



**Semantizing and visualising archival metadata:
the French PIAAF prototype
and the following projects
at the Archives nationales de France (ANF)**

Florence Clavaud
Archives nationales de France
florence.clavaud@culture.gouv.fr

Summary

- PIAAF in a few words
- The context and stakes
- Project management
- The ‘making of’
 - preparing the metadata
 - converting them into RDF datasets
 - building the web application
- Results
- What’s next at the ANF? The following projects

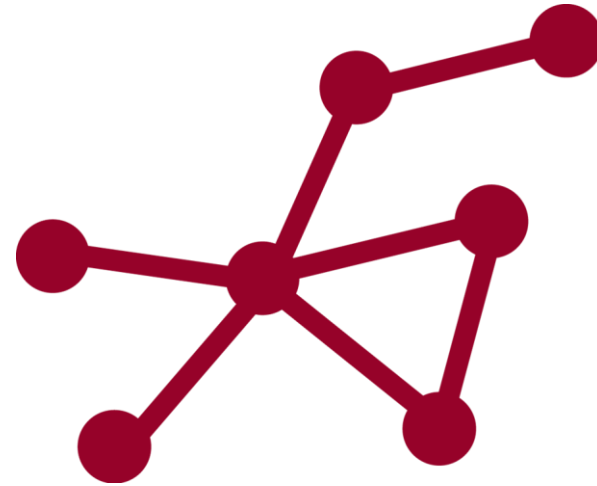


1. PIAAF in a few words

PIAAF: what does it mean?

- « Pilote d'Interopérabilité pour les Autorités Archivistiques Françaises »
= Pilot for Interoperable Archival Authorities in France
- Actually, it is not only about archival authority records, it is also about finding aids
 - But the archival authority records (and vocabularies) are key metadata components for interconnecting datasets created by different institutions or projects

A proof of concept



Deliverables


- RDF datasets & **conversion tools**
- A web application (demonstration) and data visualisation libraries
- **An assessment report** (about the results, the methodology and the prospects)

Project supported by the French Ministry of Culture

Participants



**ARCHIVES
NATIONALES**

{ BnF |  CRIHN

- Authority records
- Finding aids



Logilab



**Demo web
app.**



Within the Ministry of Culture: the Service interministériel des Archives de France (SIAF)

- Controls the network of French public archival institutions
 - > **need of a demonstrating/pedagogic tool**
- Builds **a set of records describing categories of decentralised administrative units or regional authorities as historical entities** (for instance, a district public archival institution, of a regional direction of cultural affairs...)
 - > **need to represent this graph of entities, and to connect it with the graph of central administrations created by the Archives nationales de France**
- **Responsible for the new French archives portal (France Archives, <https://francearchives.fr/>)**
 - > **need to think of the future of this portal** (using semantic technologies)

At the Archives nationales de France (ANF)

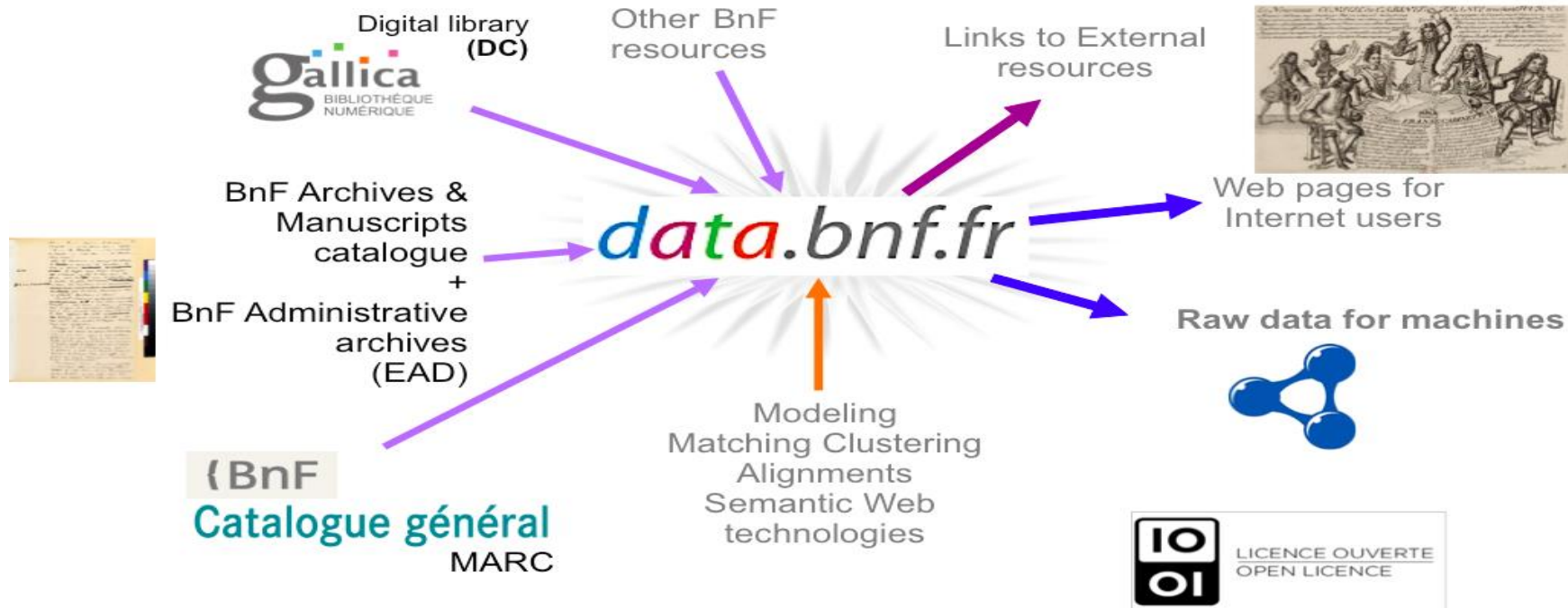
- 25 000 archival finding aids (XML/EAD 2002 structured documents)
- 15 000 authority records (in XML/EAC-CPF)
- 20 000 relations between the authority records
- 30 000 relations between the finding aids and the authority records
- about 20 vocabularies used for indexing these files

> need to reveal, process and query this multidimensional oriented graph

> need to think of opening the information silo, and to connect these data to other data sets (Linked open data)

> need to think of the future of the whole information system

At the Bibliothèque nationale de France (BnF)



- > Need a method and tools for representing archival metadata and authority records in RDF
- > Need to publish and visualise these metadata
- > Need to connect these data with those of archival institutions, starting with the ANF

In short: different situations and needs, but a strategic project for the three partners

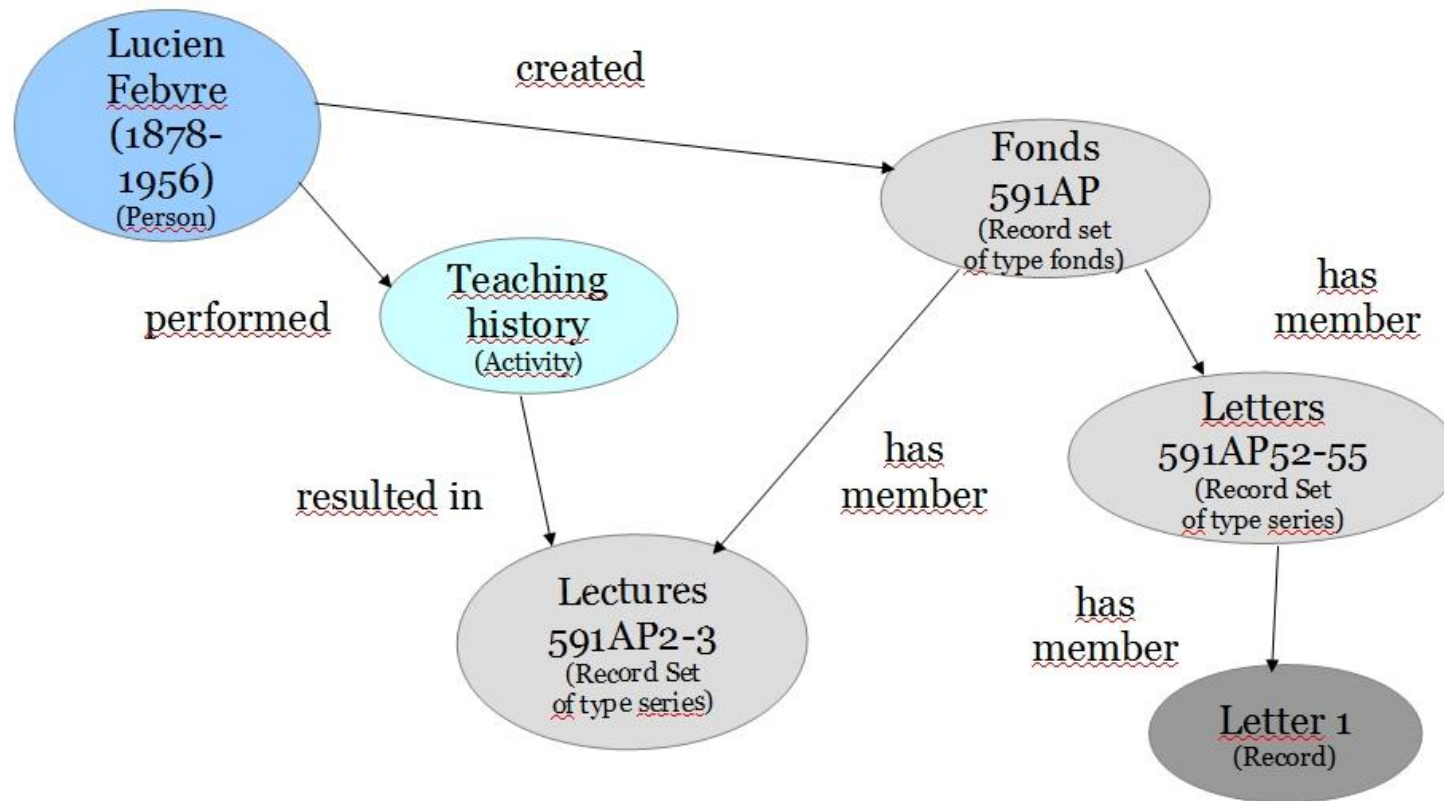
- Prove it is feasible and useful to:
Convert (to RDF, accurately) – Interconnect – Make accessible and visualize, **real** archival metadata sets
- A small, experimental, but quite ambitious and challenging project
- **Qualitative requirements mainly** (it was not about the whole, huge, amount of metadata we already had)
- Strong interest for the **methodology and outcomes**
- Other institutions might also be interested

The « technical » context

At last, an OWL ontology for creating RDF datasets describing archives

- RiC-O (**Records in Contexts-Ontology**) is being developed by the Expert Group on Archival Description of the International Council on Archives (ICA-EGAD)
- It is the **second part of a new standard, Records in Contexts, that will replace the current four descriptive metadata standards** (ISAD(G), ISAAR(CPF), ISDF and ISDIAH)
- RiC-O is a formal, RDFS/OWL, implementation of RiC-Conceptual Model (the first part of RiC standard). A version 0.1 of RiC-CM was published in September 2016
<http://www.ica.org/en/egad-ric-conceptual-model>
- **RiC-O v0.1 will be published by the end of November 2019**
- Aims:
 - replace the four existing standards, (re)define the concepts;
 - focus on the entities that exist in the world of archives rather than on description rules;
 - take into account the evolution of archival theory and practice and the needs of users;
 - consider archives and their multiple layers of contexts;
 - enable archivists and records managers to move forward to a more accurate, more nuanced, easier to process, multidimensional description

The archival world seen as an oriented graph: a small example



The same assertions as in the previous diagram, expressed in RDF/Turtle syntax

```
anf:LucienFebvre      RiC-O:created      anf:fonds591AP ;
                    rdf:type      RiC-O:Person ;
                    RiC-O:performed  anf:teachingHistory ;

                    owl:sameAs
<http://data.bnf.fr/ark:/12148/cb119024609#foaf:Person> .

anf:fonds591AP      rdf:type      RiC-O:RecordSet ;
                    RiC-O:hasMember  anf:letters591AP52To55,
                                        anf:Lectures591AP2to3 .

anf:letters591AP52To55  rdf:type      RiC-O:RecordSet ;
                    RiC-O:hasMember  anf:letter1 .

anf:letter1          rdf:type      RiC-O:Record.

anf:teachingHistory  rdf:type      RiC-O:Activity ;
                    RiC-O:resultedIn  anf:Lectures591AP2to3 .

anf:Lectures591AP2to3  rdf:type      RiC-O:RecordSet.
```

Using RiC-O for the first time at this early stage: a risk we accepted to take

- A strong link between the PIAAF team and the RiC-O development team within EGAD
- PIAAF only uses a subset, that was already quite consistent in 2017, of RiC-O (and also extends it a bit)
- PIAAF, a use case / test-bed for RiC-O, that has been very helpful for EGAD (see RiC-O design principles: to be useful, and usable)

Visualising RDF datasets conforming to RiC-O on the web: a true challenge

- RiC-O is a domain model, thus it defines a significant number of object categories (classes), a lot of relations, which makes dense graphs
- Articulating complex data and the temporal dimension
- Displaying information at a very fine level of granularity, and making something that a human user can understand

Kindred Britain

KINDRED BRITAIN

PEOPLE

CONNECTIONS

STORIES

🔍 🏠 ?

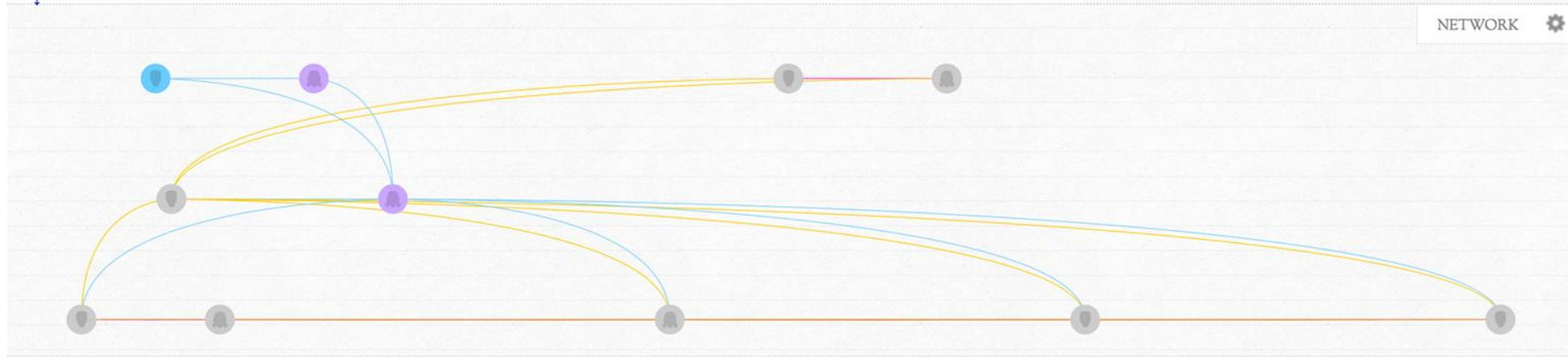
President George Washington

1732 - 1799

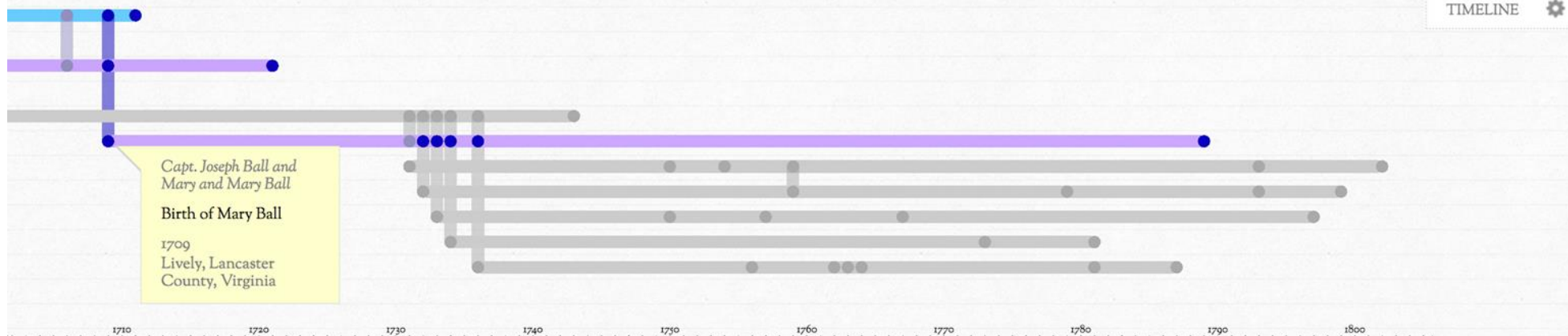


The family of President George Washington

NETWORK ⚙️



TIMELINE ⚙️





3. Project management

How we got organized

- The three institutional partners signed, in May 2016, a specific agreement
- Each partner brought:
 - 15000 euros
 - archival metadata sets
 - and human resources (time!)

How we got organized (2)

- A small project team
- Project lead by the ANF
- The ANF converted the three metadata sets to RDF
- A private company, [Logilab](#), realized the web application, within the framework of a public services contract

Specifications of the web application

- A **specific website**, distinct from the information systems of the 3 institutions
- Two categories of end users: the archival community (and other cultural heritage communities) and the public, from citizens to researchers
- An RDF triplestore and SPARQL endpoint (for enabling textual queries), allowing advanced search
- Dynamic and interactive graphical visualisation, including a timeline
- Ability to align the datasets (inferencing)
- Analysis of the input data (provided RDF datasets)
- Open source components only ; development based on small iterations and interactions with the project team (agile approach)

Main steps

- 2015 – organization
- 2016 – preparation of the source metadata sets, call to tender
- 2017 – RDF conversion, development of the web application (4 iterations), test workshops
- **2018, March 1st – the prototype is online**
(<http://piaaf.demo.logilab.fr>)



Step 1: source metadata sets curation, harmonization and enrichment

The source datasets chosen for the project: nature and scope

- **A small selection of ‘real world’ authority records describing archival creators, plus some finding aids** from each of the partners
- **Relevant, controllable** (amount, characteristics...), possible to interconnect with the data provided by the others
- Selection from two functional domains :
 - management of historical monuments and “civil buildings”
 - management of public reading and public librarieswithin the French governments from the 19th century to nowadays

Common encoding rules

Each of the partners preserves its own perspective and uses its own tools

However, we agreed on applying common encoding rules:

- metadata encoded in XML/EAD 2002 and XML/EAC-CPF;
- some mandatory information (e.g., in the EAC-CPF files, existence dates, legal status of the entity...);
- forms of names standardized according to French rules (EAC-CPF <nameEntry>);
- ISNI identifiers added in EAC-CPF <entityId>;
- expressiveness of **@cpfRelationType** enhanced beyond “hierarchical” and “associative” with a variety of @xlink:arcrole values : **isControlledBy, isDirectedBy/isDirectorOf, hasPart/isPartOf, isFunctionallyLinkedTo, etc.**
- enrichment with “**identity**” links to descriptions of the same entities in other datasets (Wikipedia, data.bnf.fr)

Building and using two small vocabularies for indexing the activities of the entities described

- Indexing the activities of the entities: we considered an activity type as the combination of a domain/object and a process/action
- In order to have shared controlled values for this, we built two small RDF/SKOS thesauri using [Ginco tool](#) (an open source thesaurus editor developed for the French Ministry of Culture)

```
<function localType="piaafFunction">  
  <term vocabularySource="http://data.culture.fr/thesaurus/resource/ark:/67717/f14e8183-5885-46d6-8fc9-17ebd8f3c27e" piaaf:rank="1"  
    piaaf:conceptId="http://data.culture.fr/thesaurus/resource/ark:/67717/4002e166-eb11-446c-a777-cb73b6def641">objets culturels</term>  
  <term vocabularySource="http://data.culture.fr/thesaurus/resource/ark:/67717/51232822-adac-4a33-aa14-29e2c701a5ee" piaaf:rank="2"  
    piaaf:conceptId="http://data.culture.fr/thesaurus/resource/ark:/67717/b06a81ff-bbe0-4bf3-b33e-e51d26eb0431">conservation</term>  
</function>
```

Results: three sets of XML structured documents

- From the BnF :
 - 55 records describing corporate bodies
 - 12 records describing persons
 - 6 finding aids
- From the SIAF:
 - 14 records (describing in an abstract way some decentralised administrative units or regional authorities)
 - 93 records on decentralised administrative units or regional authorities
 - 2 records describing persons
 - 2 finding aids (authored by the Archives départementales de la Gironde)
- From the ANF:
 - 90 records describing central government entities and their subdivisions, plus some national institutions
 - 10 records describing persons
 - 30 finding aids
- **Thus: 276 EAC/CPF authority records and 38 EAD 2002 finding aids**

**Step 2: conversion of these
three sets to RDF datasets
using RiC-0
as the reference ontology**

RiC-O classes that were used

RiC:Agent,
RiC:CorporateBody,
RiC:Person
RiC:GroupType
RiC:FunctionAbstract
RiC:Position
RiC:Mandate
RiC:Event
RiC:AgentName
RiC:Place
RiC:Relation
RiC:Description
RiC:Record
RiC:RecordSet
RiC:Proxy
RiC:FindingAid

Conversion to RDF: aims

The process should not result in loss of information

A lot of objects are described/represented/referred to (sometimes in an implicit way) in the source metadata (much more than the archival creators and records that are their main subjects):

- **in the authority records** : names, events, positions, activities, places, relations between the archival creators ;
- **in the finding aids** : through the XML node tree, aggregation and sequential relations; provenance relations; entities that are the subjects of the records described; digital instantiations of the primary records...

From one authority record encoded in EAC-CPF, and describing one person...

...one can generate and interconnect RDF resources of many different categories

```
-<eac-cpf xsi:schemaLocation="urn:isbn:1-931666-33-4 schema/cpf-modified.xsd">
+<control></control>
-<cpfDescription>
  -<identity>
    <entityId>ISNI 0000 0003 8432 0062</entityId>
    <entityType>person</entityType>
    <nameEntry localType="preferredFormForProject" scriptCode="Latn" xml:lang="fre" xml:id="FRAN_an_050789-d249e170">
      <part>Jeanneney, Jean-Noël (1942-...)</part>
    </nameEntry>
  </identity>
  -<description>
    -<existDates>
      -<dateRange>
        <fromDate standardDate="1942-04-02">2 avril 1942</fromDate>
        </dateRange>
      </existDates>
    +<occupations></occupations>
    -<biogHist>
      -<p>
        Jean-Noël Jeanneney est né en avril 1942 à Grenoble. Après des études aux lycées Montaigne et Louis-le-Grand à Paris, il est reçu à l'École normale supérieure de la rue d'Ulm où il est élève de 1961 à 1966. Diplômé contemporaine à l'université de Nanterre comme assistant puis maître-assistant. Docteur ès-lettres, en 1975, avec une thèse intitulée "François de Wendel en République, l'Argent et le pouvoir", il est nommé profes
      </p>
    </biogHist>
    <!-- etc. -->
  </description>
  -<relations>
    -<cpfRelation xlink:href="FRAN_NP_051355" xlink:type="simple" cpfRelationType="associative" xlink:arcrole="http://www.piaaf.net/ontology/piaaf#isDirectorOf" xml:id="FRAN_re_050789-d249e369">
      <relationEntry>Bibliothèque nationale de France (1994-...)</relationEntry>
      -<dateRange>
        <fromDate standardDate="2002-03-01">mars 2002</fromDate>
        <toDate standardDate="2007-03-31">mars 2007</toDate>
      </dateRange>
    </cpfRelation>
    -<cpfRelation cpfRelationType="identity" xlink:href="http://catalogue.bnfr.fr/ark:/12148/cb11908748q" xlink:type="simple" xml:id="FRAN_re_050789-d249e430">
      <relationEntry>Jeanneney, Jean-Noël (1942-...)</relationEntry>
      -<dateRange>
        <fromDate standardDate="1942-04-02">2 avril 1942</fromDate>
      </dateRange>
      -<descriptiveNote>
        <p>Notice d'autorité BnF</p>
      </descriptiveNote>
    </cpfRelation>
    -<cpfRelation xlink:href="FRAN_NP_005419" xlink:type="simple" cpfRelationType="associative" xlink:arcrole="http://www.piaaf.net/ontology/piaaf#isDirectorOf" xml:id="FRAN_re_050789-d249e473">
      <relationEntry>Radio France</relationEntry>
      -<dateRange>
        <fromDate standardDate="1982-01-01">1982</fromDate>
        <toDate standardDate="1986-12-31">1986</toDate>
      </dateRange>
      -<descriptiveNote>
        <p>Jean-Noël Jeanneney a été président de Radio France et RFI</p>
      </descriptiveNote>
    </cpfRelation>
    -<resourceRelation resourceRelationType="creatorOf" xlink:href="FRAN_IK_050629" xlink:type="simple" xml:id="FRAN_re_050789-d249e490">
      <relationEntry localType="archival">
        <relationEntry>Archives papier de Jean-Noël Jeanneney président-directeur général de Radio France et de Radio France Internationale de 1982 à 1986 (1964-1986)</relationEntry>
      </relationEntry>
    </resourceRelation>
  </relations>
</cpfDescription>
</eac-cpf>
```

The diagram illustrates the mapping of XML elements to RDF categories:

- Name** (blue hexagon) points to the `<nameEntry>` element.
- CorporateBody** (red hexagon) points to the `<relationEntry>` element in the first `<cpfRelation>` block.
- AgentLeadershipRelation** (grey hexagon) points to the `<relationEntry>` element in the second `<cpfRelation>` block.
- Position** (blue hexagon) points to the `<relationEntry>` element in the third `<cpfRelation>` block.
- RecordSet** (brown hexagon) points to the `<relationEntry>` element in the `<resourceRelation>` block.

From one unique authority record encoded in EAC-CPF, and describing a person...

...one can generate and interconnect RDF resources of many different categories



Issues that had to be solved

- Which objects are to be described as autonomous RDF resources? What URIs should we assign to them?
- How should we represent the (dated and documented) relations between agents ?
- How should we handle the quite abstract entities described by the SIAF?
- How should we handle the objects resulting from the combination of a domain and an activity?
- How should we handle the sequences within record set aggregations?

Representing a relation between two agents

```
<cpfRelation xlink:href="FRAN_NP_005419" xlink:type="simple" cpfRelationType="associative"
  xlink:arcrole="http://www.piaaf.net/ontology/piaaf#isDirectorOf">
  <relationEntry>Radio France</relationEntry>
  <dateRange>
    <fromDate standardDate="1982-01-01">1982</fromDate>
    <toDate standardDate="1986-12-31">1986</toDate>
  </dateRange>
  <descriptiveNote>
    <p>Jean-Noël Jeanneney a été président de Radio France et RFI</p>
  </descriptiveNote>
</cpfRelation>
```

```
<rdf:Description rdf:about="http://www.piaaf.net/rerelations/FRAN_agent_relation_050789-d249e473">
  <rdf:type rdf:resource="http://www.ica.org/standards/RiC/ontology#AgentLeadershipRelation"/>
  <rdfs:label xml:lang="fr">Relation
    de "leadership" ou de direction entre les agents "Jeanneney, Jean-Noël (1942-....)" et "Radio France"</rdfs:label>
  <RiC:description xml:lang="fr">Jean-Noël Jeanneney a été président de Radio France et RFI</RiC:description>
  <RiC:beginningDate rdf:datatype="http://www.w3.org/2001/XMLSchema#date">1982-01-01</RiC:beginningDate>
  <RiC:endDate rdf:datatype="http://www.w3.org/2001/XMLSchema#date">1986-12-31</RiC:endDate>
  <RiC:leadershipBy rdf:resource="http://www.piaaf.net/agents/FRAN_person_050789"/>
  <RiC:leadershipOn rdf:resource="http://www.piaaf.net/agents/FRAN_corporate-body_005419"/>
  <!--coll='FRAN'-->
  <RiC:leadershipWithPosition rdf:resource="http://www.piaaf.net/positions/FRAN_position_050789-d249e473"/>
</rdf:Description>
```

The conversion process



0-generateIdentifiers.xsl
1-processFunctions.xsl
2-processLegalStatuses.xsl
3-processMandates.xsl
4-processEvents.xsl
5-processAgentNames.xsl
6-processPlaces.xsl
7-processPositions.xsl
8a-8d-processRelations...xsl
9-processGroupTypes.xsl
10-processAgents.xsl
11-12-recordsSets.xsl
13-processProxies.xsl



agent-names
agents
events
group-types
mandates
places
positions
proxies
record-sets
records
relations
vocabularies
functions.rdf
legal-statuses.rdf

The conversion results

940 agents (from 262 EAC-CPF records)

20 group types (from 14 EAC-CPF records)

547 agent names

335 mandates

114 events

140 places

28 positions

16 legal statuses, 91 functional domains

2062 record sets, 291 records (from 2187 <c> elements in the EAD files)

About 1600 relations (entities)

See also: <http://piaaf.demo.logilab.fr/stats/>

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <rdf:RDF xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:RiC="http://www.ica.org/standards/RiC/ontology#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:piaaf-onto="http://www.piaaf.net/ontology#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:isni="http://isni.org/ontology#"
  xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
3   <rdf:Description rdf:about="http://www.piaaf.net/agents/FRAN_person_050789">
4     <rdf:type rdf:resource="http://www.ica.org/standards/RiC/ontology#Person"/>
5     <rdfs:label xml:lang="fr">Jeanneney, Jean-Noël (1942-....)</rdfs:label>
6     <RiC:hasDenominationRelation
7   rdf:resource="http://www.piaaf.net/relations/FRAN_DenominationRelation_050789_d78e96"/>
8     <RiC:birthDate rdf:datatype="http://www.w3.org/2001/XMLSchema#date">1942-04-02</RiC:birthDate>
9     <RiC:history xml:lang="fr">Jean-Noël Jeanneney est né en avril 1942 à Grenoble. Après des études aux
10 lycées Montaigne et Louis-le-Grand à Paris, il est reçu à l'École normale supérieure de la rue d'Ulm où il est
11 élève de 1961 à 1966. Diplômé de l'Institut d'études politiques de Paris en 1964, agrégé d'histoire en 1965, il
12 consacre une année,
13 [...]
14 Il a été enfin président de la Bibliothèque nationale de France de mars 2002 à mars 2007, y conduisant en
15 particulier un combat pour la création d'une bibliothèque numérique européenne. Il est docteur honoris causa de
16 l'Université libre de Bruxelles.</RiC:history>
17     <owl:sameAs rdf:resource="http://data.bnf.fr/ark:/12148/cb11908748q#foaf:Person"/>
18     <owl:sameAs rdf:resource="http://fr.dbpedia.org/resource/Jean-No%C3%ABl_Jeanneney"/>
19     <isni:identifiantValid>0000000384320062</isni:identifiantValid>
20     <RiC:headOf rdf:resource="http://www.piaaf.net/relations/FRAN_agent_re
21     <RiC:headOf rdf:resource="http://www.piaaf.net/relations/FRAN_agent_re
22     <RiC:headOf rdf:resource="http://www.piaaf.net/relations/FRAN_agent_re
23     <RiC:headOf rdf:resource="http://www.piaaf.net/relations/FRAN_agent_re
24     <RiC:associatedWithRelation
25   rdf:resource="http://www.piaaf.net/relations/FRAN_agent_relation_050789-d249e2
26     <RiC:associatedWithRelation
27   rdf:resource="http://www.piaaf.net/relations/FRAN_agent_relation_050789-d249e2
28     <RiC:associatedWithRelation
29   rdf:resource="http://www.piaaf.net/relations/FRAN_agent_relation_050789-d249e2
30     <RiC:associatedWithRelation
31   rdf:resource="http://www.piaaf.net/relations/FRAN_agent_relation_050789-d249e3
32     <RiC:associatedWithRelation
```

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <rdf:RDF xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:RiC="http://www.ica.org/standards/RiC/ontology#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:piaaf-onto="http://www.piaaf.net/ontology#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:isni="http://isni.org/ontology#"
  xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
3   <rdf:Description rdf:about="http://www.piaaf.net/agents/FRBNF_person_0296">
4     <rdf:type rdf:resource="http://www.ica.org/standards/RiC/ontology#Person"/>
5     <rdfs:label xml:lang="fr">Jeanneney, Jean-Noël (1942-....)</rdfs:label>
6     <RiC:hasDenominationRelation
7   rdf:resource="http://www.piaaf.net/relations/FRBNF_DenominationRelation_0296_d61e109"/>
8     <RiC:birthDate rdf:datatype="http://www.w3.org/2001/XMLSchema#Year">1942</RiC:birthDate>
9     <RiC:history xml:lang="fr">Jean-Noël Jeanneney est un haut fonctionnaire de l'Etat. Il a étudié à l'École
10 normale supérieure de 1961 à 1966, est diplômé de l'Institut d'études politiques de
11 Paris en 1964, agrégé d'histoire en 1965, docteur ès lettres en 1975. Il enseigne à
12 l'Institut d'études politiques de Paris à partir de 1968 et à l'Université de Paris X
13 [...]
14 ligne de près de 39 millions de documents numérisés représentatifs du patrimoine
15 historique et culturel de l'Europe. En 2007, atteint par la limite d'âge, il quitte
16 ses fonctions.</RiC:history>
17     <RiC:conformsToRule rdf:resource="http://www.piaaf.net/mandates/FRBNF_mandate_0296-d57e216"/>
18     <RiC:conformsToRule rdf:resource="http://www.piaaf.net/mandates/FRBNF_mandate_0296-d57e222"/>
19     <owl:sameAs rdf:resource="http://data.bnf.fr/ark:/12148/cb11908748q#foaf:Person"/>
20     <isni:identifiantValid>0000000384320062</isni:identifiantValid>
21     <RiC:headOf rdf:resource="http://www.piaaf.net/relations/FRBNF_agent_relation_0296-d57e272"/>
22     <RiC:agentFollows rdf:resource="http://www.piaaf.net/relations/FRBNF_agent_relation_0296-d57e293"/>
23     <RiC:agentPrecedes rdf:resource="http://www.piaaf.net/relations/FRBNF_agent_relation_0296-d57e314"/>
24     <RiC:hasProvenanceRelation
25   rdf:resource="http://www.piaaf.net/relations/FRBNF_ProvenanceRelation_0296-d57e329"/>
26     <RiC:hasProvenanceRelation
27   rdf:resource="http://www.piaaf.net/relations/FRBNF_ProvenanceRelation_0296-d57e335"/>
28     <RiC:describedBy rdf:resource="http://www.piaaf.net/descriptions/FRBNF_description_0296"/>
29     </rdf:Description>
30   </rdf:Description rdf:about="http://www.piaaf.net/relations/FRBNF_DenominationRelation_0296_d61e109">
```

Jean-Noël Jeanneney
as he is represented in the RDF
dataset (in the archival world) of
the ANF and of the BnF

Findings

This proved to be not trivial, but could be realized and resulted in rich RDF datasets, that had the expected level of quality

- The experience with data processing by implementing RiC-O in this early state of development was **extremely encouraging**.
- It showed that:
 - **RiC-O works to express the complexity of archival description!**
 - we (specially at the ANF) could imagine going much further

Specific findings (and issues that should be carefully considered before any RDF conversion of archival metadata)

- The EAC-CPF and EAD 2002 models have some limitations
- **The objects described should have persistent local identifiers**
- As concerns the EAD files, a strategy for representing the levels of description should be defined (should all the archival units be converted or not? should the lower levels inherit some metadata of higher levels or not? If so, what metadata? Etc.)
- **The source metadata should include more controlled access points (that should be defined and described using vocabularies or authority files)**
- And, above all, **it is essential to check the quality of the metadata to be converted**

Step 3: building the web application

Specifications of the web application (reminder)

- A **specific website**, distinct from the information systems of the 3 institutions
- **Two categories of end users**: the archival community (and other cultural heritage communities) and the public, from citizens to researchers
Actually we did not have time for thinking of the general public or including it in the team (except during a workshop); we chose to focus on the needs of professionals and researchers
- **An RDF triplestore and SPARQL endpoint** (for enabling textual queries), allowing advanced search
- **Dynamic and interactive graphical visualisation, including a timeline**
- Ability to align the datasets (inferencing)
- Analysis of the input data (provided RDF datasets)
- Open source components only ; development based on small iterations and interactions with the project team (**agile approach**)

Four iterations

Iteration 1

- Kickoff meeting
- Input data analysis
- Setting up the technical infrastructure
- RDF data import
- First, partial, graph visualisation

Iterations 2 and 3

- Graph visualisation including the « diagramme chronologico-hiérarchique »
- import new version of the data + other complementary datasets
- textual research, examples of SPARQL queries, including for aligning some resources
- enhancements on default display

Iteration 4

- workshops for testing the web app with colleagues and a few other persons
- documentation

Graphes

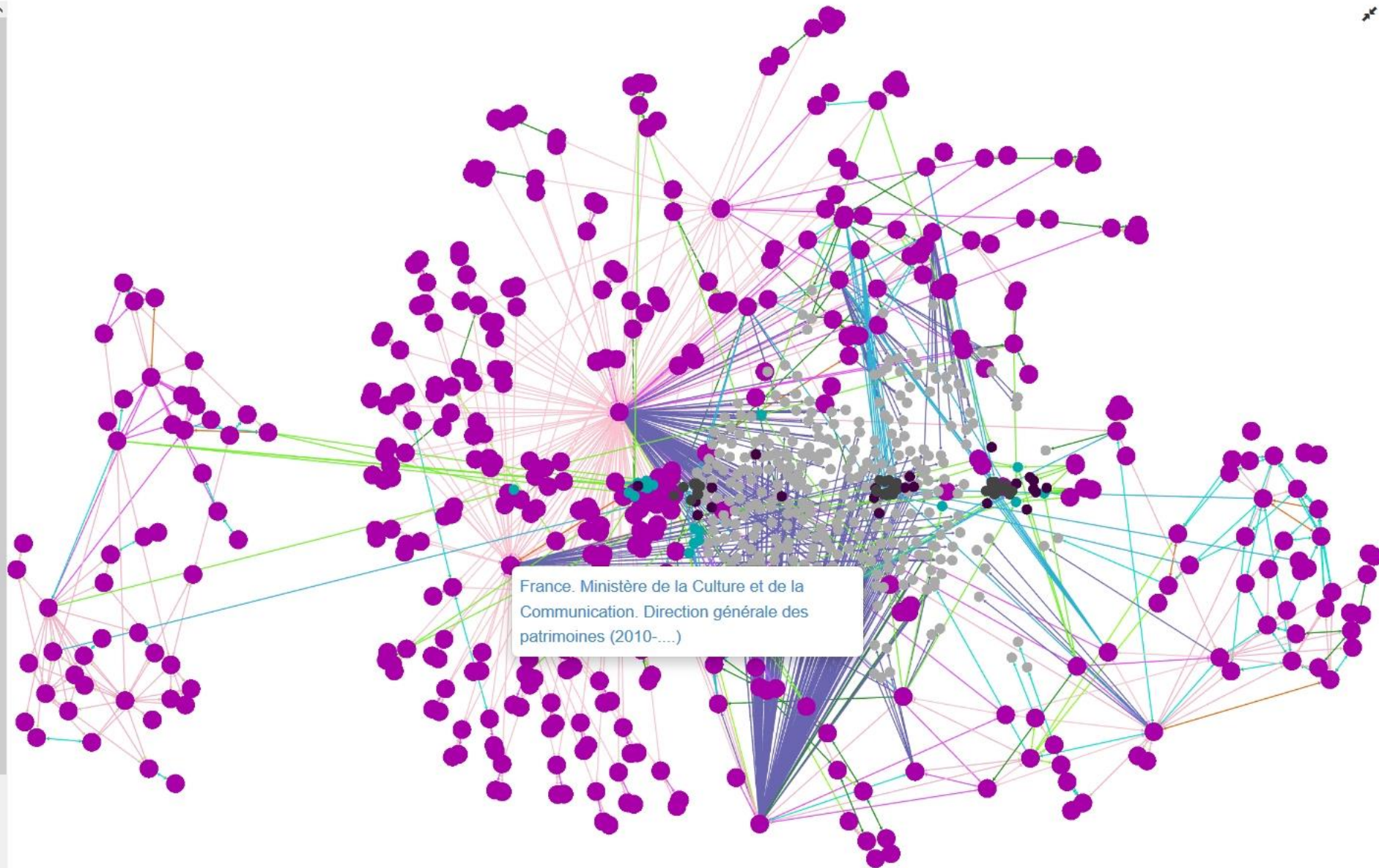
- urn:piaaf:an (813/1889)
- urn:piaaf:bnf (168/466)
- urn:piaaf:data (516/526)
- urn:piaaf:siaf (388/818)

Ressources

- <inconnu> (0/167)
- Agent (312/312)
- ★ Collectivité (438/438)
- Domaine d'activité (0/91)
- Catégorie de collectivité (0/8)
- Personne (20/20)
- Lieu (0/140)
- Poste (22/22)
- Groupe de documents (0/447)
- Groupe de documents au plus haut niveau (36/36)
- skos:Concept (0/15)

Relations

- identique à (0/186)
- a pour domaine d'activité (0/609)
- sous le contrôle ou la tutelle de (260/260)
- successeur de (143/143)
- subdivision de (295/295)
- a pour entité hiérarchiquement supérieure (17/17)
- a pour entité hiérarchiquement inférieure (325/325)
- prédécesseur de (140/140)
- a pour membre (7/7)



Dates

1568 - 2016

Jeanneney, Jean-Noël (1942-....)

http://www.plaaf.net/agents/FRAN_person_050789

Start layout

Graphes

um:plaafan

Resources

<inconnu>

ric:Agent

ric:CorporateBody

ric:Person

ric:Position

ric:RecordSet

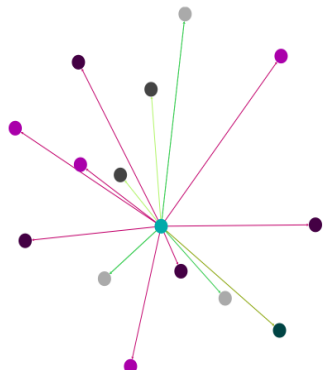
Relations

associatedWithRelation

hasProvenanceRelation

headOf

sameAs



Attributs

rdfs:label	Jeanneney, Jean-Noël (1942-....)
ric:birthDate	1942-04-02
ric:history	<p>Jean-Noël Jeanneney est né en avril 1942 à Grenoble. Après des études aux lycées Montaigne et Louis-le-Grand à Paris, il est reçu à l'École normale supérieure de la rue d'Ulm où il est élève de 1961 à 1966. Diplômé de l'Institut d'études politiques de Paris en 1964, agrégé d'histoire en 1965, il consacre une année, avant son service militaire, à faire le tour du monde grâce à une bourse Singer-Polignac. Il enseigne ensuite l'histoire contemporaine à l'université de Nanterre comme assistant puis maître-assistant. Docteur ès-lettres, en 1975, avec une thèse intitulée "François de Wendel en République, l'Argent et le pouvoir", il est nommé professeur des universités à Sciences Po en 1977. Il y enseigne depuis lors l'histoire politique, l'histoire culturelle et l'histoire des médias, sujets sur lesquels il a publié de multiples ouvrages, parmi divers livres touchant à d'autres domaines. Il est l'auteur de nombreux documentaires historiques pour la télévision, de 1973 à 2009. Il appartient à l'équipe fondatrice de la revue L'Histoire. Il a été co-président fondateur du Festival international de Radio France et de Montpellier (1985-1986). Il a présidé, entre 1991 et 2000, le Conseil scientifique de l'Institut d'Histoire du temps présent. Il a été chroniqueur au Journal du Dimanche de 1999 à 2001. Il a été membre du Conseil d'administration (1967-1991 puis 1993-1995) puis de surveillance (1995-2002) des éditions du Seuil et président du Conseil d'orientation de la chaîne de télévision câblée « Histoire » de 1997 à 2004, membre du Conseil d'administration de l'École normale supérieure entre 1998 et 2004, membre du Haut Conseil de l'Institut du monde arabe de 2003 à 2008, membre du Conseil d'administration du Fonds d'action SACEM de 1982 à 2013. Il préside le Conseil scientifique des « Rendez-vous de l'Histoire de Blois » depuis 2003. Il est président d'honneur du Festival international du film historique de Pessac depuis 1990. Il préside le jury du Prix du livre d'Histoire du Sénat depuis 2007, les Rencontres de la photographie d'Arles depuis septembre 2009. Depuis mars 1998, il co-préside l'association Europartenaies. Il produit l'émission hebdomadaire « Concoissance des temps », sur France Culture, depuis 1999. Il a exercé diverses activités publiques : président de Radio France et de Radio France Internationale de 1982 à 1986, président de la Mission du Bicentenaire de la Révolution française et de la Déclaration des droits de l'homme et du citoyen (1988-1989), secrétaire d'État au Commerce extérieur (1991-1992) puis à la Communication (1992-1993) sous la seconde présidence de François Mitterrand. Il a été conseiller régional de Franche-Comté de 1992 à 1998. Il a été enfin président de la Bibliothèque nationale de France de mars 2002 à mars 2007, y conduisant en particulier un combat pour la création d'une bibliothèque numérique européenne. Il est docteur honoris causa de l'Université libre de Bruxelles.</p>

Relations ayant cette ressource comme sujet

rdf:type	http://www.ica.org/standards/RIC/ontology#Person
owl:sameAs	http://data.bnf.fr/ark:/12148/cb11908748q#oaf:Person http://fr.dbpedia.org/resource/Jean-No%C3%A9_Jeanneney
ric:associatedWithRelation	http://www.plaaf.net/reliations/FRAN_agent_relation_050789-d249e256 http://www.plaaf.net/reliations/FRAN_agent_relation_050789-d249e279 http://www.plaaf.net/reliations/FRAN_agent_relation_050789-d249e302 http://www.plaaf.net/reliations/FRAN_agent_relation_050789-d249e322

Archives papier de Jean-Noël Jeanneney, Président-directeur général de Radio France (1964-1986)

Start layout

Graphes

um:plaaf:an

Resources

ric:Person

ric:RecordSet

Relations

originatedBy

Attributs

ric:title	Archives papier de Jean-Noël Jeanneney président-directeur général de Radio France et de Radio France Internationale de 1982 à 1986 (1964-1986)
-----------	---

Relations ayant cette ressource comme sujet

rdf:type	http://www.ica.org/standards/RIC/ontology#RecordSet
ric:mainSubjectOf	https://www.siv.archives-nationales.culture.gouv.fr/siv/IR/FRAN_IR_050629
ric:originatedBy	http://www.plaaf.net/reliations/FRAN_ProvenanceRelation_050789-d249e496

Relations ayant cette ressource comme objet

ric:originatedObject	http://www.plaaf.net/reliations/FRAN_ProvenanceRelation_050789-d249e496
----------------------	---

Demonstration

<http://piaaf.demo.logilab.fr>



A success, and its limitations

- Proof of concept done!
Though we did not have enough time for doing everything that had been foreseen
- Prototype was welcome and considered both very interesting and opening new pathes
- Another benefit of the project came from collaboration: we learned more about the practice, and perspective, of the three institutions, and could see how much they determine, and even shape, the metadata created
- Of course it remains a proof of concept...



6. What's next at the ANF? The following projects

An institutional programme: entering and successfully completing an « archival (metadata) transition »

- The ANF are now designing their information system master plan for the next years
- This plan will include a set of tasks aiming to **switch to one unique global descriptive metadata model**
Several metadata silos, several 'models', redundancies, inconsistencies, difficulties as concerns sharing knowledge, several end user interfaces...
Moving to a unique, relevant, fully documented, scalable, framework
- **Moving from semi-structured metadata to a graph of data**
- **This model will be based on RiC**
- First discussions, workshops and concrete tasks already ongoing inside the institution

Preparing data for a Linked Data repository

Home > Getting Started

Getting Started

Test drive the converter

To run your first conversion :

1. double-click on `ricoconverter.bat` (or `ricoconverter.sh` on Linux or Mac - make sure it is executable),
2. hit `Enter` to leave the default values for the input parameters of the script;
3. wait a couple of seconds for the conversion to finish, and look in the generated output folder `output-<date_of_day>` for the result of the conversion;

The few sample files provided by default in the `input` folder have been converted to RiC-O.

Display the Help message

To print the help message and get the details of all the possible commands and options, run `ricoconverter.bat` and type "help" at the prompt for the command :

```
"" :: Welcome to Ric-O Converter 20190604 ::
```

```
Enter command to execute (convert_arrange, convert, validate, test, version, help) [press enter for 'convert_arrange']:help ""
```

Script folder content

In the script folder you will find the following directories :

- `documentation` : contains this documentation :-)
- `input`, `input-50`, `input-500`, `input-5000` and `input-all` : contains samples EAC files to be converted with the `convert_arrange` and `convert` commands;
- `input-validation` : the default input folder for validating data using the `validate` command;
- `parameters` : contains the parameter files for each command; you need to modify these files if you want to adjust the options of a command;
- `shapes` : contains the shapes definition file used by the `validate` command;
- `unit-tests` : contains the unit test files used by the `test` command;
- `vocabularies`
- `xslt` : contains the XSLT files used by the `convert_arrange` and `convert` commands to convert EAC to RiC-O; the main one is `eac2rico.xslt`;

Commands and parameters

The converter contains a few commands, the major ones are `convert_arrange`, `convert`, and `validate`.

Each command is associated to a set of parameters, and these parameters are listed in a property file under the `parameters` subfolder, in a file with the same name as the command; e.g. the options for the command `convert_arrange` are in `parameters/convert_arrange.properties`.

Designing a « semantic web module » is a subset of the IS master plan.

This will most probably be a front end module.

Meanwhile, we are working on converting the whole of our EAD and EAC-CPF metadata to RDF.

Developing **ricoconverter**:

- an **open source** tool
- **easy to install and run**
- configurable
- documented in English
- including unit tests
- fast and efficient
- **public official release in January**

2020

Enhancing the quality of our descriptive metadata

Consistency and accuracy of metadata depend on using authority records and controlled vocabularies.

Also, authority data:

- are bridges allowing interconnexions within our metadata and to other metadata;
- can help to build efficient, easy to understand, end user interfaces

Enhancing quality implies:

- that our authority data be richer, better structured, aligned to other ones

Several projects, including: enriching the lists of places so that they are described as geo-historical entities ; building a thesaurus for indexing the functions of corporate bodies; creating authority records for any agent which has relations with the records we keep; etc.

- that they are much more used by our colleagues

Several projects including: changing the IS functionalities; training courses, events on data quality, changing practice; a named entities recognition project

In short

Medium to long-term projects

Collective and collaborative by essence

They also are articulated with the strategies and metadata repositories of other institutions, teams, portals and research projects.



Thank you for your attention!