

Le nouveau standard *Records in Contexts* : un modèle conceptuel et une ontologie OWL pour la description des archives

Support de formation distribué sous licence Creative Commons [CC-BY-SA 4.0](#).

Préparé par Florence Clavaud

(Archives nationales, Lab | Conseil international des Archives, groupe d'experts EGAD ;
courriel : florence.clavaud@culture.gouv.fr)

pour les étudiants du Master 2 TNAH de l'Ecole nationale des Chartes (cours ENC 2408 sur la description des archives)

Date : février-mars 2022

Contenu

1. Records in Contexts : informations générales
2. Le modèle conceptuel RiC-CM
3. L'ontologie RiC-O
4. Le cas d'usage des Archives nationales
5. Proposition d'exercices : produire et utiliser des données RDF/RiC-O
6. Conclusion

Quelques précisions pour commencer

En principe, vous devez déjà savoir :

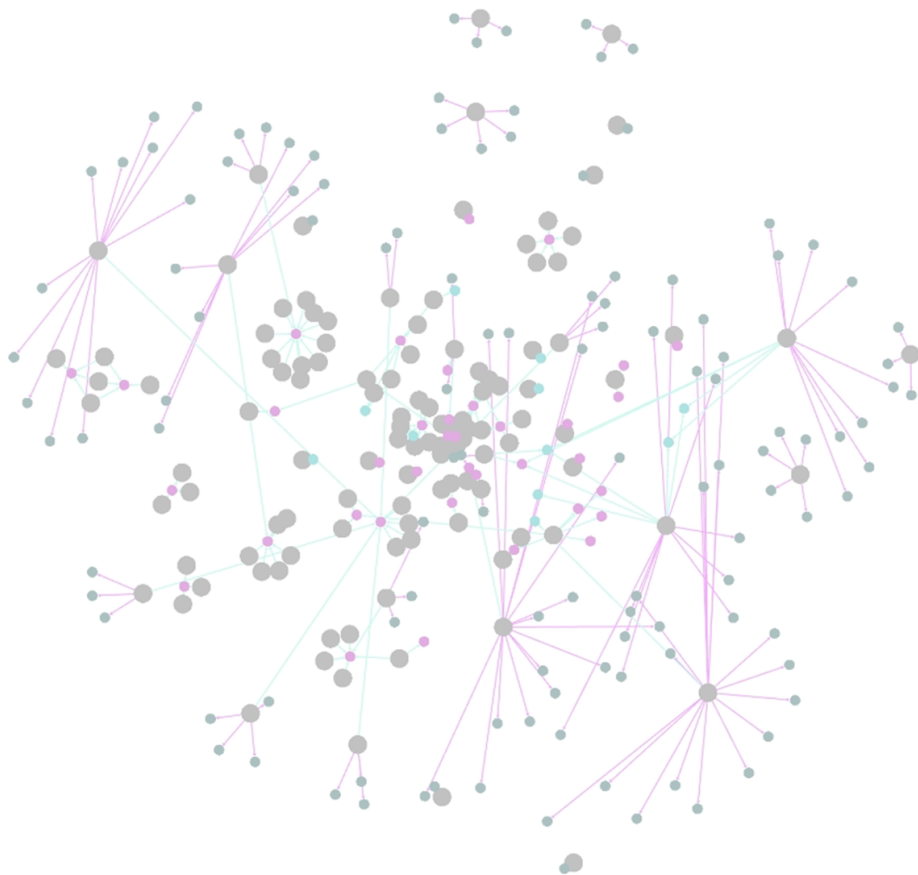
- Ce que sont les **archives**
- Ce que c'est qu'une **norme** et pourquoi, de manière générale, il est bon de connaître les normes applicables à son domaine professionnel
- Quelles sont les **normes qui encadrent actuellement la description archivistique**, et en particulier ce qu'il faut absolument retenir de la norme principale (dont le nom est ?)
- Quels sont les **formats techniques existants qui correspondent à ces normes**, et en gros ce qu'ils permettent
- Ce que c'est que le **web de données, RDF** et **une ontologie**

À la fin de ce cours, vous devriez :

- Connaître l'essentiel du standard RiC et des raisons et enjeux de l'évolution des pratiques de la description archivistique
- Avoir une vision globale de RiC-O
- Disposer de repères pour réfléchir à une stratégie de mise en œuvre de RiC-O et pour produire des données RDF à partir de métadonnées archivistiques existantes (notamment, comment passer d'EAD à RDF)

Noter cependant que ce cours est juste une courte formation introductive.

1. Records in Contexts (RiC) : informations générales

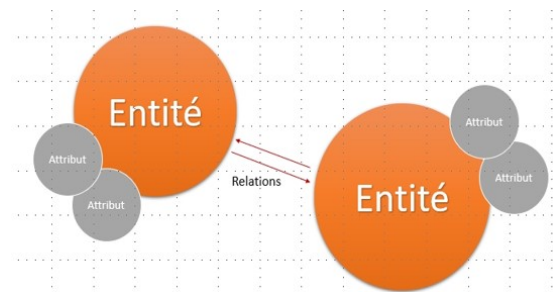
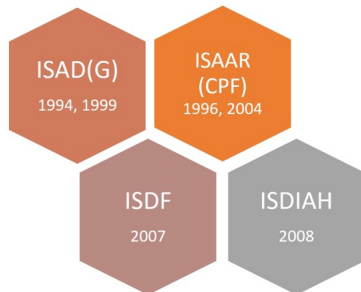


Records in Contexts (RiC), une norme en quatre parties

- un court document d'**introduction à la description archivistique** (RiC-Introduction to Archival Description ou RiC-IAD)
- Un **modèle conceptuel abstrait**, indépendant de toute implémentation (RiC-Conceptual Model ou RiC-CM)
- Une **ontologie OWL**, transposition de RiC-CM, pour définir le vocabulaire et les règles applicables aux métadonnées archivistiques conformes à RiC et ayant la forme de jeux de données RDF (RiC-Ontology ou RiC-O)
- Un **manuel de mise en œuvre** (RiC-Application Guidelines, RiC-AG)

Records in Contexts : pourquoi ?

- Pour **remplacer les quatre standards internationaux existants**



- Pour passer de principes et règles de description à **une norme unique fondée sur une approche entités-relations**
- Pour **mieux représenter les archives, leur histoire et les multiples couches de *contextes* dans lesquelles elles s'inscrivent**
- Pour permettre aux archivistes et aux records managers de passer à une **description plus précise, plus nuancée, plus facile à exploiter, multidimensionnelle**
- Et pour, de ce fait, **faciliter l'accès aux métadonnées archivistiques donc aux ressources documentaires qu'elles décrivent**

Cf. la transition en cours des métadonnées bibliographiques vers des graphes de données (modèle IFLA-LRM et nouvelles règles de catalogage RDA)

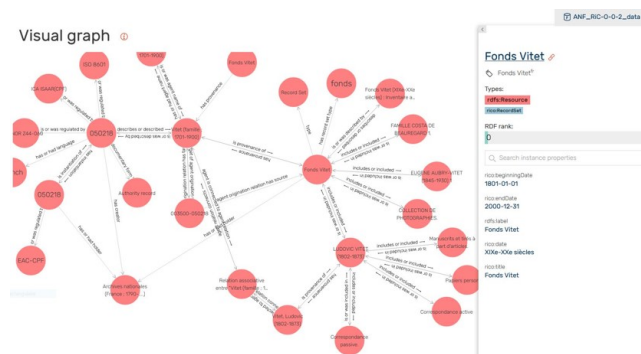
Records in Contexts : voir le monde des archives comme un graphe orienté (exemple)

A gauche, la vue HTML, dans la SIV des Archives nationales, de l'inventaire du fonds de la famille Vitet conservé aux AN sous la cote 572AP (source : un fichier EAD) https://www.siv.archives-nationales.culture.gouv.fr/siv/IR/FRAN_IR_003500

A droite, les informations de haut niveau du même inventaire, après conversion en RDF/RiC-O et import dans un triplestore RDF (image source disponible ici : https://raw.githubusercontent.com/ICA-EGAD/RiC-O/master/examples/examples_v0-2/NationalArchivesOfFrance/GraphDB-screenshot_recordResource-003500-top_FondsVitet_vRiC-O-0-2.jpg)

The screenshot shows the 'INVENTAIRE' page for 'Fonds Vitet (XIXe-XXe siècles)' with the cote '572AP/1-572AP/122'. It features a search bar and a list of inventory items with their respective cotes and descriptions.

Cotes	Description	Action
572AP/1-572AP/20	LUDOVIC VITET (1802-1873)	Voir le détail
572AP/21-572AP/95	EUGENE AUBRY-VITET (1845-1930) 1	Voir le détail
572AP/96-572AP/111	FAMILLE COSTA DE BEAUREGARD 1.	Voir le détail
572AP/112-572AP/122	COLLECTION DE PHOTOGRAPHIES.	Voir le détail



RiC est développé par le groupe EGAD

- EGAD : *Expert Group on Archival Description* (voir <https://www.ica.org/en/egad-steering-committee-0>)
- Un des groupes d'experts du Conseil international des Archives (ICA)
- Mandat : 2013-2016, renouvelé depuis
- Actuellement 21 personnes, venant de nombreux pays ; différentes fonctions et compétences, différents points de vue
- Tous les discussions et travaux sont menés en anglais
- Travail régulièrement soumis à commentaires pour faire participer les communautés professionnelles concernées
- Devra impérativement être poursuivi en continu ensuite

Historique et feuille de route globale de RiC

- Août 2016 : **publication de RiC-CM 0.1** avec un appel à commentaires
 Entre août 2016 et janvier 2017, EGAD a reçu de très nombreux commentaires, et les a pris en compte depuis
- Février-novembre 2019 : plusieurs versions beta de RiC-O successivement envoyées aux personnes qui avaient répondu à un appel à commentateurs
- 12 décembre 2019 : **publication d'une première version officielle de l'ontologie, RiC-O 0.1** (<https://www.ica.org/standards/RiC/ontology>) assortie d'un appel à commentaires ; **publication d'une prévisualisation de RiC-CM 0.2 qui remplace RiC-CM 0.1** devenu obsolète
- Mars 2020 : **ouverture de l'entrepôt de sources de RiC-O sur GitHub**
- 12 février 2021 : **publication de RiC-O 0.2**, qui remplace RiC-O 0.1 (toujours à <https://www.ica.org/standards/RiC/ontology>) et qui est conforme à RiC-CM 0.2
- Juillet 2021 : **publication du modèle conceptuel RiC-CM 0.2** ; appel à commentaires
- Janvier 2022 : **publication de RiC-IAD**

La **publication de la version 1.0 de RiC**, qui aura le statut de recommandation officielle de l'ICA, est prévue pour fin 2022 ou début 2023.

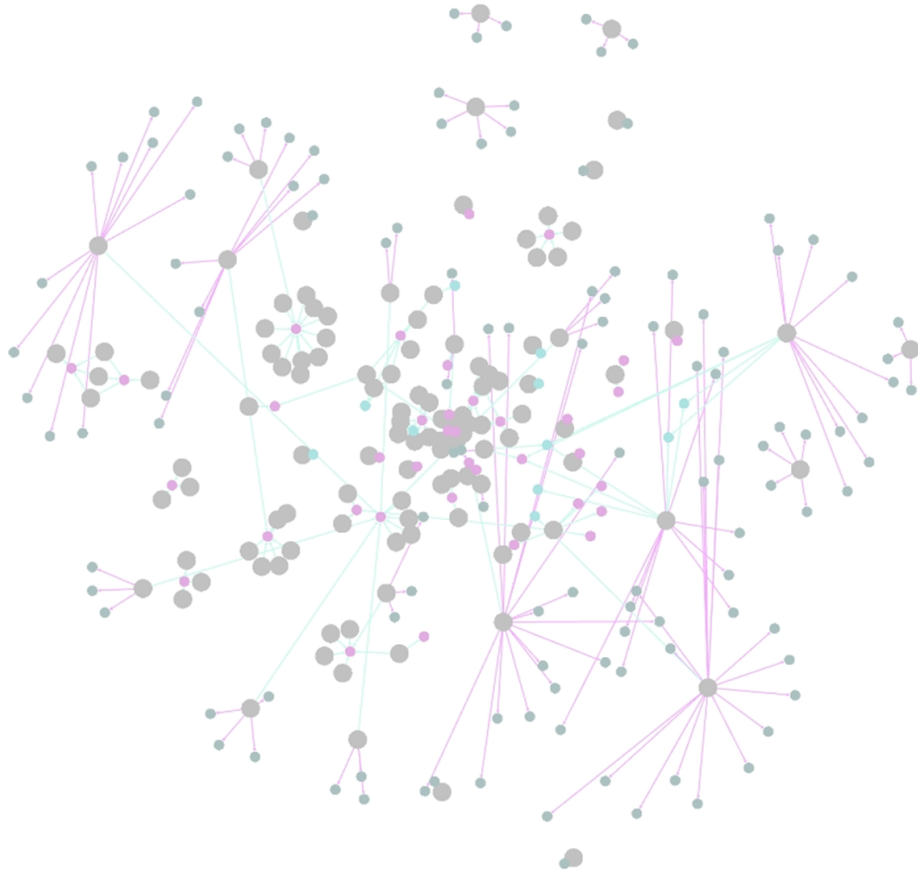
Toutes les versions de RiC antérieures à la v1.0 sont par définition des *drafts* (projets). Cependant **RiC-CM et RiC-O 0.2 sont proches de cette version 1.0 et peuvent déjà être employés.**

Comment se tenir au courant des avancées de RiC ?

Les principales actualités sur RiC sont publiées sur le [site web du CIA](#).

- Sur RiC-CM : surveiller le site web du CIA et les listes professionnelles françaises de discussion des archivistes (AAF, SIAF), ou la liste de l'association Humanistica
- Sur RiC-O :
 - Consulter régulièrement le [site web officiel de RiC-O](#), qui fournit également des informations sur les projets, événements et outils
 - Si on souhaite récupérer le dernier état de RiC-O, des exemples, participer aux discussions et travaux en cours : consulter [l'entrepôt de sources sur GitHub](#)
 - Il est possible de s'abonner à une liste des utilisateurs de RiC-O
- **Ne pas hésiter à me contacter** par courriel (florence.clavaud@culture.gouv.fr), pour toute question, remarque, demande de conseil ou de présentation

2. Le modèle conceptuel RiC-CM



RiC-CM définit et organise 22 entités

Les entités dans un modèle conceptuel sont les **principaux objets d'intérêt du domaine envisagé**. Les entités identifiées dans RiC-CM sont celles qui permettent de fournir les éléments de contexte nécessaires pour la gestion physique, la conservation, la découverte, l'utilisation et la compréhension des documents d'archives. Les quatre entités considérées comme essentielles sont indiquées en gras dans le tableau ci-dessous (Record Resource, Instantiation, Agent, Activity). Les entités sont organisées en hiérarchie d'objets. Les attributs et relations d'une entité de haut niveau sont héritées par les entités de niveau inférieur.

RiC Entities Hierarchy			
First Level	Second Level	Third Level	Fourth Level
RiC-E01 Thing	RiC-E02 Record Resource	RiC-E03 Record Set	
		RiC-E04 Record	
		RiC-E05 Record Part	
	RiC-E06 Instantiation		
	RiC-E07 Agent	RiC-E08 Person	
		RiC-E09 Group	RiC-E10 Family
			RiC-E11 Corporate Body
		RiC-E12 Position	
		RiC-E13 Mechanism	
	RiC-E14 Event	RiC-E15 Activity	
	RiC-E16 Rule	RiC-E17 Mandate	
	RiC-E18 Date	RiC-E19 Single Date	
		RiC-E20 Date Range	
		RiC-E21 Date Set	
	RiC-E22 Place		

Les entités RiC-CM ont des attributs

Il y a 41 attributs :

- Des attributs communs à toutes les entités : identifier, name, descriptive note
- Des attributs pour chacune des entités de haut niveau (les sous-entités « héritent » de ces attributs)
- Des attributs propres à certaines entités

Certains attributs sont répétables ou extensibles.

ID	Assigned identifier beginning with RIC-A and followed by a number.
Name	Natural language label.
Definition	Brief definition.
Domain	Entity or Entities which may have the Attribute
Specifications	The possible specifications, precisions, annotations or qualifiers of the meaning of the value an attribute in a description.
Extensibility	The possible differentiation of the attribute using extensions or specializations in sub-attributes.
Repeatability	Whether the attribute can be repeated or not in the description of a particular entity or relation.
Value Schema	The rules for selecting or formulating the value of the attribute.
Scope	It provides additional information to understand and use the attribute.
Examples	It presents examples of application of each attribute.

Exemple : les attributs des entités RiC-E04 Record et de RiC-E06 Instantiation

RiC-E04	Record
Attribute ID	Attribute Name
RiC-A16	Descriptive Note
RiC-A22	Identifier
RiC-A28	Name
RiC-A21	History
RiC-A35	Record Resource Extent
RiC-A38	Scope and content
RiC-A40	Structure
RiC-A03	Authenticity Note
RiC-A07	Classification
RiC-A08	Conditions of Access
RiC-A09	Conditions of Use
RiC-A10	Content Type
RiC-A17	Documentary Form Type
RiC-A24	Integrity
RiC-A25	Language
RiC-A26	Legal Status
RiC-A39	State

RiC-E06	Instantiation
Attribute ID	Attribute Name
RiC-A16	Descriptive Note
RiC-A22	Identifier
RiC-A28	Name
RiC-A03	Authenticity Note
RiC-A04	Carrier Extent
RiC-A05	Carrier Type
RiC-A08	Conditions of Access
RiC-A09	Conditions of Use
RiC-A21	History
RiC-A23	Instantiation Extent
RiC-A24	Integrity
RiC-A31	Physical Characteristics
RiC-A33	Production Technique
RiC-A34	Quality of Representation
RiC-A37	Representation Type
RiC-A40	Structure

Les entités RiC-CM sont reliées entre elles par des relations

Dans RiC-CM 0.2, il y a **78 relations** (et les relations inverses le cas échéant), organisées en **système poly-hiérarchique**. La relation de haut niveau est 'is related to'.

Elles ont des attributs (identifier, description, date, certainty, source).

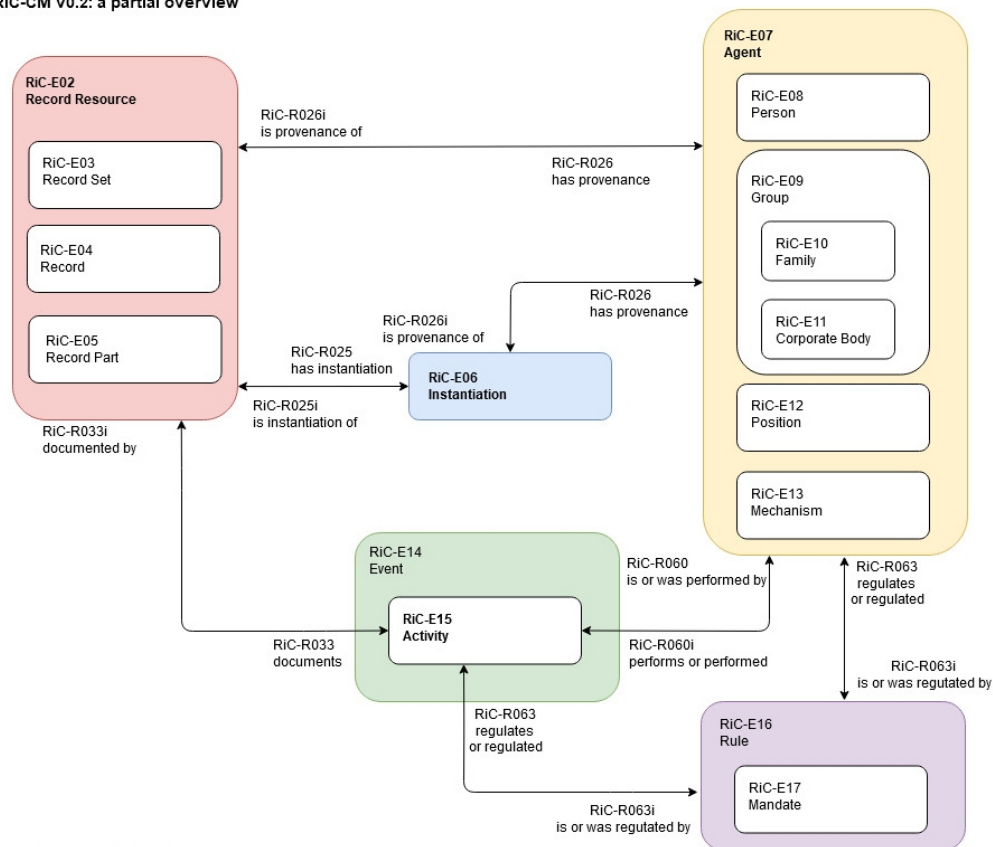
- Whole-part relations
- Sequential and temporal relations
- Subject relations
- Record Resource to Record Resource relations
- Record Resource to Instantiation relations
- Provenance relations
- Instantiation to Instantiation relations
- Management relations
- Agent to Agent relations
- Event relations
- Rule relations
- Date relations
- Spatial relations

<p>Type: <u>Record Resource to Record Resource relations</u></p> <p>RIC-R022: Record Resource <i>is record resource associated with</i> record resource Record Resource</p>	<p>RIC-R023: Record resource <i>has genetic link to</i> Record Resource Record Resource</p> <p>RIC-R013: Record Resource <i>has reply</i> Record Resource (see also above)</p> <p>RIC-R003: Record <i>has or had constituent</i> Record Part (see also above)</p> <p>RIC-R024: Record Set <i>includes or included</i> Record or Record Set</p>	<p>RIC-R010: Record <i>is original of</i> Record (see also above)</p> <p>RIC-R011: Record <i>is draft of</i> Record (see also above)</p> <p>RIC-R012: Record Resource <i>has copy</i> Record Resource (see also above)</p>	
<p>Type: <u>Record Resource to Instantiation relations</u></p> <p>RIC-R025: Record Resource <i>has instantiation</i> <u>Instantiation</u></p>			

[Sommaire](#)

Les principales entités RiC-CM et quelques-unes des relations existantes

RiC-CM v0.2: a partial overview

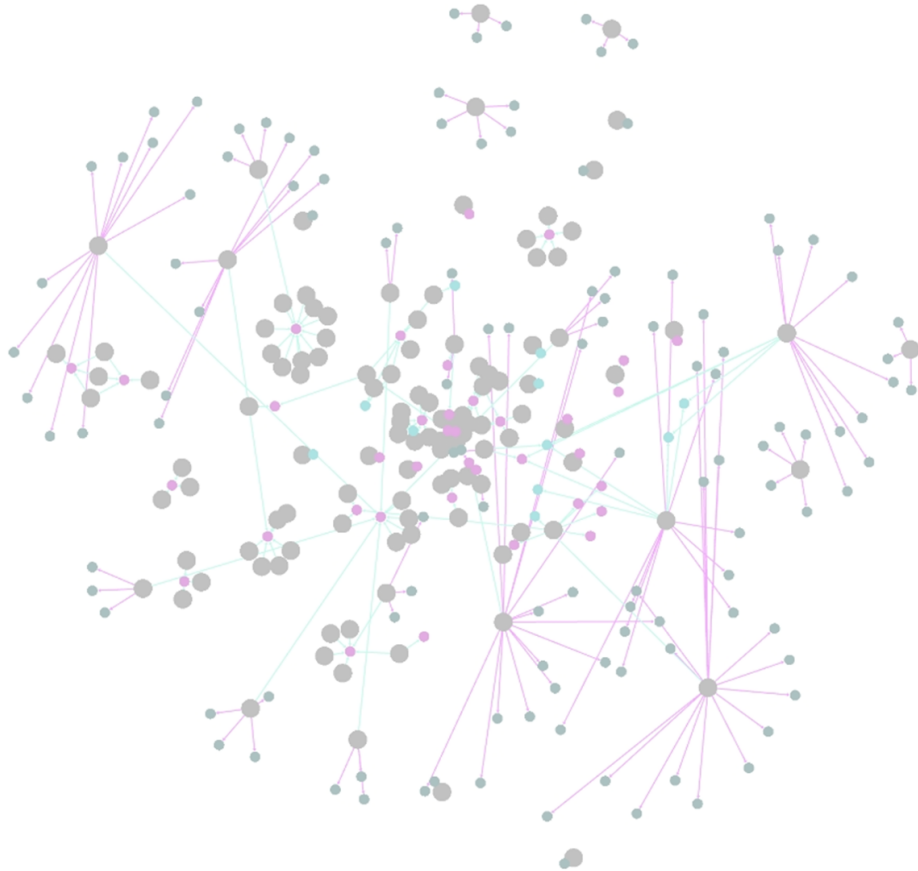


Created by ICA-EGAD in September 2020.

This diagram is licensed under a Creative Commons Attribution 4.0 International License.



3. L'ontologie RiC-0



RiC-O, qu'est-ce que c'est (rappel)?

- une **ontologie OWL générique pour la description des archives et de leurs entités contextuelles**
- une **transposition formelle du modèle conceptuel Records in Contexts-Conceptual Model (RiC-CM)**
 - totalement conforme avec RiC-CM 0.2
 - a joué et joue toujours un rôle majeur dans les tests relatifs au contenu et à la logique de RiC-CM (développement simultané)
 - documentée en anglais, avec introduction détaillée et spécifications de chaque composant

RDF, RDFS, OWL

- RDF (Resource Description Framework ; une recommandation du W3C depuis 1999, version actuelle : RDF 1.1 ; voir <http://www.w3.org/RDF/>) ;
- RDFS (the RDF Schema language ; une recommandation du W3C depuis 2004 ; voir <http://www.w3.org/TR/rdf-schema/>);
- OWL (Web Ontology Language, première version en 2004, 2e version en 2009 ; voir <http://www.w3.org/2001/sw/wiki/OWL>)

Une ontologie, en sciences de l'information, est un modèle de données représentatif d'un ensemble de concepts dans un domaine, et des relations entre ces concepts. Elle permet de raisonner à propos des objets du domaine concerné.

Voir aussi la définition de Gruber : "An ontology is an explicit specification of a conceptualization. [...] A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose." Thomas R. Gruber, Towards Principles for the Design of Ontologies Used for Knowledge Sharing in Formal Ontology in Conceptual Analysis and Knowledge Representation, Kluwer Academic Publishers, 1993, [en ligne](#).

Principes d'élaboration de RiC-O

Disposer d'un modèle OWL qui soit :

- une transposition formelle de RiC-CM
- une **ontologie de référence**, de domaine
- **immédiatement utilisable**
- flexible
- **utile**
- extensible

Voir aussi :

- [RiC-O design principles](#)
- [Why use RiC-O?](#)

RiC-O doit être utile : faire des recherches sur le graphe produit

1 # Les entités qui ont succédé à une entité donnée (ici, n° 51120, France. Ministère de l'Intérieur. Direction des Beaux-Arts. Bureau des Monuments historiques (1839-1848))
2
3 PREFIX rico: <http://www.ica.org/standards/RiC/ontology#>
4 PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
5 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
6
7 SELECT ?entite ?nom ?dateDebut ?dateFin WHERE {
8 <http://data.archives-nationales.culture.gouv.fr/agent/051120> (rico:agentIsSourceOfAgentTemporalRelation/rico:agentTemporalRelationHasTarget)+ ?entite .
9 ?entite rdfs:label ?nom; rico:beginningDate ?dateDebut.
10 OPTIONAL { ?entite rico:endDate ?dateFin. }
11
12
13

Table Raw Response Pivot Table Google Chart Download as

Filter query results Showing results from 1 to 10 of 10. Query took 0.2s, today at 14:48.

	entite	nom	dateDebut	dateFin
1	http://data.archives-nationales.culture.gouv.fr/agent/051121	France. Ministère de l'Intérieur. Direction des Beaux-Arts. Bureau des Beaux-Arts. des Musées et des Monuments historiques (1848-1852) ⁹⁹	*1848-01-01**xsd:date	*1852-12-31**xsd:date
2	http://data.archives-nationales.culture.gouv.fr/agent/051122	France. Ministère d'État. Division des Beaux-Arts. Service des Monuments historiques (1853-1858) ⁹⁹	*1853-01-01**xsd:date	*1858-12-31**xsd:date
3	http://data.archives-nationales.culture.gouv.fr/agent/051123	France. Ministère d'État. Bureau des Monuments historiques (1858-1863) ⁹⁹	*1858-01-01**xsd:date	*1863-12-31**xsd:date
4	http://data.archives-nationales.culture.gouv.fr/agent/051124	France. Ministère de la maison de l'Empereur. Surintendance des Beaux-Arts. Bureau des Monuments historiques (1863-1870) ⁹⁹	*1863-01-01**xsd:date	*1870-12-31**xsd:date
5	http://data.archives-nationales.culture.gouv.fr/agent/051125	France. Ministère de l'Instruction publique. Direction des Beaux-Arts. Bureau des Monuments historiques (1870-1907) ⁹⁹	*1870-01-01**xsd:date	*1907-12-31**xsd:date
6	http://data.archives-nationales.culture.gouv.fr/agent/051126	France. Ministère de l'Éducation nationale. Direction des Beaux-Arts. Bureau des Monuments historiques (1907-1939) ⁹⁹	*1907-01-01**xsd:date	*1939-12-31**xsd:date
7	http://data.archives-nationales.culture.gouv.fr/agent/051149	France. Ministère de l'Éducation nationale. Secrétariat général des Beaux-Arts. Direction des services d'architecture. Bureau des Monuments historiques et Sites (1940-1944) ⁹⁹	*1940-01-01**xsd:date	*1944-12-31**xsd:date
8	http://data.archives-nationales.culture.gouv.fr/agent/051150	France. Ministère de l'Éducation nationale. Direction de l'architecture. Sous-direction des Monuments Historiques et des Sites (18 août 1945-8 janvier 1959) ⁹⁹	*1945-08-18**xsd:date	*1959-01-08**xsd:date
9	http://data.archives-nationales.culture.gouv.fr/agent/051151	France. Ministère de la Culture et de la Communication. Direction de l'Architecture et du Patrimoine. Sous-direction des Monuments Historiques et des Espaces protégés (1959-2009) ⁹⁹	*1959-01-08**xsd:date	*2009-12-31**xsd:date
10	http://data.archives-nationales.culture.gouv.fr/agent/051153	France. Ministère de la Culture et de la Communication. Direction générale des patrimoines. Service du patrimoine. Sous-direction des Monuments historiques et des Espaces protégés (2009-...) ⁹⁹	*2009-01-01**xsd:date	

keyboard shortcuts

Une requête SPARQL sur le graphe des producteurs d'archives des Archives nationales (actuellement, 15000 notices d'autorité en EAC-CPF) (copie d'écran d'une instance locale de GraphDB). On peut ainsi poser de nouvelles questions (en parcourant le graphe via les arcs/reliations)

Accéder à RiC-O (pour les humains)

La version officielle de RiC-O (donc, en l'occurrence, RiC-O 0.2) est accessible via son URI:
<https://www.ica.org/standards/RiC/ontology>.



International Council on Archives Records in Contexts Ontology (ICA RiC-O) version 0.2

Ontology IRI: <https://www.ica.org/standards/RiC/ontology>

This HTML view: https://www.ica.org/standards/RiC/RIC-O_v0-2.html

OWL source file: https://www.ica.org/standards/RiC/RIC-O_v0-2.rdf

Ontology metadata

This version: Version 0.2 - 2021-02-12.

See also the project on GitHub, where other resources can be found, and the accompanying website: <https://github.com/ICA-EGAD/RiC-O> (the website is at <https://ica-egad.github.io/>)

Prior version: https://www.ica.org/standards/RiC/RIC-O_v0-1.rdf

Creators:

Florence Clavaud (Archives nationales de France), member of EGAD, lead of EGAD RiC-O team
International Council on Archives Expert Group on Archival Description (ICA EGAD)

Contributors:

Aaron Rubinstein (University of Massachusetts Amherst, USA), member of EGAD
Daniel Pitt (University of Virginia, USA), chair of EGAD
Mia Herrala (National Archives of Finland), member of EGAD
Tobias Wild (Docuteam GmbH, Switzerland), member of EGAD

Publisher: International Council on Archives

Rights: Copyright 2019-..... International Council on Archives (ICA)

License: <https://creativecommons.org/licenses/by/4.0/>

Abstract

RiC-O (Records in Contexts-Ontology) is an OWL ontology for describing archival record resources. As the second part of Records in Contexts standard, it is a formal representation (CM).

The current official version is **v0.2**, it is compliant with RiC-CM v0.2 full draft, that will be published in February or March 2021, and that is slightly different from [RiC-CM v0.2 preview](#).

Cet URI permet d'atteindre une représentation HTML de l'ontologie, qui donne à lire toute sa documentation interne, actuellement uniquement en anglais.

L'entrepôt de sources de RiC-O permet en outre de télécharger :

- la [dernière version en date de RiC-O](#)
- Des [exemples](#)
- Des [diagrammes](#)

Le fichier source de la version officielle de RiC-O

Pour le télécharger directement, utiliser le lien présent en haut de la vue HTML.

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns="https://www.ica.org/standards/RiC/ontology#"
  xmlns:cc="http://creativecommons.org/ns#" xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#" xmlns:html="http://www.w3.org/1999/xhtml"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rico="https://www.ica.org/standards/RiC/ontology#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:vann="http://purl.org/vocab/vann/"
  xmlns:voaf="http://purl.org/vococommons/voaf#" xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:ric-dft="https://www.ica.org/standards/RiC/vocabularies/documentaryFormTypes#"
  xmlns:ric-rst="https://www.ica.org/standards/RiC/vocabularies/recordSetTypes#"
  xml:base="https://www.ica.org/standards/RiC/ontology">
  <owl:Ontology rdf:about="https://www.ica.org/standards/RiC/ontology">
    <cc:license rdf:resource="https://creativecommons.org/licenses/by/4.0/" />
    <dc:contributor xml:lang="en">Aaron Rubinstein (University of Massachusetts Amherst, USA),
      member of EGAD</dc:contributor>
    <dc:contributor xml:lang="en">Daniel Pitti (University of Virginia, USA), chair of
      EGAD</dc:contributor>
    <dc:contributor xml:lang="en">Miia Herrala (National Archives of Finland), member of
      EGAD</dc:contributor>
    <dc:contributor xml:lang="en">Tobias Wildi (Docuteam GmbH, Switzerland), member of
      EGAD</dc:contributor>
    <dc:creator xml:lang="en">Florence Clavaud (Archives nationales de France), member of EGAD,
      lead of EGAD RiC-O team</dc:creator>
    <dc:creator xml:lang="en">International Council on Archives Expert Group on Archival
      Description (ICA EGAD)</dc:creator>
    <dc:publisher xml:lang="en">International Council on Archives</dc:publisher>
    <dc:rights xml:lang="en">Copyright 2019-..., International Council on Archives
      (ICA)</dc:rights>
    <dcterms:abstract rdf:parseType="Literal">{16 lines}
    <dcterms:contributor rdf:resource="http://www.isni.org/0000000039584825"/>
    <dcterms:creator rdf:resource="https://data.bnf.fr/ark:/12148/cbl69702205#about"/>
    <dcterms:description rdf:parseType="Literal">{393 lines}
    <dcterms:publisher rdf:resource="http://www.wikidata.org/entity/Q1421986"/>
    <dcterms:title xml:lang="en">International Council on Archives Records in Contexts Ontology
      (ICA RiC-O) version 0.2</dcterms:title>
    <vann:preferredNamespacePrefix>rico</vann:preferredNamespacePrefix>
    <vann:preferredNamespaceUri>https://www.ica.org/standards/RiC/ontology#</vann:preferredNamespaceUri>
    <rdfs:label xml:lang="en">International Council on Archives Records in Contexts Ontology
      (ICA RiC-O) version 0.2</rdfs:label>
    <owl:priorVersion rdf:resource="https://www.ica.org/standards/RiC/RiC-O_v0-1.rdf"/>
    <owl:versionInfo xml:lang="en">Version 0.2 - 2021-02-12.</owl:versionInfo>
    <skos:historyNote rdf:parseType="Literal">
      <html:div xml:lang="en">
        <html:p>A first beta version of RiC-O was developed in 2017 and used by the National
          Archives of France for building a proof of concept (<html:a
```

C'est un fichier XML conforme à [OWL 2](#) (qui peut bien sûr être converti aisément dans d'autres formats de sérialisation de RDF, comme [Turtle](#)). Utiliser au moins un éditeur XML pour l'ouvrir et le manipuler, ou, mieux, un **éditeur d'ontologies** comme [Protégé](#).

Il y a plus de composants dans RiC-O que dans RiC-CM

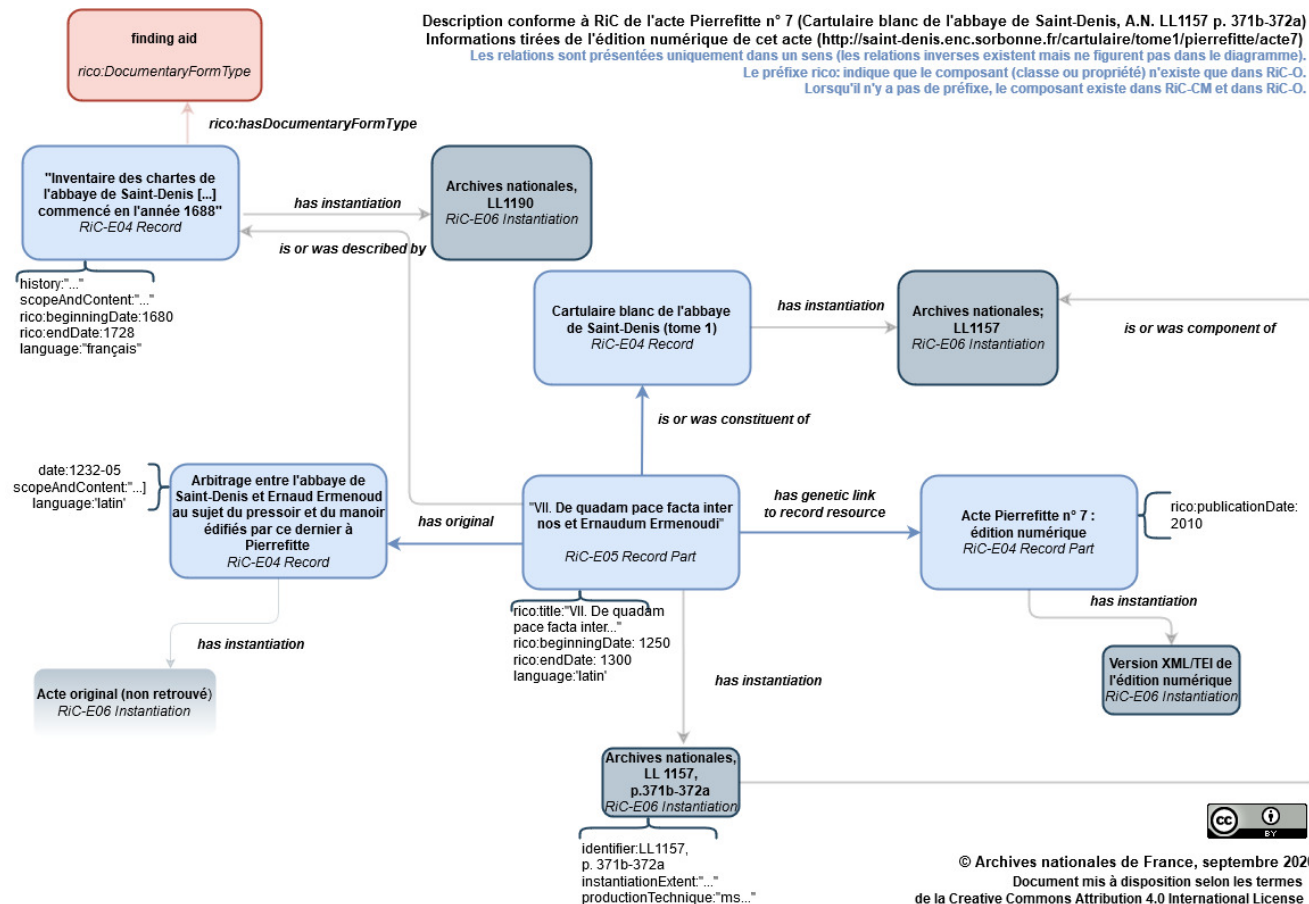
Comme dans de nombreuses représentations formelles, techniques, d'un modèle, RiC-O est plus précise et plus riche que RiC-CM.

- Les entités RiC-CM y sont représentées par des classes
- D'autres classes ont été ajoutées pour pouvoir donner une définition plus précise d'une entité. Ainsi l'entité RiC-CM Place est représentée dans RiC-O par les classes Place, Physical Location et Coordinates
- La plupart des attributs dont le type de contenu est du texte libre sont des 'datatype properties' dans RiC-O (elles ont pour objet du texte). Par ex. *rico:history* ou *rico:scopeAndContent*
- Par ailleurs, certains attributs sont aussi représentés par des classes, lorsqu'on a besoin de leur assigner des propriétés donc de les considérer comme des objets à part entière. C'est le cas pour les attributs *type* qui ont des valeurs contrôlées dans RiC-CM. C'est aussi le cas pour les attributs Language, Name et Identifier de RiC-CM.
- Dans certains cas, pour assurer une certaine flexibilité au modèle et permettre de l'utiliser immédiatement, RiC-O propose plusieurs solutions de modélisation. C'est le cas pour les relations RiC-CM, qui sont représentées par des propriétés binaires dans RiC-O, mais aussi par un système de classes.

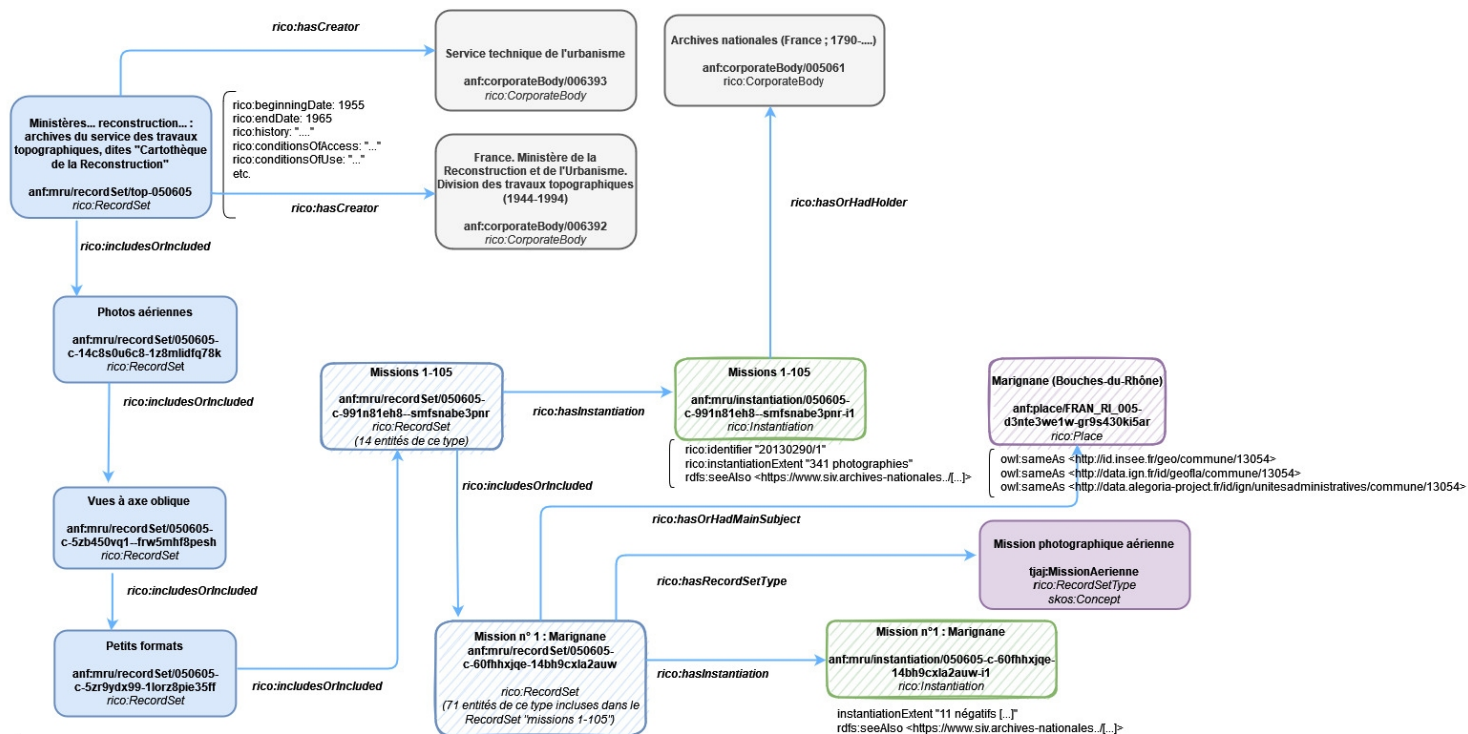
Des explications plus précises figurent ici : <https://www.ica.org/standards/RiC/ontology.html#fromRiCCM-to-RiCO>.

Nous donnons ci-après quelques précisions et exemples.

Record resources and instantiations: the example of a deed in a cartulary



Record resources and instantiations : the example of a series of aerial photographs



ALEGORIA : métadonnées descriptives des photos de la cartothèque de la Reconstruction (alias cartothèque MRU)
Aperçu général du modèle de description, du fonds jusqu'au niveau des missions.

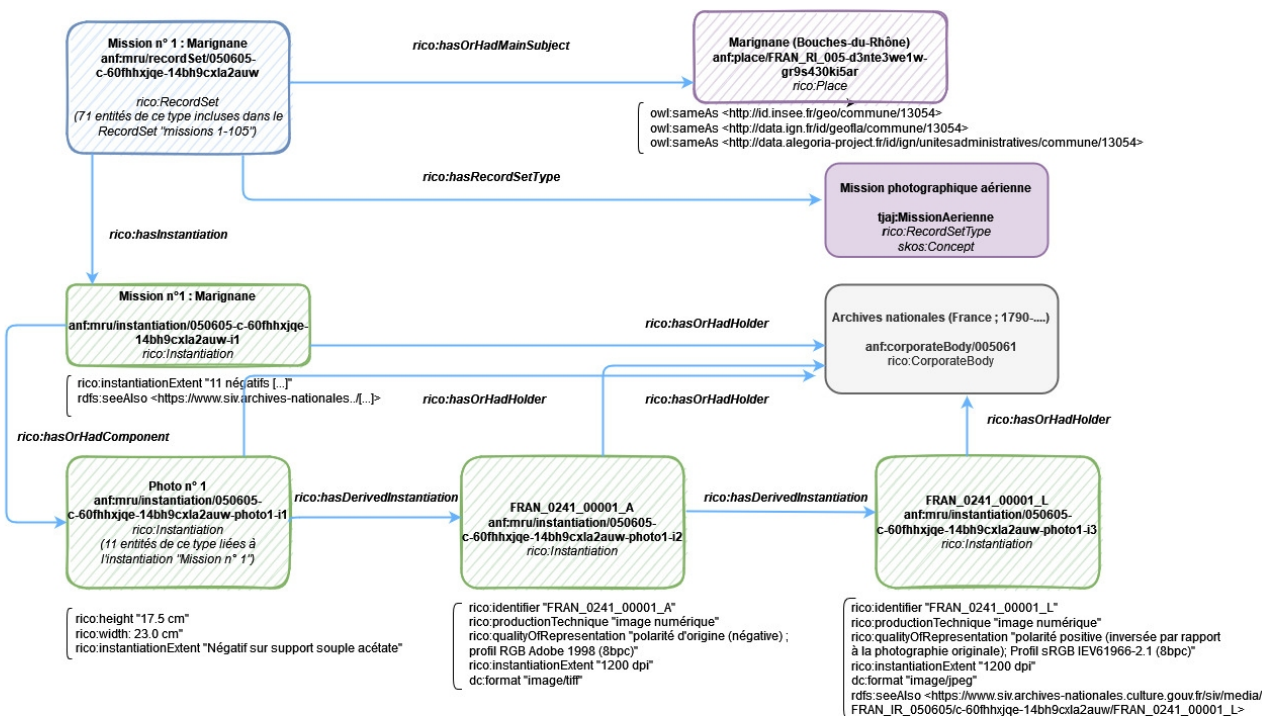
Les entités représentées sur fond rayé sont un exemple d'un nombre important d'entités de même nature.
Les entités représentées sur fond rose sont représentatives d'entités qui permettent de lier les données MRU avec celles de l'IGN
Au-dessous des rectangles, figurent quelques-unes des propriétés des entités ayant pour cible des valeurs textuelles, sauf pour owt: sameAs et rdfs: seeAlso (dont la cible est toujours un permalien vers la description classique de la ressource dans le SI des Archives nationales)

Les relations sont présentées uniquement dans un sens (les relations inverses existent mais ne figurent pas dans le diagramme).
@prefix tja: <http://data.alegoria-project.fr/id/codes/ign/geotheque/typejeuxdarchivedimagerie/>
@prefix rico: <https://www.ica.org/standards/RIC/ontology/>
@prefix anf: <http://data.alegoria-project.fr/id/anf/>
Lorsqu'il n'y a pas de préfixe, le composant existe dans RIC-CM et dans RIC-O.



© Archives nationales de France, décembre 2021
Document mis à disposition selon les termes de la Creative Commons Attribution 4.0 International License

Record ressources et instantiations : la description d'une photo dans la série de photographies



ALEGORIA : métadonnées descriptives des photos de la cartothèque de la Reconstruction (alias cartothèque MRU)
Aperçu général du modèle de description, pour les items (les photographies aériennes).

Le modèle ci-dessus illustre le cas le plus fréquent ; dans d'autres cas une mission consiste en un simple item (rico:Record) qui a alors trois instantiations. Certains Record Set de type mission aérienne peuvent aussi avoir pour sujet (rico:hasOrHadSubject) un concept (instance de rico:Thing et de skos:Concept) : ils concernent un objet de type gare, château, église...
Les entités représentées sur fond rose sont un exemple d'un nombre important d'entités de même nature.
Les entités représentées sur fond rose sont représentatives d'entités qui permettent de lier les données MRU avec celles de l'IGN
Au-dessous des rectangles, figurent quelques-unes des propriétés des entités ayant pour cible des valeurs textuelles, sauf pour owl:sameAs et rdls:seeAlso (dont la cible est toujours un permalien vers la description classique de la ressource dans le SI des Archives nationales)

Les relations sont présentées uniquement dans un sens (les relations inverses existent mais ne figurent pas dans le diagramme).
@prefix tjaj: <http://data.alegoria-project.fr/id/codes/ign/geotheque/type/pejoux/archivesdimagerie/>
@prefix rico: <https://www.ica.org/standards/RiC/ontology/>
@prefix anf: <http://data.alegoria-project.fr/id/anf/>



© Archives nationales de France, décembre 2021
Document mis à disposition selon les termes de la **Creative Commons Attribution 4.0 International License**

[Diagrammes tirés de la présentation faite le 16 décembre 2021 lors du séminaire de clôture du projet ANR ALEGORIA]

RiC-CM et RiC-O, des modèles flexibles

Il s'agit notamment de rendre RiC-O **immédiatement utilisable**. Par exemple :

- pour consigner **l'histoire d'une entité** (comme un Record), RiC-CM propose soit l'attribut RiC-A21 History, soit la relation RiC-R059i (*is or was affected by*) et une instance de l'entité Event. Dans RiC-O, c'est pareil, on a le choix.
- pour spécifier que **telle entité (comme une activité) est régulée par un texte réglementaire officiel**, RiC-CM propose la relation RiC-R063i (*is or was regulated by*) et une instance de l'entité Rule. Dans RiC-O, c'est pareil ; mais on trouve aussi la propriété *rico:ruleFollowed* qu'on peut utiliser pour stocker du texte.
- pour consigner **l'appartenance d'une entité à une catégorie**, RiC-CM propose des attributs, comme, pour les documents, RiC-A17 Documentary Form Type. Dans RiC-O, on peut faire le même choix et se contenter de stocker du texte comme objet de la propriété *rico:type*. Mais on peut aussi choisir d'utiliser *rico:hasOrHadCategory*+ une instance de la classe *rico>Type* (donc, par ex., un concept qui serait une instance de *rico:DocumentaryFormType*).
- Si on souhaite indiquer que **telle personne a dirigé tel groupe**, on trouvera dans RiC-CM la relation RiC-R042 *is or was leader of*. La même possibilité existe dans RiC-O, qui propose également de représenter cette relation sous la forme d'un graphe plus complexe, impliquant une instance de la classe *rico:LeadershipRelation*

Les 'types' dans RiC-O (1)

Dans RiC-CM, des attributs (comme Activity Type, Carrier Type, Corporate Body Type, Documentary Form Type, Event Type, Place Type, Record Set Type) **permettent de catégoriser les entités. Leur valeur devrait normalement être prise dans des vocabulaires contrôlés.**

Les concepts de tels vocabulaires sont des **points d'accès majeurs aux métadonnées**, et constituent potentiellement des **clés de liage essentielles** entre jeux de métadonnées archivistiques.

En RDF, le modèle de référence pour représenter un vocabulaire contrôlé est [SKOS](#). Il permet de définir des concepts et les relations sémantiques (hiérarchiques ou associatives) qu'ils ont entre eux, ainsi que des relations d'alignement avec des concepts d'autres vocabulaires.

Il existe déjà des vocabulaires SKOS nationaux et internationaux pour certaines de ces notions (et d'autres vont voir le jour, en France par exemple ; on le souhaite en tout cas). Voir par exemple le thésaurus matières pour l'indexation des archives locales.

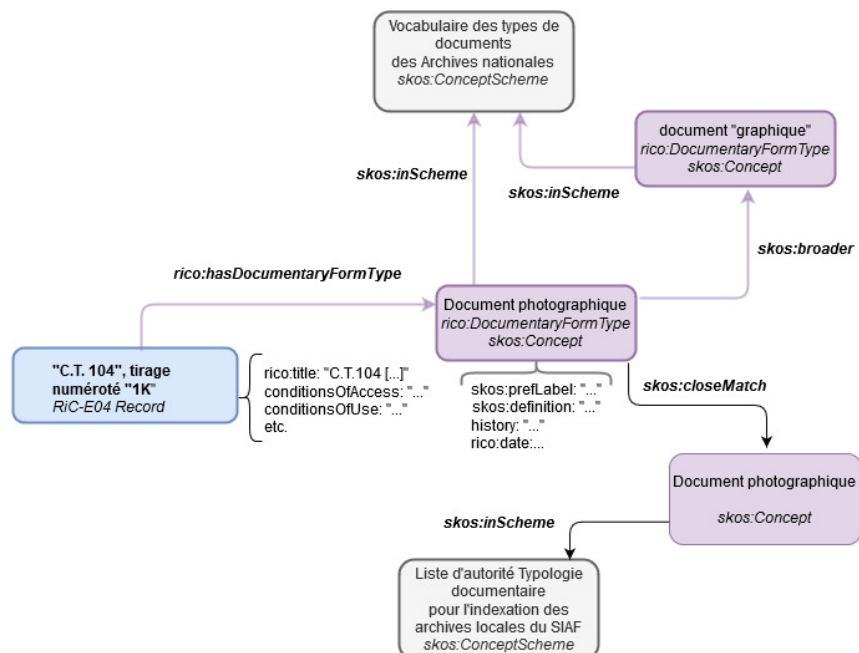
Les 'types' dans RiC-O (2)

Si l'on en fait une analyse scientifique fine, certaines notions au moins (ou les contenus de certains de ces vocabulaires) peuvent être considérées comme des entités historiques, ayant une histoire, voire des dates d'existence et une dimension spatiale, précédant ou suivant d'autres concepts. C'est le cas notamment pour `rico:DocumentaryFormType` ou `rico:ActivityType`. Le modèle SKOS n'inclut pas les propriétés nécessaires pour exprimer ces facettes contextuelles, historiques.

Afin de permettre, si et quand ce sera nécessaire, de décrire ces notions comme des entités historiques, **RiC-O définit donc ses propres classes, sous-classes de `rico:Type`, pour représenter les attributs de catégorisation de RiC-CM.**

Il est très simple d'articuler un vocabulaire SKOS existant avec une des classes RiC-O ainsi définie. Il suffit de spécifier que les concepts de ce vocabulaire sont aussi des instances de la classe RiC-O correspondante. RiC-O est accompagnée de l'embryon de deux vocabulaires, un pour les Record Set Type (types d'ensemble documentaire), un pour les Documentary Form Type (types de document).

Les 'types' dans RiC-O (3)



**Exemple d'utilisation d'un type de document (rico:DocumentaryFormType),
ici défini comme étant également un concept SKOS
dans le référentiel des types de documents des Archives nationales**

Les relations RiC-O sont présentées uniquement dans un sens (les relations inverses existent mais ne figurent pas dans le diagramme).

Le préfixe rico: indique que le composant (classe ou propriété) n'existe que dans RiC-O.

Lorsqu'il n'y a pas de préfixe, le composant existe dans RiC-CM et dans RiC-O.



© Archives nationales de France, février 2020
Document mis à disposition selon les termes
de la Creative Commons Attribution 4.0 International License

La modélisation des lieux dans RiC-O (1)

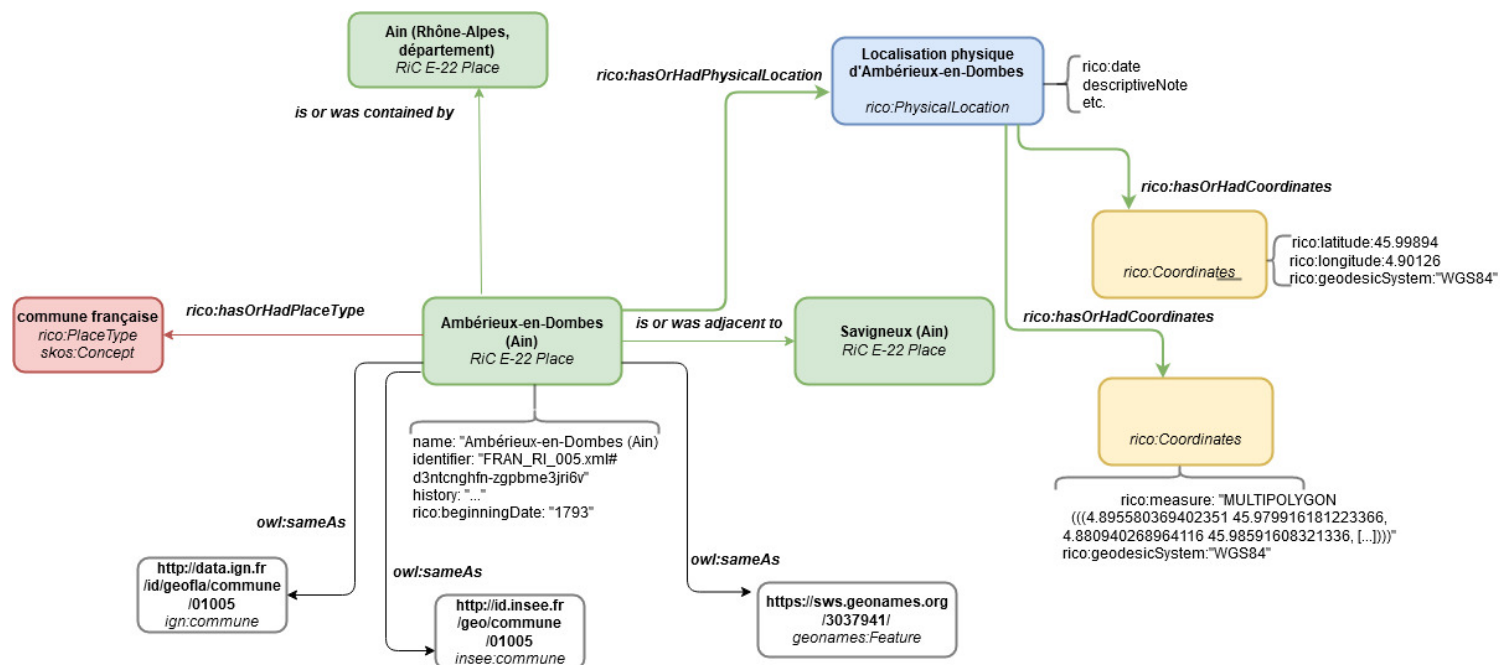
Les lieux dans RiC sont considérés, non pas comme des concepts, mais comme des **entités géo-historiques**.

RiC-O propose d'utiliser jusqu'à trois classes pour représenter un lieu, si besoin.

Ce modèle est assez proche d'autres modèles en cours d'élaboration, comme [Linked Places Format](#).

La modélisation des lieux dans RiC-O (2)

Exemple ici avec la commune d'Ambérieux-en-Dombes (dans le référentiel des lieux des Archives nationales).



Description partielle, conforme à RiC-O, d'un lieu (la commune d'Ambérieux-en-Dombes) dans la version RDF du référentiel des lieux des Archives nationales.

Les relations sont présentées uniquement dans un sens (les relations inverses existent mais ne figurent pas dans le diagramme).
Le préfixe rico: indique que le composant (classe ou propriété) n'existe que dans RiC-O.
Lorsqu'il n'y a pas de préfixe, le composant existe dans RiC-CM et dans RiC-O.

© Archives nationales de France, janvier 2021

Document mis à disposition selon les termes de la Creative Commons Attribution 4.0 International License



On aurait pu ajouter des relations vers les noms (instances de rico:PlaceName) de la commune, vers des lieux qui ont précédé cette commune, vers des lieux localisés dans cette commune... Et bien sûr, on peut ensuite spécifier, par exemple, des relations vers des agents dont le ressort ou le siège social est dans cette commune, ou des documents ayant cette commune pour sujet.

Les classes de relations dans RiC-O (1)

Une relation peut être considérée comme un objet à part entière, ayant par exemple une date et une description. C'est particulièrement vrai pour certaines relations dont la portée est importante en archivistique, comme les relations entre agents, ou celles entre agents et activités. Et dans de nombreux autres cas !

Exemple : « Jean-Noël Jeanneney (1942-....) a été président de Radio France et RFI de 1982 à 1986. »

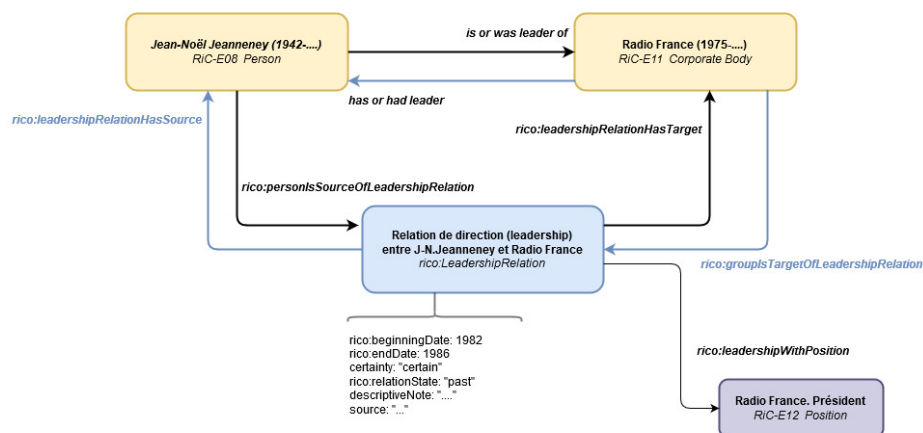
En EAC-CPF (ISAAR-CPF) (adapté de la notice d'autorité 050789 des Archives nationales) :

```
<cpfRelation xmlns:xlink="http://www.w3.org/1999/xlink" cpfRelationType="associative"
  xlink:href="FRAN_NP_005419" xlink:type="simple">
  <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>
```

Les classes de relations dans RiC-O (2)

Dans RiC-O, les 78 relations binaires définies dans RiC-CM sont présentes. Cependant, comme une propriété RDF (par ex. *rico:hasProvenance*, transposition de la relation RiC-R026 de RiC-CM) ne peut pas avoir d'attributs et qu'ici on a besoin notamment de décrire et de dater la relation, les relations sont aussi représentées comme des classes.

Cette méthode de modélisation permet aussi de connecter, à l'aide d'une telle relation, de nombreuses entités autres que la source et la cible d'une relation binaire. La relation binaire est alors par ailleurs formellement définie dans RiC-O comme un chemin simple, déclaré comme déductible, par inférence, de la combinaison de deux triplets dans lesquels une instance d'une classe Relation est centrale.



Représentation conforme à RiC-O d'une des relations de direction (*rico:LeadershipRelation*) spécifiée dans la notice EAC-CPF de J.-N. Jeanneney aux Archives nationales (https://www.siv.archives-nationales.culture.gouv.fr/siv/NP/FRAN_NP_050789). Cette relation relie trois entités (une personne, une collectivité et un poste).

Le préfixe *rico:* indique que le composant (classe ou propriété) n'existe que dans RiC-O. Lorsqu'il n'y a pas de préfixe, le composant existe dans RiC-CM et dans RiC-O.



© Archives nationales de France, septembre 2020
Document mis à disposition selon les termes
de la Creative Commons Attribution 4.0 International License

De la théorie à la pratique : mises en œuvre actuelles et futures de RiC-O

RiC-CM et RiC-O n'en sont qu'à leurs **débuts** et il y a peu d'implémentations de RiC-O.

Voir la page [RiC-O projects and tools](#) qui sera mise à jour prochainement.

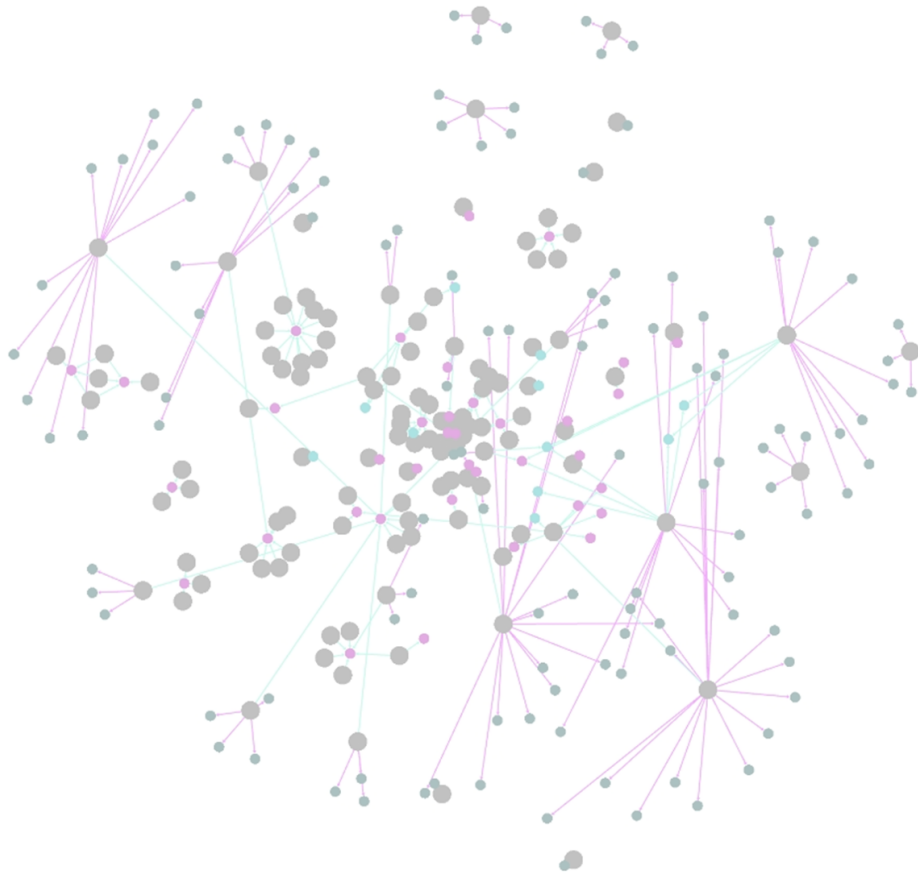
Parmi les utilisateurs présents ou futurs que le groupe EGAD connaît :

- le projet [Social Networks and Archival Context \(SNAC\) Cooperative](#)
- la société suisse [Docuteam](#), éditeur de logiciels, et plus généralement plusieurs projets en Suisse
- le service des archives de la ville d'Amsterdam, dont le futur système d'information sera fondé sur RiC-O et les technologies sémantiques
- le projet de recherche allemand [SONAR\(idh\)](#)
- à la suite d'une étude précise et indépendante qui a préconisé le choix de RiC-O, et après réalisation des évolutions nécessaires, le **portail FranceArchives**
- et les **Archives nationales de France**

Il est très probable qu'aucun projet ou institution n'utilise la totalité des composants RiC-CM ou a fortiori RiC-O.

Ils en utiliseront un sous-ensemble, et l'étendront aussi sans doute (ajouteront des sous-classes et sous-propriétés pour satisfaire des besoins spécifiques).

4. *Records in Contexts* aux Archives nationales



Cette partie du cours...

...s'appuie sur une présentation (fournie dans le dossier du cours) faite en octobre 2021 lors de la conférence internationale organisée par l'ICA (ICA Virtual Conference 2021, "Empowering knowledge societies") dont les enregistrements vidéo sont accessibles sur inscription seulement.

Voir aussi (c'est plus ancien) la présentation faite le 28 janvier 2020 aux Archives nationales sur ce sujet, lors de la journée d'étude sur « Les métadonnées archivistiques en transition : le nouveau cadre normatif, les enjeux et les premières réalisations ». Cette présentation est téléchargeable ici : https://f.hypotheses.org/wp-content/blogs.dir/2167/files/2020/02/20200128_3_RiCauxAN_EnjeuxPremieresRealisations.pdf.

Voir aussi l'article suivant : Clavaud (Florence), « Transformer les métadonnées des Archives nationales en graphe de données : enjeux et premières réalisations », dans *Les Archives nationales, une refondation pour le XXI^e siècle*, numéro spécial de La Gazette des Archives, n° 254 (2019-2), p. 59-88.

Ce qui sera fait cette année aux Archives nationales (en ce qui concerne les outils et infrastructures)

- la mise à jour du logiciel open source [RiC-O Converter](#) pour convertir l'ensemble des fichiers EAD et EAC-CPF des AN en données RDF/RiC-O 0.2
- le développement, dans le cadre d'un projet impliquant aussi le ministère de la Culture, la BnF et la société Sparna, d'une nouvelle version de [Sparnatural](#), un éditeur visuel de requêtes SPARQL ; la réalisation et la mise en ligne (officiellement, en mai 2022), d'un démonstrateur utilisant Sparnatural pour la recherche au sein des métadonnées relatives au tiers des archives notariales conservées aux Archives nationales
- la poursuite des travaux entrepris en 2022 avec l'Inria pour disposer à terme d'algorithmes de reconnaissance d'entités nommées (machine learning) dans les fichiers EAD (projet NER4Archives ; voir [la présentation faite en décembre 2021 lors de la conférence FF21](#))
- des spécifications et scénarios pour l'évolution du modèle de données et de l'infrastructure de gestion des référentiels
- des actions de conseil pour la réalisation de la version RiC-O du portail FranceArchives

Ce qui sera fait cette année aux Archives nationales (en ce qui concerne les données)

- la production et la publication, sous la forme de releases successives sur GitHub uniquement, des référentiels des Archives nationales, au format RDF/RiC-O 0.2 au moins, pour répondre à divers besoins de réutilisation (voir <https://github.com/ArchivesNationalesFR/Referentiels>)
- la production et la publication des jeux de données RDF/RiC-O 0.2 décrivant des séries de photographies aériennes des fonds Lapie et des ministères de la Reconstruction et de l'Urbanisme (dit fonds MRU), dans le cadre du projet [ALEGORIA](#) ; ce pour le [SPARQL endpoint du projet](#) et sous la forme de dumps sur GitHub. Voir à ce sujet [la page décrivant ces métadonnées](#)
- la production et la publication, dans le démonstrateur Sparnatural, de jeux de métadonnées décrivant le tiers des archives notariales conservées aux AN (voir diapo précédente)
- à partir des expériences ci-dessus, un travail d'amélioration de la qualité des métadonnées source (fichiers EAD) de description des archives notariales, et une réflexion plus large sur ce sujet
- la production et la publication de nouvelles données pour les référentiels, notamment ceux des agents, des lieux de Paris, des circonscriptions administratives françaises, des types de documents, des fonctions, dans le cadre de projets scientifiques divers ; la production d'une première version d'un référentiel des relations, basé sur l'existant aux AN, sur les besoins actuels et sur RiC
- une première mise en oeuvre de RiC-CM et RiC-O sur les métadonnées de description des archives nativement numériques
- dans le cadre de divers projets, des travaux de modélisation, de production de données et de liage avec d'autres jeux de données (cas du projet [ORESME](#))

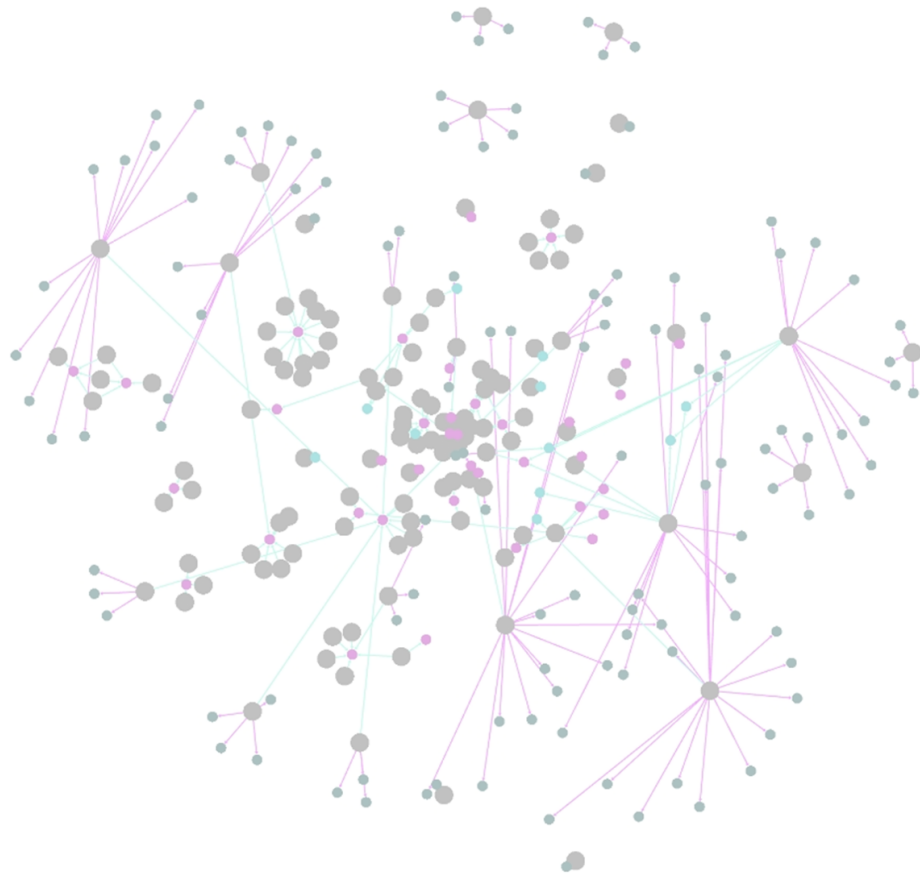
Tout cela, en attendant...

...une décision politique pour la mise en oeuvre des principes FAIR et la réalisation d'un data.archives-nationales.culture.gouv.fr (comme data.bnf.fr) permettant de passer à l'échelle, et la définition d'une stratégie, incluant :

- la mise en oeuvre de dispositifs d'assignation et de gestion d'identifiants pérennes pour l'ensemble des ressources à décrire ;
- l'intégration des dispositifs de production de données RDF dans le SI des AN ;
- la mise en place d'un SPARQL endpoint et le développement d'une interface complète de recherche et de consultation pour les données RDF ;
- le développement d'APIs dédiées permettant aux machines d'accéder à distance aux métadonnées des AN ;
- et plus globalement le découplage des silos internes de métadonnées, le changement des infrastructures de stockage et de gestion, pour passer à une architecture orientée données
- des outils et programmes pour améliorer la qualité des métadonnées et passer à un modèle de graphe d'entités fondé sur RiC

On espère que cette décision politique sera prise fin 2022 ou en 2023, que la stratégie sera élaborée courant 2023 et qu'elle sera suivie de l'affectation des moyens financiers et humains nécessaires pour réaliser ces importants changements.

5. Exercices



5.1. Exercice 1 : étudier un petit instrument de recherche en EAD, du point de vue de RiC

Ouvrir le fichier XML/EAD nommé FRAN_IR_054848.xml

- il est consultable dans la salle des inventaires virtuelle des AN ; son permalien est le suivant : https://www.siv.archives-nationales.culture.gouv.fr/siv/IR/FRAN_IR_054848.
- il est associé à la notice d'autorité (en EAC-CPF) FRAN_NP_005422.xml, dont la vue HTML est consultable via le permalien https://www.siv.archives-nationales.culture.gouv.fr/siv/NP/FRAN_NP_005422.

Quelles **entités** RiC y sont décrites ? Quelles propriétés ont ces entités ? Quelles **relations** existent entre ces entités ?

D'après cet unique exemple, quelles questions ou problématiques soulève le "passage à RiC" selon vous ?

Quelles réflexions ce petit exercice vous inspire-t-il ? Et si vous deviez, pour un service d'archives ou pour un projet fondé sur des métadonnées archivistiques, publier en RDF ces jeux de métadonnées, comment aborderiez-vous le sujet ?

5.2. Exercice 2 : utiliser le logiciel RiC-O Converter

Télécharger la dernière release de RiC-O Converter : <https://github.com/ArchivesNationalesFR/rico-converter/releases/tag/v1.0.0>.

Dézipper le dossier. On y trouve :

- dans le dossier 'documentation', un site web de documentation. Dans votre navigateur, ouvrir le fichier index.html qui est dans ce dossier. Vous y retrouvez notamment les mappings EAD et EAC vers RiC-O 0.1.
- dans le dossier unit-tests, les tests unitaires qui ont permis d'établir les spécifications détaillées du logiciel
- dans le dossier parameters, des fichiers de paramètres, adaptés au cas des AN, mais qui peuvent être modifiés
- dans le dossier input-ead, des fichiers EAD prêts à convertir
- idem dans le dossier input-eac

Dans un premier temps, exécuter la conversion des fichiers EAD (sur un ordinateur sous Windows, double-cliquer sur le fichier ricoconverter.bat et répondre 'convert_ead' à la première question, puis retour chariot à la question suivante).

La conversion prend quelques secondes, le résultat est stocké dans le dossier output-ead-{AAAAAMMJJ}.

Exercice 2, suite

Dans le dossier RDF de sortie, il y a un fichier RDF par instrument de recherche. Ouvrir le fichier `FRAN_RecordResource_054848.rdf` et le fichier EAD source (`FRAN_IR_054848.xml`) dans votre éditeur XML.

Des questions ?

Faire la même chose pour les 101 fichiers EAC-CPF stockés dans le dossier `input-eac`. Dans le dossier RDF de sortie, on trouve 3 sous-dossiers, dont un pour les agents (101 fichiers RDF, un pour chaque entité décrite) et un pour les relations entre agents et les relations de provenance. Ouvrir le fichier `FRAN_Agent_005422.rdf` et le fichier EAC-CPF source.

Des questions ?

Paramétrer la conversion

Pour cela on utilise les fichiers XML du dossier parameters

Ouvrir le fichier convert_ead.properties

On y trouve des paramètres pour la gestion des fichiers à traiter et des fichiers générés, et des paramètres qui vont définir certaines caractéristiques des fichiers de sortie, comme le segment invariant de l'URI.

Si l'on veut vraiment modifier le comportement de RiC-O Converter pour l'adapter à un cas d'utilisation d'EAD différent de celui des AN, il faudra modifier les scripts de transformation. Pour cela, il faut être un bon développeur XSLT 2.0, ou trouver un tel développeur !

Les AN vont faire vivre le logiciel :

- produire, tout d'abord, un 2e release, prenant en compte RiC-O 0.2 (courant 2022)
- continuer à enrichir le logiciel. Une des pistes serait que RiC-O Converter prenne en compte aussi des fichiers XML/SEDA (le SEDA, [Standard d'Echange de Données pour l'Archivage](#), schéma XML (transposé en norme AFNOR - NF Z440-22 « Modélisation des échanges de données pour l'archivage » - en janvier 2014, puis en mars 2018 en norme ISO - NF ISO 20614 « Protocole d'échange de données pour l'interopérabilité et la préservation ») qui définit le modèle de données des messages échangés lors du transfert d'un paquet d'archives numériques vers un système d'archivage électronique. Ces messages incluent impérativement les métadonnées de gestion et les métadonnées descriptives des archives transférées).

Elles seront heureuses de pouvoir étudier toute demande de modification, ou de travailler avec des ingénieurs d'autres institutions.