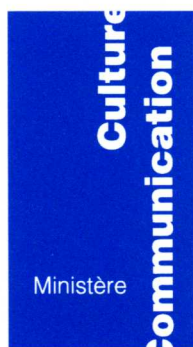


Paris, le 14 mars 2011



**Direction générale
des patrimoines**

Service interministériel des Archives de
France

Affaire suivie par : Michel Jacobson

Poste : 01 40 27 62 64

Référence : DGP/SIAF/2011/006

56, rue des Francs-Bourgeois
75141 Paris Cedex 03
France

Téléphone 01 40 27
Télécopie 01 40 27

Le directeur chargé des Archives de France

à

Mesdames les directrices des services des Archives nationales

Mesdames et Messieurs les directeurs d'archives départementales
sous couvert de Mesdames et Messieurs les présidents des conseils généraux

Mesdames et Messieurs les archivistes communaux
sous couvert de Mesdames et Messieurs les maires

Mesdames et Messieurs les archivistes régionaux
sous couvert de Mesdames et Messieurs les présidents des conseils régionaux

Note d'information DGP/SIAF/2011/006 relative à la mise en ligne d'une application de consultation du Thésaurus pour la description et l'indexation des archives locales anciennes, modernes et contemporaines.

Référence: Instruction DITN/RES/2009/006 relative à la publication du Thésaurus pour la description et l'indexation des archives locales anciennes, modernes et contemporaines

J'ai le plaisir de vous informer de la mise en ligne sur le site des archives de France à l'adresse <http://www.archivesdefrance.culture.gouv.fr/thesaurus> d'une

application de consultation du Thésaurus pour la description et l'indexation des archives locales anciennes, modernes et contemporaines.

L'instruction DITN/RES/2009/006 en date du 8 juin 2009 qui décrivait les principes ayant guidé la révision du thésaurus présentait également une première ébauche de standardisation du codage du thésaurus avec le langage SKOS¹. Depuis cette date, ce langage, défini au sein du W3C², a atteint le 18 août 2009 le niveau de « recommandation » ce qui représente un gage de grande stabilité et de large audience. Le codage du thésaurus a donc été modifié afin qu'il respecte ces recommandations.

Tirant parti à la fois des nouvelles possibilités d'expression de ce langage et de la nouvelle architecture du Web (appelé le « Web sémantique » ou plus récemment le « Web de données »³), les Archives de France ont établi un alignement du thésaurus avec deux autres sources de données que sont les vocabulaires Rameau⁴ et Dbpedia⁵.

Parallèlement à ce travail d'alignement, les Archives de France ont fait développer une application de publication de ce thésaurus afin d'en faciliter l'accès et l'utilisation tant par les archivistes, les internautes que directement par les éditeurs de logiciels documentaires d'archives. Cette application illustre les principes de la nouvelle architecture du Web de données.

- Chaque concept du thésaurus s'est vu attribuer un identifiant unique sous forme d'une URI⁶ (par exemple, pour le concept d'AGRICULTURE, il s'agit de l'url <http://www.archivesdefrance.culture.gouv.fr/thesaurus/ressource/T1-246>). En attendant la mise en place d'un système d'identifiants pérennes unifié à une échelle plus grande, nous avons choisi de baser ceux du thésaurus sur le nom de domaine archivesdefrance.
- Chaque concept est directement accessible par le web par le biais de son identifiant. Suivant le contexte de la demande (opérateur humain ou machine), la réponse sera formatée sous forme d'une page web en HTML (pour une personne) ou d'un flux XML en RDF (pour une machine). Un opérateur humain peut également demander un formatage XML/RDF concernant un identifiant ou même récupérer l'ensemble du thésaurus dans cette forme afin de pouvoir le recharger dans une autre application.
- Le thésaurus peut également être interrogé par l'intermédiaire d'un service web. En effet le langage SKOS et plus généralement RDF⁷ (sur lequel celui-ci est basé) peut être interrogé avec le protocole et langage de requête SPARQL⁸

1 SKOS: « Simple Knowledge Organization System » est un modèle de représentation des connaissances pour des vocabulaires simples de type thésaurus, taxonomies, etc. Recommandation du W3C <http://www.w3.org/TR/skos-reference/>

2 W3C: Le World Wide Web Consortium est un organisme à but non-lucratif, fondé en octobre 1994 afin de développer des standards tels que HTML, HTTP, XML ou encore RDF permettant d'assurer une croissance à long terme du Web.

3 Pour plus d'explication sur le « web de données »:
<http://www.archivesdefrance.culture.gouv.fr/thesaurus/web-de-donnees.html>

4 RAMEAU: Répertoire d'autorité-matière encyclopédique et alphabétique unifié

5 DBPEDIA: catégories utilisées dans l'encyclopédie collaborative Wikipédia

6 Une URI (Uniform Resource Identifier, soit littéralement identifiant uniforme de ressource), est une courte chaîne de caractères identifiant une ressource sur un réseau (par exemple une ressource Web).

7 RDF: « Resource Description Framework » est le langage de base du Web sémantique. Recommandation du W3C <http://www.w3.org/TR/rdf-syntax-grammar/>

8 SPARQL: acronyme récursif pour « SPARQL Protocol and RDF Query Language » est un langage de requête pour des infractions. Recommandation du W3C <http://www.w3.org/TR/rdf-sparql-query/>

également standardisé par le W3C. Vous trouverez sur les pages web du thésaurus un lien « SPARQL endpoint » qui pointe sur ce service web et sur des exemples de requêtes.

L'application en ligne vous permettra donc à la fois de consulter et de naviguer dans le thésaurus en suivant les liens hypertextuels, de pointer sur ces pages directement à partir d'identifiants que vous placerez dans vos applications ou pages statiques. L'application peut également être utilisée pour récupérer le contenu du thésaurus dans une forme qui en facilite la réutilisation mais l'utilisation du service Web doit permettre également aux applications d'utiliser le thésaurus sans pour autant avoir besoin de le récupérer dès lors que ces applications fonctionnent en mode connecté au réseau Internet. Enfin l'application comporte une partie documentation donnant des explications tant sur le thésaurus que sur les technologies mises en œuvre.

Parallèlement, les Archives de France ont procédé à une enquête auprès des services d'archives territoriaux, pour avoir un premier retour sur les difficultés rencontrées dans l'application du thésaurus. Les données contenues dans l'application seront réactualisées en fonction des demandes d'évolution adressées par les services territoriaux ayant répondu à l'enquête.

Le directeur, chargé des Archives de France



Hervé LEMOINE