

Les référentiels aux Archives nationales

Florence Clavaud,
conservateur en chef du patrimoine,
responsable de la mission référentiels
Archives nationales, direction des fonds

*Rencontre annuelle des correspondants archives
des opérateurs et grands corps de l'État
12 octobre 2016*

Les vocabulaires

- Des vocabulaires (listes d'autorité ou thésaurus) élaborés pendant la phase de conception du SI, pour permettre **une indexation cohérente et contrôlée** :
 - des instruments de recherche
 - des notices de producteurs (vocabulaires lieux, activités des personnes, fonctions des collectivités, statuts juridiques des collectivités – qui mis à part le dernier sont aussi utilisables pour les instruments de recherche)En tout 17 référentiels.
- Depuis 2013, administrés, enrichis et publiés via un module spécifique du SI ; ils ne servent pour l'instant que lors de la recherche sur les ressources qu'ils indexent (ne sont pas consultables directement)
- En cours d'enrichissement et d'évolution (en particulier pour ce qui concerne les statuts, les activités et les fonctions)
- Globalement, sauf exceptions, ils n'ont pas encore la portée, la précision et la fréquence d'utilisation nécessaires pour devenir des outils puissants d'accès à l'information locale, ou des outils pour l'articulation de nos données avec les données d'autres institutions. Mais certains devraient l'être un jour.

Le référentiel des producteurs et les référentiels liés

Le référentiel des producteurs : enjeux (pourquoi ?)

- Pour mieux **connaître** :
 - le contexte de production des fonds et groupes de documents conservés aux Archives nationales ;
 - les fonds eux-mêmes !Pour consolider et valoriser une **expertise** en histoire des administrations et des institutions
- Pour aider à la construction d'une politique proactive de collecte
- Pour **faciliter l'accès** aux métadonnées de description des documents (dans la [salle des inventaires virtuelle](#))
- Pour disposer de données sur lesquelles fonder des dispositifs d'**interopérabilité** (l'interopérabilité se construira aux Archives nationales quasiment exclusivement sur les référentiels)

Le référentiel des producteurs : quelques éléments quantitatifs

- Environ 14100 notices validées et publiées
(sans compter, en permanence, environ 160 notices à valider, et 300 en cours de rédaction)
- Noyau dur du référentiel formé par des jeux de notices plus anciens :
 - 3426 notices sur les études et notaires de Paris du XV^e siècle à nos jours
 - **9700 notices environ issues de la base PRIAM du Centre des Archives contemporaines** (nomenclature créée pour la gestion des versements entrés aux AN entre 1970 et 2012)
- Environ 1000 nouvelles notices depuis le déploiement des modules documentaires du SI en 2013
- **Près des trois quarts des notices décrivent des collectivités**
- En fait le référentiel est très loin d'être une simple collection de notices (en nov. 2015, plus de 35000 relations entre notices et plus de 27000 relations avec des unités de description de documents au sein des instruments de recherche)

Par qui ?

- Rédacteurs :
 - l'ensemble des agents scientifiques des Archives nationales en charge de la collecte et du traitement des fonds
Soit une centaine de personnes potentiellement, essentiellement à la direction des fonds
 - les archivistes en poste dans les missions des archives en ministère (ou auprès des opérateurs et autres organismes de l'État) et au bureau des missions au SIAF)
Pour les services qui n'ont pas accès au SIA : **la rédaction se fait au traitement de texte ; il est recommandé de suivre les consignes générales de rédaction**
- Un chef de projet, responsable de la qualité de chaque notice, et de la cohérence et de la qualité globale du référentiel
- Depuis la mi-septembre 2016, des responsables scientifiques au sein des services (rôle intermédiaire entre le rédacteur et le chef de projet ; gestion « de proximité » de la qualité)

Quand (les « événements » déclencheurs) ?

- L'entrée d'un fonds ou d'un fragment de fonds :

- versement d'archives publiques
- entrée de fonds privés

Des événements très fréquents

Aucune entrée de fonds ne se fait tant que la notice descriptive du producteur n'existe pas

- Un chantier de description systématique de l'ensemble des producteurs

- inscrit au programme scientifique de l'établissement depuis 2013
- bases méthodologiques posées ou en cours de finalisation (cartographie prospective de l'existant, outils statistiques, manuel, expérimentation, outils simples de visualisation)
- un chantier de longue haleine, inédit, complexe, pour lequel des sous-chantiers sont en cours de définition dans les services de la DIRF

Comment ? Le format et l'outil

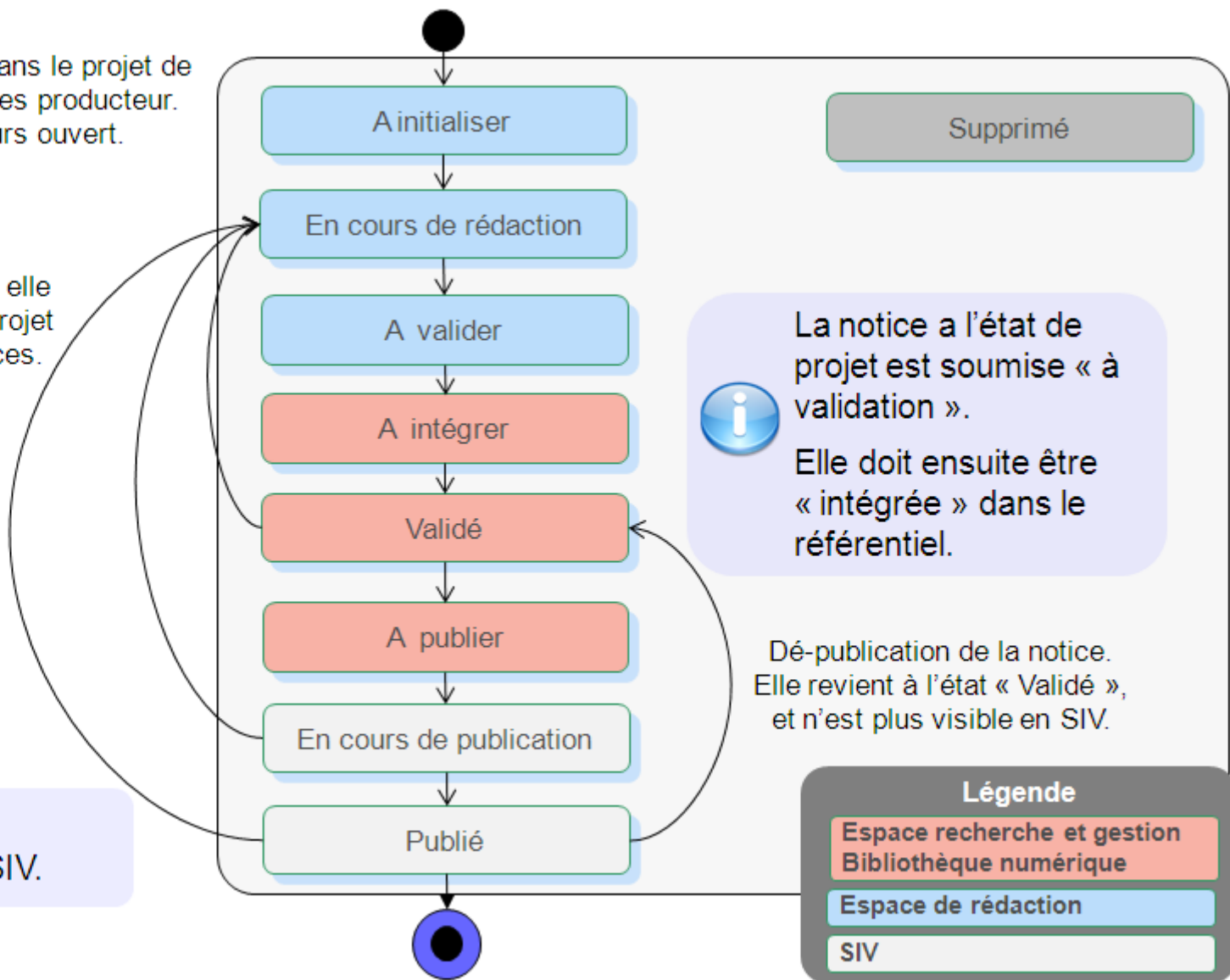
- **XML/EAC-CPF** : schéma publié en 2010, adopté en 2011 comme standard international par la Society of American Archivists
- Conforme à la norme conceptuelle **ISAAR(CPF)**, 2^e édition publiée par le CIA en 2004
- Au sein du SIA des Archives nationales, déployé en 2012-2013
- Stockage en XML/EAC-CPF, fichiers conformes au schéma EAC-CPF, implémentation légèrement restrictive
- Notices producteurs stockées dans l'entrepôt documentaire, utilisées en base de gestion (gestion des entrées et des articles) et publiées en front office
- Outil de rédaction fondé sur l'éditeur XML oXygen

Comment ? Le workflow

Les notices sont créées dans le projet de rédaction du référentiel des producteurs.
Le projet reste toujours ouvert.

Révision de la notice : elle est déplacée dans le projet de rédaction des notices.

La notice est « Publiée » en SIV.



Comment ? Principes

- Identifier :
 - recherches sur le producteur, son histoire, son statut, son organisation, ses fonctions et activités, son environnement administratif
 - perspective historique
- Principes de base :
 - **unicité** (un producteur, une notice) ;
 - **réalité** (pas d'abstraction : pas de collectivités fictives ou de notices de regroupement ; pas de notices pour des fonctions)
- Questions à traiter avant toute chose :
 - est-il pertinent de décrire l'entité ?
 - sur le temps long, a-t-on affaire à la même entité ?Décision sur la base de critères déjà éprouvés
 - > forte **granularité (profondeur, et précision temporelle)**
- Principes déjà mis à l'épreuve à l'occasion de l'élaboration et de la saisie de groupes de notices formant réseau

Comment ? Principes

- Décrire :
 - des informations ou éléments obligatoires :
Pour les collectivités et leurs subdivisions : forme autorisée du nom ; identifiant ISNI s'il existe ; dates d'existence ; statut juridique ; historique ; textes de référence ; indexation fonctions ; relations hiérarchiques, chronologiques, associatives, d'identité depuis peu ; sources utilisées pour rédiger la notice
 - des règles pour leur rédaction
 - des contrôles par l'outil, puis par le responsable scientifique ou le chef de projet
 - des mises à jour nécessaires
- Décrire exclusivement au sein du référentiel
 - donc remplacer les notices biographiques des producteurs qui seraient présentes dans les IR par des liens vers les notices producteurs du référentiel
 - idem pour le plan d'orientation général

Exemples de notices (via la salle des inventaires virtuelle)

- Une notice de collectivité :
 - notice 51396 : « France. Conseil supérieur de la navigation de plaisance et des sports nautiques (1967-....) »
https://www.siv.archives-nationales.culture.gouv.fr/siv/NP/FRAN_NP_051396
- Une notice de personne :
 - notice 51392 : « Biasini, Émile (1922-2011) »
https://www.siv.archives-nationales.culture.gouv.fr/siv/NP/FRAN_NP_051392
- Une notice de famille :
 - notice 50577 : « Murat (famille) »
https://www.siv.archives-nationales.culture.gouv.fr/siv/NP/FRAN_NP_050577

Le manuel pour la rédaction des notices

- Première version diffusée en interne aux AN en décembre 2015
Incomplète : contient essentiellement des recommandations pour la rédaction des notices, zone par zone et champ par champ, étayées par un argumentaire et assorties d'exemples
- Une 2^e version sera diffusée avant la fin de l'année pour commentaires, d'abord au sein des AN, puis dans le réseau des services liés aux AN
Contenu :
 - introduction (enjeux ; contenu du manuel ; conventions ; standards, normes et réalisations de référence)
 - chapitre sur les principes
 - chapitre sur la rédaction des notices (mis à jour et complété)
 - chapitre méthodologique
- Une fois validé ce manuel sera diffusé à tous les rédacteurs

Perspectives 2017-2020

- Avancer sur le chantier systématique
- Ouvrir le référentiel (interopérabilité) :
 - Parvenir à des échanges manuels assistés (ponctuels, déclenchés par le rédacteur dans l'outil de rédaction) ou automatiques (par batch) de données, dans les deux sens si pertinent, entre le référentiel des producteurs et :
 - la base de données internationale des identifiants ISNI ;
 - ONOMA, le futur référentiel acteurs du ministère de la culture (livraison fin 2016) ;
 - les notices d'autorité de la BnF (ou le possible futur référentiel national BnF-ABES)
 - tout autre référentiel chevauchant celui des AN ; pourquoi pas des référentiels des ministères ou des opérateurs nationaux
- Mieux l'exploiter, mieux le donner à voir (vers une interface représentant mieux ce que nos métadonnées sont : un réseau de données, un graphe)

Avancer sur le chantier de description systématique des producteurs

- Comment ?

- en définissant et réalisant des opérations ciblées et cohérentes d'enrichissement systématique

- producteurs de statut privé : avant tout, finir le traitement du « passif »

- administrations de l'Etat : s'appuyer sur tous les autres programmes du PSCE ; par ailleurs, décrire les entités du plus haut niveau de l'État et si possible de leurs subdivisions de premier niveau (cabinets, directions opérationnelles ou de support...), en considérant la longue durée (jusqu'à nos jours) et le contexte, et en reliant les notices entre elles et avec celles des personnes qui ont joué un rôle éminent au sein de ces entités.

Quelques exemples de chantiers déjà décidés

> dans ce travail, les responsables des archives des opérateurs nationaux et ministères pourraient apporter leur connaissance de la situation contemporaine ! Appel à volontaires !

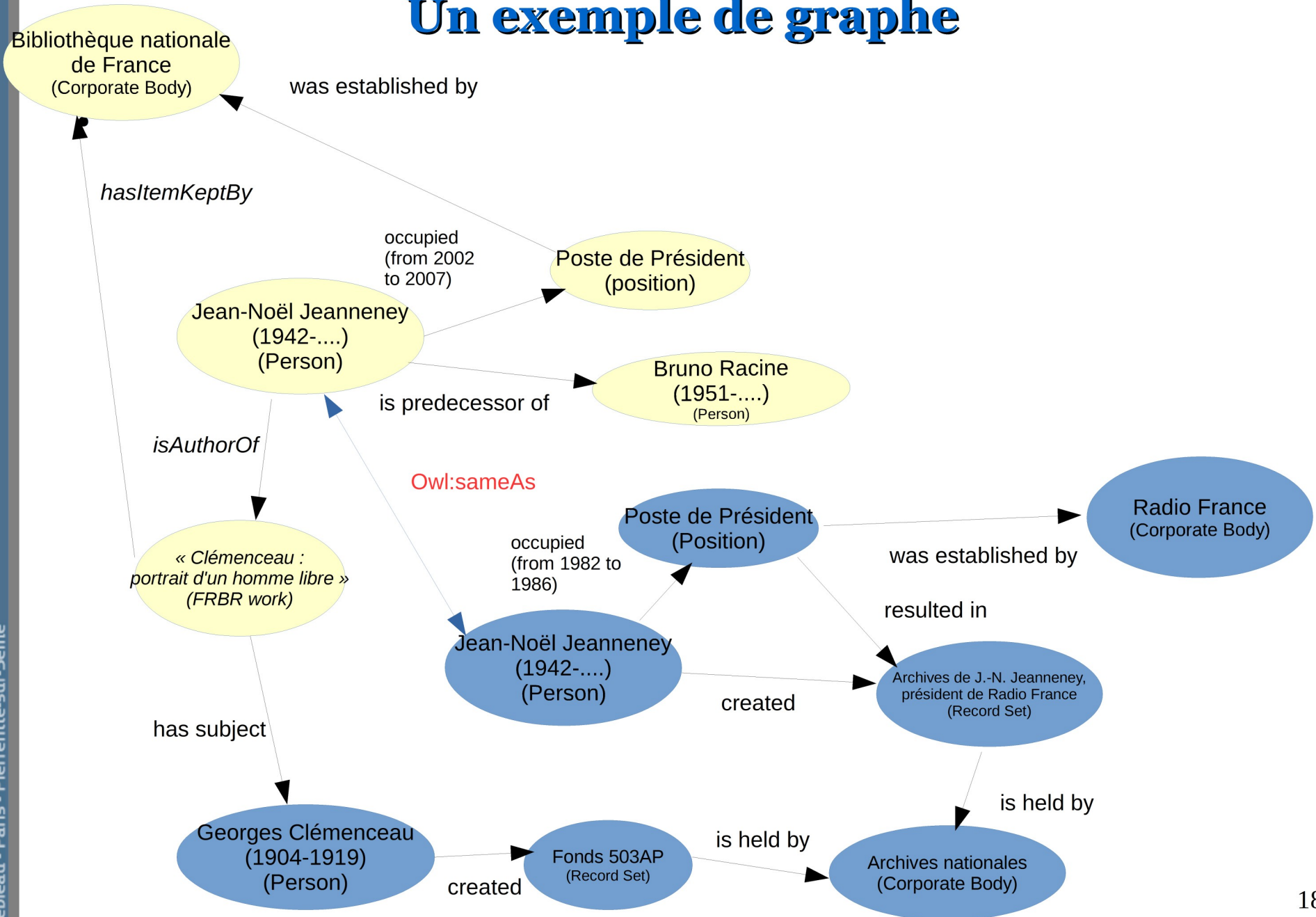
Consolider et valoriser l'expertise méthodologique acquise

- Comment ?
 - en mettant en ligne le manuel sur le site web des AN ;
 - en faisant évoluer ce manuel ;
 - en publiant les outils de représentation graphique réalisés ;
 - en participant à la formation initiale ou continue des archivistes ;
 - en organisant une journée d'études (fin 2017 ?) pour questionner et comparer les pratiques avec celles d'autres opérateurs – avant tout nationaux, tels que le NARA ou les AN du Luxembourg, et pour entendre les chercheurs en histoire de l'administration

Mieux exploiter, mieux donner à voir le référentiel des producteurs (et l'ensemble des métadonnées des AN)

- Comment ? en s'appuyant sur les technologies du web de données (Linked Open Data) ;
- Projet de prototype en cours de réalisation, pour la RDFisation, l'interconnexion et la visualisation en graphe de métadonnées archivistiques de provenances multiples :
 - projet né à l'été 2014, fin prévue : fin 2017
 - partenariat Archives nationales – BnF – SIAF (projet porté administrativement et piloté par les Archives nationales ; partenaire de la mission référentiels dans cette aventure : le DECAS)
 - preuve de concept et d'utilité (démonstrateur) : les partenaires ont choisi des métadonnées (notices d'autorité, IR, images numériques, vocabulaire des fonctions) et vont les convertir en triplets RDF ; un prestataire, qui est en cours de sélection, mettra en place une application web pour requêter ces triplets, construire par raisonnement de nouveaux triplets, proposer à l'utilisateur des dispositifs de visualisation analytique et en graphe
 - Architecture et bibliothèques open source, réutilisables dans les limites d'une solution non industrielle, déconnectée des SI des partenaires
 - Fin 2017 : présentation des résultats ; fin 2017 ou début 2018 : publication d'un bilan méthodologique et prospectif
- Ce projet prendra pour base de travail le nouveau modèle conceptuel (Records In Contexts-Conceptual Model) publié par l'ICA en septembre 2016 (voir <http://www.ica.org/fr/egad-ric>) et l'ontologie associée (Records In Contexts-Ontology, qui sera publiée fin 2016)
- Il s'agira ensuite de mettre en œuvre l'ontologie sur d'autres cas spécifiques, puis de généraliser, pour obtenir une ontologie pour les AN... et de faire évoluer le SIA pour qu'il expose (pour les machines) et présente (pour les humains) les métadonnées archivistiques sous cette nouvelle forme

Un exemple de graphe



Autres référentiels liés au référentiel des producteurs

- Les référentiels des activités (des personnes) et des fonctions (des collectivités)
 - déjà en cours d'enrichissement
 - devraient passer de structure plate à structure hiérarchique (thésaurus)
 - expérience en cours, d'élaboration de deux thésaurus pour certaines fonctions du domaine culture
 - à terme, ces référentiels devraient vraiment faciliter la recherche au sein des notices producteurs
- Le plan d'orientation général

Pour les services de l'État du Moyen Âge à nos jours, l'idée est de passer à un référentiel décrivant les fonctions en tant qu'entités contextuelles, ayant une histoire

> faciliter les recherches au sein du SIA en proposant le parcours suivant : du POG par fonctions aux notices producteurs et aux IR

Le vocabulaire des fonctions pourrait à terme servir à indexer ces notices descriptives de fonctions

N'hésitez pas à prendre contact (pour toute discussion ou question sur ces sujets)

- Contactez la mission référentiels des AN :
 - les demandes ponctuelles de création ou modification de notices producteurs sont à faire par courriel à :
referentiels.archives-nationales@culture.gouv.fr
 - pour toute autre question, vous pouvez aussi me contacter à :
florence.clavaud@culture.gouv.fr

- **Merci de votre attention !**