

Bilan du concours Sémanticpédia 2014

La Délégation générale à la langue française et aux langues de France (DGLFLF) et l'association Wikimedia France ont organisé aux mois de mars et avril 2014 un concours intitulé « Sémanticpédia » dont l'objectif était de proposer des projets visant à réutiliser de manière innovante les données de Wikipédia.

Ces données sont extraites de la version française d'une base de données appelée DBpédia, DBpédia en français. Il s'agit d'un dispositif d'extraction de données culturelles à partir des ressources offertes en libre accès par l'encyclopédie Wikipédia francophone. La finalité des projets du concours Sémanticpédia est d'utiliser ce dispositif pour créer des projets culturels et offrir ainsi aux musées, aux bibliothèques et plus généralement aux opérateurs culturels, des perspectives inédites de diffusion et de partage de leurs ressources en ligne auprès du grand public.



Délégation générale à la langue française et aux langues de France



Concours sémanticpédia

Réutiliser et innover autour des données de Wikipédia

Dans le cadre de l'édition 2014 du concours Sémanticpédia, douze projets ont été soumis au jury composé de cinq membres :

- > **Thibault GROUAS**, chef de la mission *Langues et numérique* à la Délégation générale à la langue française et aux langues de France (DGLFLF)
- > **Jean-Séverin LAIR**, sous-directeur des systèmes d'information (SDSI) au ministère de la Culture et de la Communication
- > **Fabien GANDON**, directeur de l'équipe de recherche en web sémantique WIMMICS à l'établissement public de recherche Inria
- > **Camille DOMANGE**, chef du département des programmes numériques au sein du secrétariat général du ministère de la Culture et de la Communication
- > **Rémi MATHIS**, président de l'association Wikimedia France



Délégation
générale à la
langue française
et aux langues
de France

6 rue des Pyramides
75001 Paris
Téléphone : 01 40 15 73 00
Télécopie : 01 40 15 36 76
Courriel : dglflf@culture.gouv.fr
Internet : www.dglf.culture.gouv.fr
ISSN en ligne : 2260-2887

Ces douze projets correspondent à douze possibilités différentes de réutiliser de manière innovante les données, notamment culturelles, extraites de la base de données DBpédia en français, et de les diffuser.

Ci-contre figure la liste de ces douze projets éligibles et examinés par le jury, accompagnés du nom de leur porteur et d'un résumé explicatif.



1. Humanpédia

David Guez

Humanpedia est un projet qui propose à chacun de participer à la constitution d'une mémoire universelle hors tout contexte technologique en partageant oralement les grandes œuvres littéraires du patrimoine mondial et, dans une version plus généraliste, toute la connaissance représentée par l'encyclopédie en ligne « Wikipédia ».

Prototype visible sur :

<http://www.guez.org/semanticpedia>

2. Projet de valorisation du patrimoine religieux

Peter Potrowl

Ce projet a pour but à la fois la mise en valeur du patrimoine religieux français sur Wikipédia et l'exploitation de ces données pour l'information du public. Il cherche à inciter le public à la fois à s'intéresser au patrimoine religieux français et à contribuer à Wikipédia, notamment à son illustration. Une fois le projet achevé, le public pourra consulter sur Wikipédia la liste des édifices religieux de son département et connaître, grâce à une application mobile géolocalisée, les édifices religieux situés à proximité. Ce projet présente des similitudes avec Wiki Loves Monuments, qu'il prétend compléter.

3. Reador.NET

Christophe Desclaux

Reador.NET propose une solution à destination des personnes abonnées à des flux. L'outil les aide à suivre rapidement et simplement des informations filtrées publiées sur internet et les réseaux sociaux. En effet, Reador.NET récupère les actualités et messages postés sur le web et effectue une analyse sémantique de chaque information. À travers cette reconnaissance automatique (à l'aide de DBpédia), l'outil est capable de comprendre le sens des mots présents dans chaque texte et ainsi de proposer aux utilisateurs un filtrage optimal des actualités, un gain de temps et une meilleure productivité !

Prototype visible sur :

<https://github.com/descl/zone>

4. Musées en France

Julien Homo et Gautier Poupeau
Antidot

Le site Web « Musées de France » est un assemblage (mashup), c'est-à-dire un site obtenu à partir du traitement, de la mise en relation et de la réconciliation de différentes sources. Toutes les données sont récupérées directement et automatiquement, puis prises en charge par des chaînes de traitement construites avec le produit Antidot Information Factory (AIF). Le site Web en lui-même est généré à partir de l'interrogation d'un moteur de recherche basé sur les données traitées précédemment et le produit Antidot.

Prototype visible sur :

<http://labs.antidot.net/museesdefrance/>

5. Plateforme sur l'architecture moderne et contemporaine

Enora Prioul, Nolwenn Rannou et
Renaud Sagot
Cité de l'Architecture et du Patrimoine

La Cité de l'architecture & du patrimoine a pour projet de concevoir une plate-forme collaborative en ligne sur l'architecture. Le projet, qui vise à créer un corpus de notices descriptives et d'articles analytiques sur l'histoire et la création architecturales et urbaines du XX^e et du XXI^e siècle, s'adresse à tous : enseignants et étudiants, spécialistes, amateurs avertis et publics curieux.

6. Pumgrana

Bruno Loton

Pumgrana a pour objectif de fournir aux utilisateurs une nouvelle expérience dans la navigation et la consultation des contenus et leurs données sur internet. Il s'agit d'associer les contenus en fonction de leur sémantique et leurs données communes, et ainsi fournir à l'utilisateur un réseau pertinent et cohérent à travers lequel il se déplacera et pourra consulter chaque élément. L'aspect essentiel de ce projet est de concevoir une organisation de tous ces contenus (articles, vidéos, images, etc.) et donc une organisation du web qui soit pertinente, et ainsi offrir une vision différente des cadres d'accès actuels.

Prototype visible sur : <http://163.5.84.222/>

7. Atlasmuseum

Virginie Pringuet
Association A-PACK

Atlasmuseum est une plateforme wiki de collaboration entre artistes, amateurs, professionnels de l'art, développeurs, historiens de l'art, professionnels de la documentation et de l'archivage, visant le développement de différents outils d'inventaire de l'art public, de documentation de l'art public, de visualisation de l'histoire de l'art public, et s'appuyant sur MediaWiki et son extension SemanticmediaWiki.

Prototype visible sur :

<http://www.publicartmuseum.net/>

8. OpenDataSoft

David Thoumas
OpenDataSoft

OpenDataSoft propose à ses clients une plateforme en ligne, clé en main, de publication, partage et réutilisation de données. OpenDataSoft souhaite offrir à ses clients la possibilité d'intégrer simplement à leur plateforme des données issues de Wikipédia leur permettant ainsi de bénéficier de la richesse inestimable de Wikipédia et d'enrichir la gamme de services développés sur leur plateforme.

Prototype visible sur :

<http://semanticpedia.opendatasoft.com/>

9. CROTOS

Benoît Deshayes

CROTOS est un moteur de recherche sur les œuvres d'art. Les données utilisées sont exclusivement issues de la version francophone de DBpédia. Les champs indexés et requêtables sont les titres, auteurs et lieux de conservation. L'interface a été présentée en 6 langues (français, anglais, allemand, russe, espagnol, russe) mais est aujourd'hui disponible en 33 langues. Quand les titres ou auteurs ont une traduction présente dans DBpédia, celle-ci est indiquée, sinon la version française est donnée. Le choix des langues a été établi sur les volumes de traductions disponibles.

Prototype visible sur :

<http://www.zone47.com/crotos/>

10. Muséoscope

Stéphane Laurière
Ubimix

Muséoscope.net est une carte des musées et des monuments historiques de France, enrichie de données issues de Wikipédia. Cette carte a été initiée par Ubimix dans le cadre du concours Sémantipédia. Son objectif est de proposer une nouvelle expérience de découverte et d'exploration de ce patrimoine, par sa représentation interactive sur le territoire. Deux types de données sont représentés sur la carte : les musées de France et les immeubles ou monuments historiques. La liste initiale des musées de France est celle publiée pour l'année 2012 sur le site Data.gouv.fr.

Prototype visible sur :

<http://museoscope.net/>

11. AKpédia

Catherine Lenoble et Stéphanie
Vilayphiou

AKpédia s'inscrit dans le projet Anna K, mythologies personnelles, une fiction développée à partir d'une collection de données autour de la vie et l'œuvre de la romancière Anna Kavan. Cette proposition vise le développement d'outils et méthodes complémentaires pour aider à la création ou à la complétion d'articles Wikipédia reposant sur l'exploration du web sur Anna Kavan ; l'extraction de ces savoirs et la production de métadonnées et de code wiki par des agents logiciels ; sur la production de pages Wikipédia, optimisée pour la réutilisation par DBpédia, à l'occasion de la mise en place et l'animation d'un *wiki sprint*.

12. Optimisation de campagnes publicitaires Facebook grâce aux données Dbpédia

Cédric Soares
OLi Digital

L'objectif de ce projet est d'améliorer les campagnes de publicité sur Facebook à partir du jeu de données DBpédia. DBpédia peut permettre d'étendre le champ sémantique des concepts et mots clés à acheter dans les campagnes marketing sur les réseaux sociaux, l'objectif étant d'augmenter la visibilité et l'efficacité des campagnes en réduisant leur coût.

Ces douze projets ont été présentés lors d'une soirée de remise des prix organisée par le ministère de la Culture et de la Communication et l'association Wikimedia France le 29 avril 2014 dans les locaux de Numa à Paris.

Trois projets ont été retenus et leurs porteurs se sont vus attribuer chacun un prix d'une valeur de 3 000 euros.

Les projets lauréats du concours Sémantipédia sont :

> dans la catégorie « création d'outils pour Wikipédia » :
ATLASMUSEUM

> dans la catégorie « valorisation de données culturelles et linguistiques » :
CROTOS

> dans la catégorie « réutilisation et diffusion de ressources Wikipédia » :
MUSÉES EN FRANCE

Ce concours s'inscrit également dans le cadre de l'événement **Silicon Valois** qui a eu lieu du 16 au 28 mai au ministère de la Culture et de la Communication. Cet événement a constitué un moment privilégié de rencontres entre des créateurs, artistes, entrepreneurs, et plus largement les acteurs numériques.

Pour en savoir plus, consulter le site du **Silicon Valois**.

Sitographie

DBpédia en français

fr.dbpedia.org

Version francophone de la base de données DBpédia.

Blog Sémanticpédia

semanticpedia.org

Plate-forme de collaboration entre le ministère de la Culture et de la Communication, l'Inria et Wikimédia France, permettant de développer un écosystème de données culturelles liées et publiquement accessibles.

JocondeLab

<http://jocondelab.iri-research.org/jocondelab/>

Ce site donne accès à plus de 300 000 notices illustrées, extraites de Joconde, en 14 langues, dont 4 régionales. Cette expérimentation permet d'explorer de nouvelles formes de navigation au sein des ressources culturelles en ligne, de découvrir l'indexation collaborative et de compléter les informations de Joconde avec des données de l'encyclopédie en ligne Wikipédia.

Muséosphère.net

museosphere.net

Ce site, consultable en 9 langues, permet de rechercher des musées à proximité, à partir de préférences personnelles. Il est réalisé dans le cadre de la plate-forme collaborative Sémanticpédia dont les données sont issues de la version francophone de DBpédia.

HDA-Lab

hda-lab.iri-research.org

Le Portail Histoire des arts a pour vocation de mettre à la disposition de publics prescripteurs une sélection de ressources culturelles et éducatives riches en lien avec le programme officiel de l'enseignement de l'Histoire des arts. Ce projet de recherche et développement a permis la réalisation d'une preuve de concept destinée à montrer le potentiel heuristique du balisage sémantique.

Bibliographie

Des données culturelles accessibles à tous :

Sémanticpédia, Repères 2012

[http://www.dglff.culture.gouv.fr/publications/](http://www.dglff.culture.gouv.fr/publications/Reperes12_Semanticpedia.pdf)

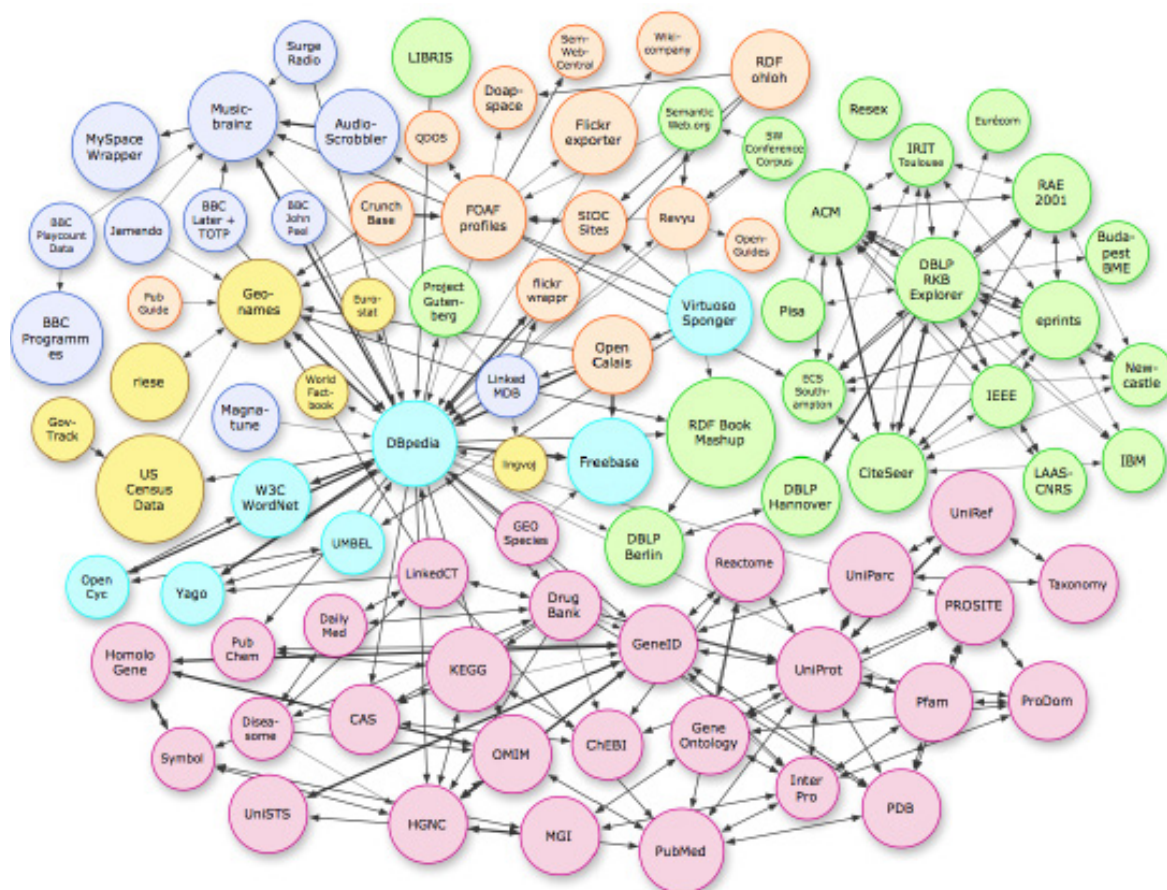
[Reperes12_Semanticpedia.pdf](http://www.dglff.culture.gouv.fr/publications/Reperes12_Semanticpedia.pdf)

Le numérique au service de la langue française

et des langues de France, Repères 2012

[http://www.culture.gouv.fr/culture/dglf/publications/](http://www.culture.gouv.fr/culture/dglf/publications/Reperes12_Numerique.pdf)

[Reperes12_Numerique.pdf](http://www.culture.gouv.fr/culture/dglf/publications/Reperes12_Numerique.pdf)



Représentation des relations entre DBpedia et divers autres projets de la toile