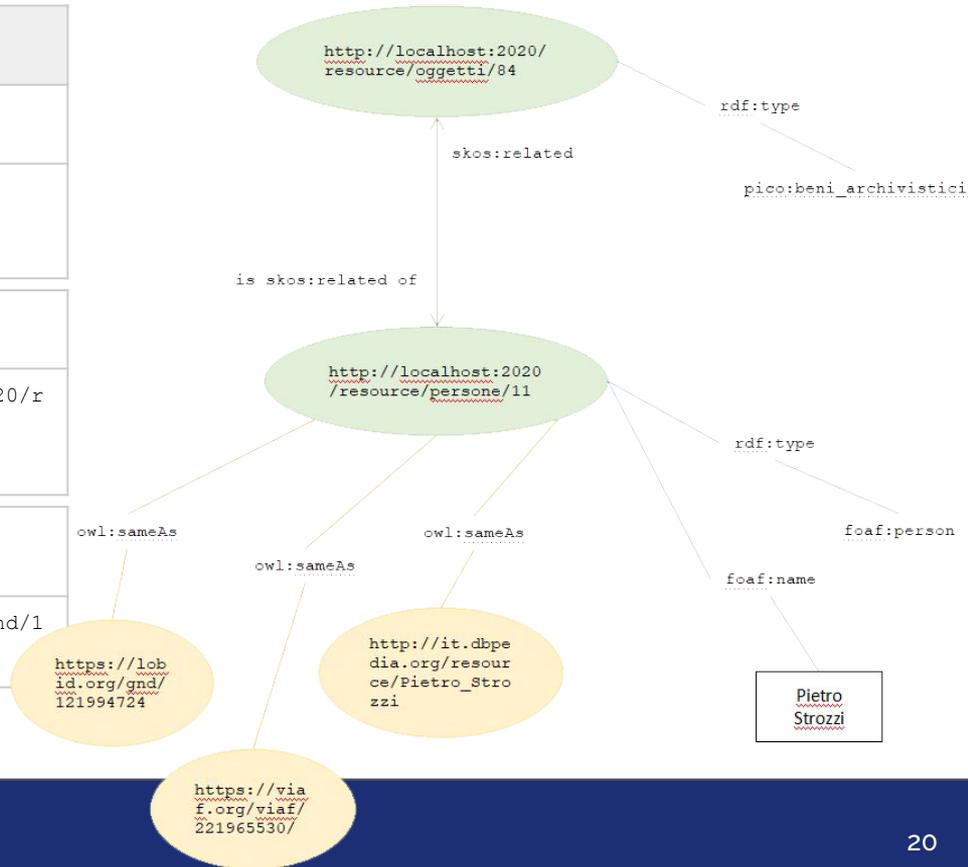
- Dump del database tramite D2RQ in RDF/XML, N-TRIPLE e Turtle
- Costruzione di un **sito web** con cui è possibile navigare all'interno dell'archivio ed effettuare query tramite uno SPARQL Endpoint.
- L'intero progetto è Open Source ed è disponibile su un repository GitHub al seguente indirizzo:

https://github.com/alod83/fondo_novi



Esempi di triple RDF

Subject	Predicate	Object
La persona con id 11	ha un nome	Pietro Strozzi
http://localhost:2020/resource/persona/11	http://xmlns.com/foaf/0.1/name	"Pietro Strozzi"
La persona con id 11	è correlato a	l'oggetto con id 84
http://localhost:2020/resource/persona/11	http://www.w3.org/2004/02/skos/core#related	http://localhost:2020/resource/oggetti/84
Questo URI	si riferisce alla stessa cosa	di questo URI
http://localhost:2020/resource/persona/11	http://www.w3.org/2002/07/owl#sameAs	https://lobid.org/gnd/121994724



Risultato: vantaggi ottenuti

- **Modello** di dati **unificante** (RDF garantisce l'identificazione univoca a livello globale delle entità)
- Meccanismo di **accesso** ai dati **standardizzato** tramite il protocollo HTTP
- **Scoperta di nuova conoscenza** percorrendo i collegamenti tra entità provenienti da diversi dataset. Questo è il modo in cui il web of data può essere navigato o scansionato da un crawler.
- Collegando tra di loro risorse provenienti da diversi dataset, si collegano anche i **dataset** stessi, così vengono resi **interoperabili**. Questo approccio supera e rende ormai deprecati i "silos" di informazioni chiuse, fornendo **nuovi modi di accesso, convalida e utilizzo di dati** provenienti da fonti diverse.
- **Varietà di espressione** che deriva dalla possibilità di utilizzare termini provenienti da vocabolari condivisi e collegati.
- **Rimosse** le **barriere intellettuali** all'apertura dei dati e al loro utilizzo
- **Rimossi** gli **ostacoli tecnologici** alla libera condivisione dei dati nel web
- **Arricchimento e integrazione** dei dati con quelli di altre istituzioni: **aumento del potenziale informativo**, più completi e riusabili in contesti anche diversi da quelli d'origine.
- **Autorevolezza**: l'incontro con altri dati nel web accresce il numero degli strumenti di controllo terminologico a disposizione, rendendo più precise e pertinenti le fonti informative, conferendo legittimità e validità ai dati;
- Maggiore impatto e nuovo pubblico far uscire fuori dal web dati locali, renderli aperti e universalmente accessibili, significa offrire **opportunità di visibilità**. Le istituzioni possono migliorare la propria presenza sul web.
- Miglioramento dell'**esperienza per gli utenti**: facilitando la scoperta di nuovi contenuti
- **Riutilizzo efficiente** delle risorse: **costruzione di servizi** innovativi che sfruttano queste tecnologie.

Valutazione qualitativa e conclusioni

Aderenza ai principi Linked Data:

- ✔ Usare URI per dare nomi alle cose.
- ✔ Usare URI HTTP in modo che gli utenti possano cercarli.
- ✔ Fornire informazioni utili utilizzando standard RDF e SPARQL.
- ✔ Includere link ad altri URI in modo da poter scoprire più cose.

Aderenza alle regole fissate dal Linking Open Data Project:

- ✔ Il dataset deve contenere almeno mille triple (ne contiene 16.889).
- ✔ Il dataset deve essere connesso con link RDF ad altri dataset già presenti nel Linking Open Data Project.
- ✔ Sono richiesti arbitrariamente almeno 50 collegamenti.
- ✔ L'accesso all'intero dataset deve essere possibile tramite crawling RDF, dump RDF o tramite un endpoint SPARQL.

Grazie a tutti dell'attenzione!



UNIVERSITÀ DI PISA