

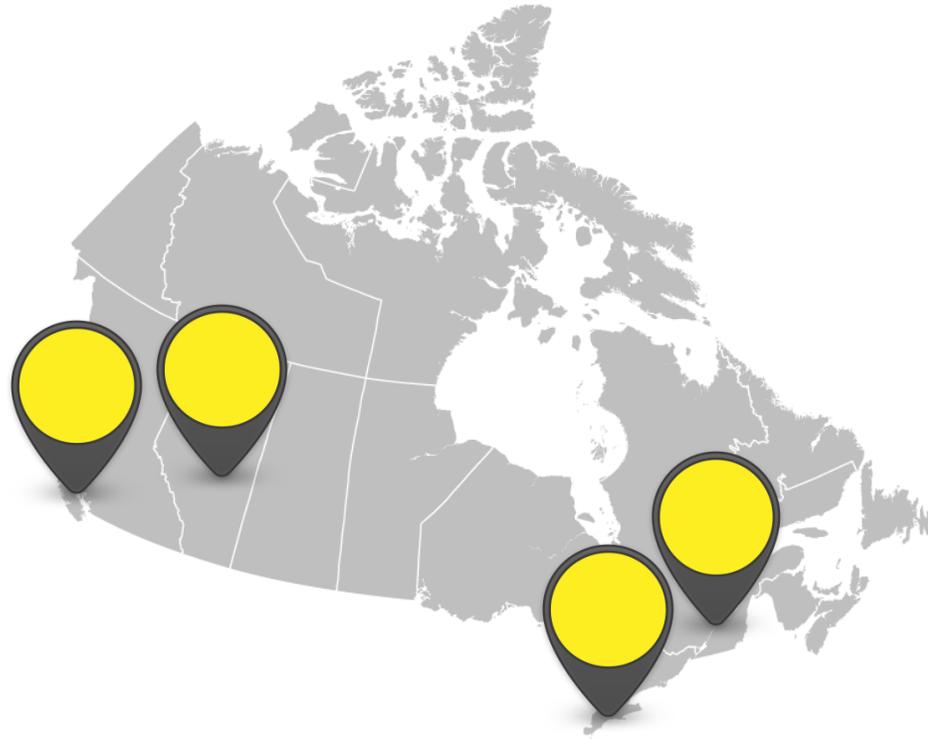


***Alignement d'ontologies reposant sur la recherche
d'information et l'analyse distributionnelle***

Eric Charton, Ph.D.

Relevance & Knowledge Lab Lead

Qu'est ce que la recherche locale ?



La recherche locale est l'activité sur Internet qui permet à des utilisateurs d'identifier dans leur proximité des entreprises de service ou de commerce.

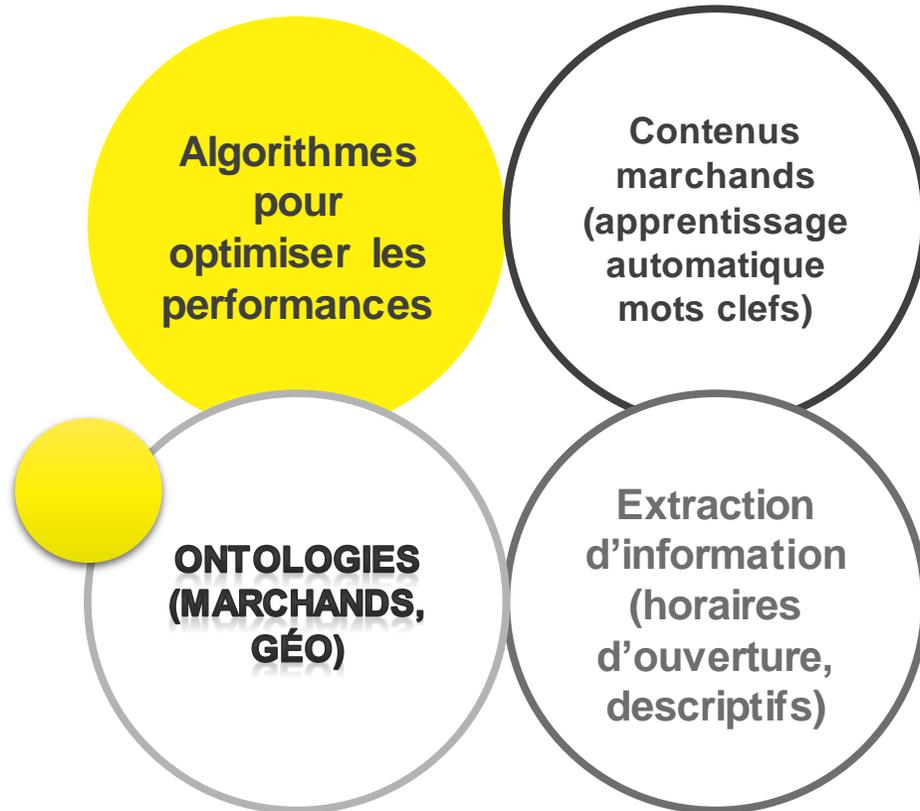
Exemples d'engins de recherche locaux

YP.ca YP.com, Pagesjaunes.fr, Yelp, Facebook for local businesses ...



500M visites

Topics : Restaurants / Home & services / legal ...



L'engin de recherche doit être supporté par des contenus de qualité (Knowledge) et des méthodes algorithmiques qui améliorent la précision et le rappel (Relevance)

Le Lab :

- Une quinzaine de collaborateurs
- Chercheurs, Linguistes, Développeurs, Analystes
- Travaillent sur :
 - Expansion de requêtes
 - Extraction d'information
 - Classification
 - Crawling
 - Ontologies
 - Analyse de données



Difficultés spécifiques d'un moteur de recherche local



Il ne dispose pas de documents à indexer

- Un profil marchand est l'équivalent d'un document.
- Une majorité de profils marchands provient des compagnies de télécommunications.

“Over one century, companies specialized in printed telephone directories have developed methods and techniques to collect and structure information about local businesses in the perspective of publishing those information through a very specific medium: printed directories.”



FILTRER PAR

- Emplacements ▾
- Ouvert
- Réservation en ligne
- Cuisine ▾

TRIER PAR

- Pertinence ▾

5 de 154 résultat(s) pour restaurant japonais (cote des neiges):

Évaluez les résultats »

Sushi Restaurant

5177, boul Décarie, Montréal QC H3W 3C2 [Itinéraire »](#)

Sushi et restaurants japonais, Restaurants

[Itinéraire](#) [Téléphone ▾](#)

Sushi Shop **OUVERT**

5640 Ch de la Côte des Neiges, Montréal QC H3T 1Z2 [Itinéraire »](#)

Sushi et restaurants japonais, Restaurants

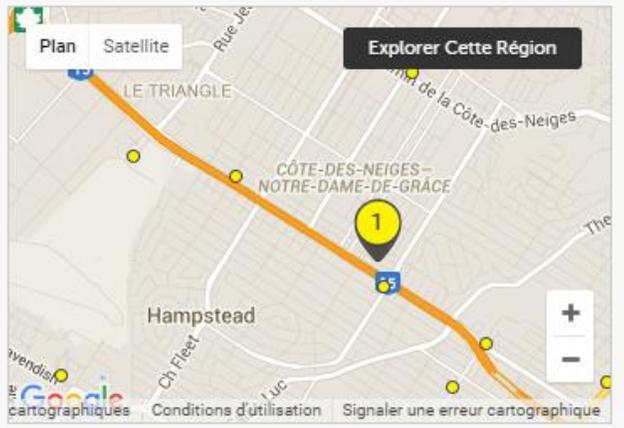
[Site web](#) [Itinéraire](#) [Téléphone ▾](#)

Sushi Shop **OUVERT**

5510, rue Jean-Talon O, Montréal QC H4P 2T5 [Itinéraire »](#)

Sushi et restaurants japonais, Restaurants

[Site web](#) [Itinéraire](#) [Téléphone ▾](#)



appli pj

Tout un quartier au bout des doigts

TÉLÉCHARGEZ L'APPLI !

Suggestions pour restaurant japonais à cote des neiges:



Difficultés spécifiques d'un moteur de recherche local



La recherche d'information avec données incomplètes ou insuffisantes est une tâche difficile.

Local businesses rarely offer exhaustive and easy to access content like websites or documentation. Usual techniques (crawl, information extraction) are sometimes inefficient. Historical system – telco data – cant be used alone.

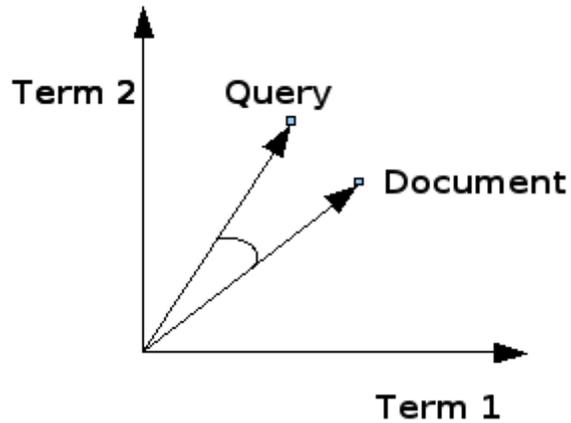


La solution : utiliser une recherche par ontologie.

Attribuer une classe à chaque marchand, et utiliser une ontologie pour fournir les mots clefs candidats



Recherche d'information classique : modèle vectoriel



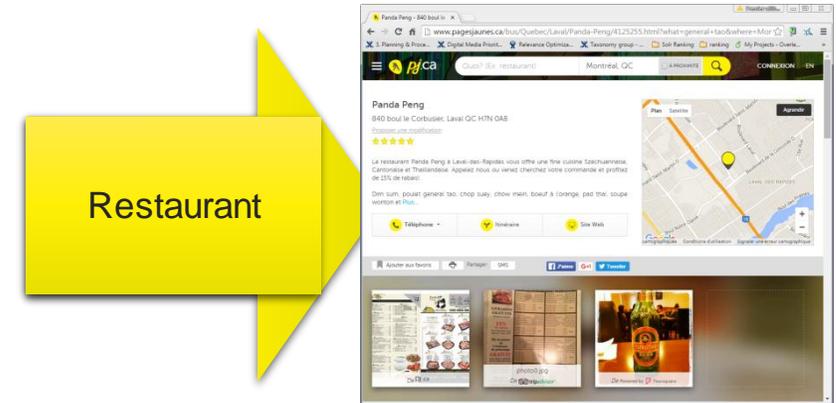
Recherche d'information dirigée par ontologie



L'ontologie définit la relation entre les mots clefs et les classes



Attribue les classes aux documents



Search





La recherche par ontologie ne s'applique pas à tout



Contraintes de la recherche par ontologie

- Le mot clef doit être généralisable
- Il doit pouvoir s'appliquer à un ensemble de documents, regroupés sous une classe unique.

Exemple : un engin de recherche dans lequel chaque document Représente un marchand ... 9

” Ednor” s’applique à un seul document décrivant un parc d’attraction



"Bak Kut Teh" s'applique à tous les documents décrivant un Restaurant Indonésien



Systeme proposé





Idée

Utiliser les contenus du web sémantique pour enrichir une ontologie propriétaire utilisée par un engin de recherche.

Ontology Search

Type	Yp Search	En Name	Fr Name	Code	AutoC	GSP	MOS	Sensiti	Rest	SEO
KW	0	bak kut teh		220095	false					false
KW	0		bak kut teh	225391	false					false

Associative Relationships for: bak kut teh

Name	Code	Relevance	Geo	Class	Status	Pcc	Comment	Filter	Weigf	GSP	MSP	LangSp	Mobil r	Sen
Chinese Food Restaurants	00283800	R	Yes	chinese restau	Approved ont	None	OMS-1431 - OHEX for YP-C	none	0.00	Yes	Yes	No	R	No

Utilisation du web sémantique pour enrichir une ontologie

Le web sémantique



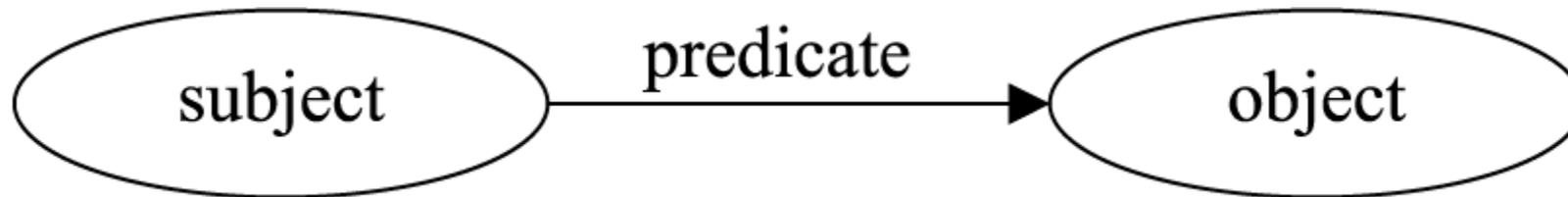
Le **web sémantique**, ou **toile sémantique**¹, est une extension du [web](#) standardisée par le [World Wide Web Consortium](#) (W3C)². Ces standards encouragent l'utilisation de formats de données et de protocoles d'échange normés sur le web, avec comme format de base le [Resource Description Framework](#) (RDF)
[Wikipedia]

Le web sémantique

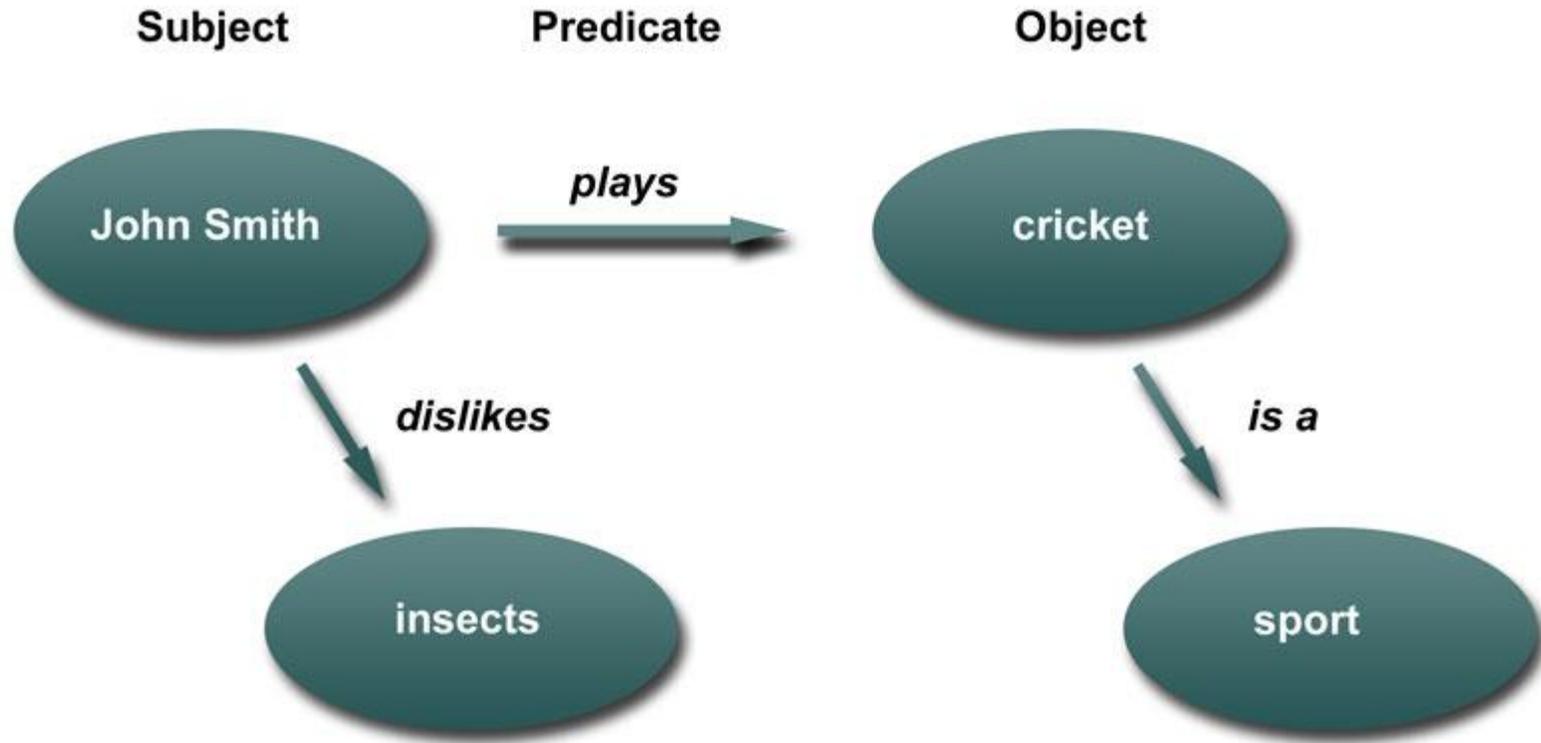


« le web sémantique fournit un modèle qui permet aux données d'être partagées et réutilisées entre plusieurs applications, entreprises et groupes d'utilisateurs »

Un principe simple : le prédicat



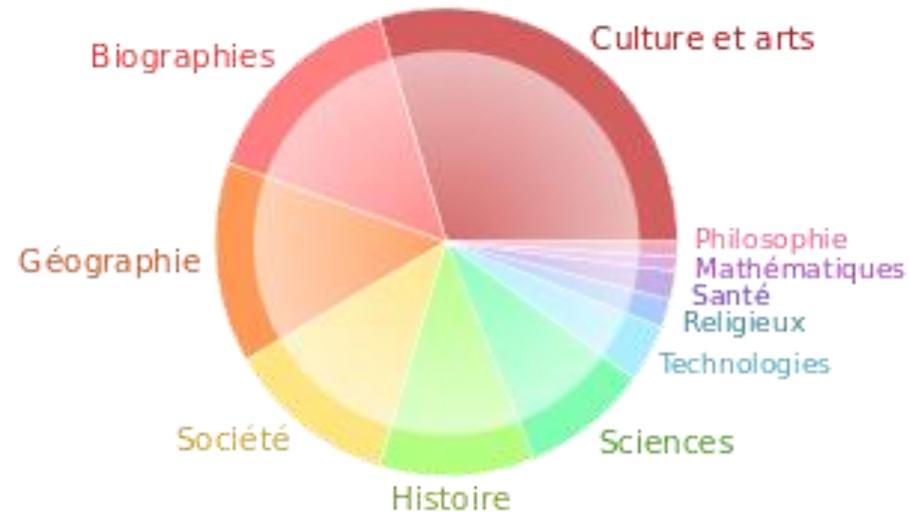
Le web sémantique



Le web sémantique



Wikipedia est une encyclopédie libre et exhaustive



DBpedia est la représentation de Wikipedia sur le web sémantique, sous forme de triples :

- 4.58 millions de 'choses' (noms de marques, de produits, d'objets)
- 1,445,000 personnes, 735,000 lieux
- 411,000 produits culturels (123,000 albums, 87,000 films et 19,000 jeux vidéo)
- 241,000 organisations
- 251,000 espèces et 6,000 maladies.



Premier utilis...
W Bak kut teh - Wikipedia, ti x

← → ↻ 🏠
https://en.wikipedia.org/wiki/Bak_kut_teh
☆ 🔍 🗨️ ☰

3. Planning & Proce...
Digital Media Priorit...
Relevance Optimiza...
Taxonomy group - ...
Solr Ranking
ranking
My Projects - Overle...
Merchant Content
OAR dev
»

Not logged in
Talk
Contributions
Create account
Log in

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes

Article **Talk**

Read **Edit** View history

Bak kut teh

From Wikipedia, the free encyclopedia

Bak-kut-teh (also spelt **bah-kut-teh**; Chinese: 肉骨茶; Peh-ōe-jī: *bah-kut-tê*) is a meat dish cooked in broth popularly served in Malaysia and Singapore, where there is a predominant Hoklo and Teochew community, and also in neighbouring areas like the Sumatra, Indonesia and Southern Thailand.

The name literally translates as "meat bone tea", and at its simplest, consists of meaty pork ribs simmered in a complex broth of herbs and spices (including star anise, cinnamon, cloves, *dang gui*, fennel seeds and garlic) for hours.^[1] Despite its name, there is in fact no tea in the dish itself; the name refers to a strong oolong Chinese tea which is usually served alongside the soup in the belief that it dilutes or dissolves the copious amount of fat consumed in this pork-laden dish.

However, additional ingredients may include offal, varieties of mushroom, choy sum, and pieces of dried tofu or fried tofu puffs. Additional Chinese herbs may include *yu zhu* (玉竹, rhizome of Solomon's Seal) and *ju zhi* (buckthorn fruit), which give the soup a sweeter, slightly stronger flavor. Light and dark soy sauce are also added to the soup during cooking, with varying amounts depending on the variant - the Teochews version is lighter than the Hokkiens'. The dish can be garnished with chopped coriander or green onions and a sprinkling of fried challets.

Bak kut teh

A closeup of bak kut teh

Place of origin	Malaysia and Singapore
Region or state	Hokkien- and Teochew-speaking areas of Malaysia, Singapore and Riau

Principe du système d'enrichissement par alignement



Principe du système

Objectifs :

- L'ontologie de PJ contient 3403 classes (Restaurants, Plombiers, Garages, Pizzeria ...).
 - *Projet :*
 - *Aligner ces 3403 classes avec les 503 000 catégories pertinentes de DBpedia*
 - *Résultats attendu :*
 - Collecter pour les 3403 classes de l'ontologie PJ toutes les relations pertinentes disponibles dans DBpedia.
 - Utiliser cette relation entre l'ontologie PJ et DBpedia pour collecter des mots clefs utilisables pour l'engin de recherche.



Principe du système

Wikipedia est utilisé comme ressource pour l'alignement pour les raisons suivantes :

- Il est possible de rechercher des documents dans Wikipedia en utilisant un moteur de recherche
- Wikipedia est en relation bijective avec DBpedia

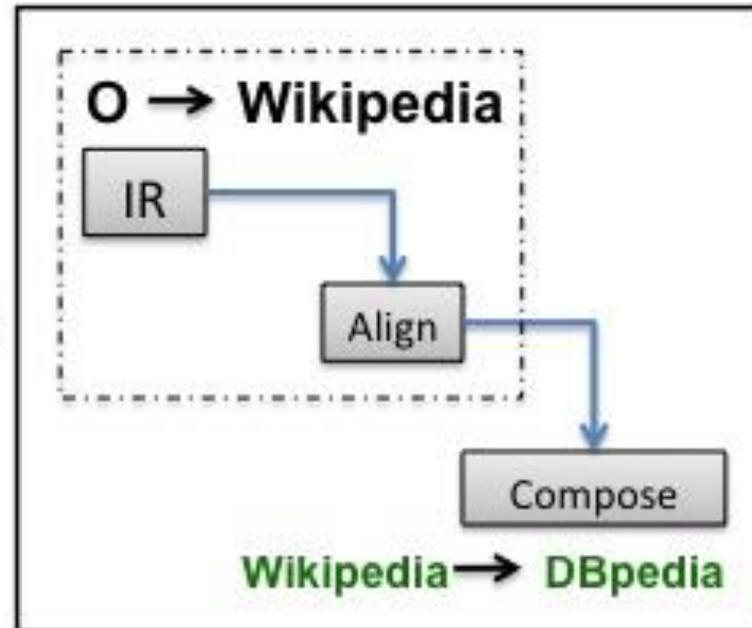
<http://www.w3.org/ns/prov#wasDerivedFrom>

▪ http://en.wikipedia.org/wiki/Category:Chinese_cuisine?oldid=546482961



Principe du système

Set of
Keywords
for a class



Alignment

O -> DBpedia



Principe du système : alignement d'ontologies

Le processus de recherche d'information utilisé pour obtenir les catégories candidates à l'alignement fonctionne comme suit :

1. Les mots clefs déjà associés à une catégorie de l'ontologie PJ sont utilisés pour chercher des pages Wikipédia.
2. Les pages Wikipédia retournées contiennent des catégories qui sont collectées.
3. Chaque catégorie possède son équivalent dans les catégories DBpedia
 1. Exemple :

La catégorie **Category:Restaurant** de Wikipedia existe dans DBpedia sous la forme **<http://dbpedia.org/ontology/Restaurant>** dans DBpedia
4. Les catégories DBpedia sont les candidates pour l'alignement



Principe du système : sélection des catégories

Les catégories DBpedia collectées sont les candidates pour l'alignement.

- Chaque catégorie de DBpedia est le sujet d'un ensemble d'entités
 - Exemple
http://dbpedia.org/page/Category:Chinese_cuisine
 - Est le sujet de toutes les entités de type 'plats' ou 'hors d'œuvre' en relation avec la cuisine chinoise

The screenshot shows a web browser window with the URL dbpedia.org/page/Category:Chinese_cuisine. The page title is "About: Chinese cuisine" and it is identified as an entity of type "Concept" from the Named Graph <http://dbpedia.org> within the Data Space dbpedia.org. The DBpedia logo is visible in the top right corner.

Property	Value
dbo:wikiPageID	698765 (xsd:integer)
dbo:wikiPageRevisionID	546482961 (xsd:integer)
rdfs:type	skos:Concept
rdfs:label	Chinese cuisine
owl:sameAs	<ul style="list-style-type: none"> http://el.dbpedia.org/resource/Κατηγορία:Κινεζική_κουζίνα http://pt.dbpedia.org/resource/Culinária_da_China http://fr.dbpedia.org/resource/Catégorie:Cuisine_chinoise http://cs.dbpedia.org/resource/Kategorie:Čínská_kuchyně http://de.dbpedia.org/resource/Kategorie:Chinesische_Küche http://es.dbpedia.org/resource/Categoría:Gastronomía_de_China http://eu.dbpedia.org/resource/Kategoria:Txinako_gastronomia http://it.dbpedia.org/resource/Categoria:Cucina_cinese http://ja.dbpedia.org/resource/Category:中国の食文化 http://ko.dbpedia.org/resource/분류:중국 요리 http://nl.dbpedia.org/resource/Categorie:Chinese_keuken http://pl.dbpedia.org/resource/Kategoria:Kuchnia_chińska http://wikidata.dbpedia.org/resource/Q7319372 http://wikidata.org/entity/Q7319372
skos:broader	<ul style="list-style-type: none"> dbc:Cuisine_by_nationality dbc:East_Asian_cuisine dbc:Food_and_drink_in_China
skos:prefLabel	Chinese cuisine
http://www.w3.org/ns/prov#wasDerivedFrom	http://en.wikipedia.org/wiki/Category:Chinese_cuisine?oldid=546482961
is dct:subject of	<ul style="list-style-type: none"> dbr:Soy_sauce dbr:Stir_frying dbr:Wheat_gluten_(food) dbr:Congee dbr:American_Chinese_cuisine dbr:Chinese_cuisine dbr:Japanese_Chinese_cuisine dbr:Egg_roll dbr:Spring_roll dbr:Changzhou_Big_Magao dbr:Tenshondon dbr:Deer_penis dbr:Dried_shredded_squid dbr:Balut_(food) dbr:Chinese_spoon dbr:Bon_bon_chicken dbr:Ciba_cake dbr:Kuai_(dish) dbr:Canadian_Chinese_cuisine dbr:Water_Banquet dbr:Caribbean_Chinese_cuisine dbr:Donkey_Burger dbr:Doupi



Principe du système : sélection des catégories

Toutes les entités en relation avec une catégorie DBpedia sont collectées et leurs propriétés sont utilisées en tant que *sac de mots* :

About: [Rouleau de printemps](#)

SameAs : [freebase:Rouleau de printemps](#)

Foaf.name : Spring roll

is [dbo:wikiPageRedirects](#) of :[dbr:Spring_rolls](#)

Etc ...

The screenshot shows a web browser window with the address bar displaying `dbpedia.org/page/Spring_roll`. The page title is "About: Rouleau de printemps" and it identifies the entity as "An Entity of Type : [Hors d'oeuvre](#), from Named Graph : <http://dbpedia.org>, within Data Space : [dbpedia.org](#)".

The main content area contains a paragraph describing the dish: "Le rouleau de printemps, appelé Goi Cuôn, Bo Cuôn ou Bi Cuôn selon la farce, est une spécialité culinaire du Vietnam. C'est également le nom d'un plat chinois, nommé 'chūnjuǎn' (春卷), très connu en occident. La garniture est enroulée dans une galette de riz crue. Il y a trois variétés vietnamiennes : La farce du Bi Cuôn se compose de viande de porc rôtie émincée, de filaments de couenne de porc, de concombre et de poudre de riz grillé. La farce du Bo Cuôn se compose de viande de bœuf émincée et de légumes. La farce du Goi Cuôn se compose de porc cuit dans un bouillon et émincé, de crevettes, de l'omelette, de la vermicelle de riz et de légumes. Il ne faut pas le confondre avec le Nem ran, ou Cha gio, spécialité du Vietnam, connu en France sous le nom de Nem, de pâté impérial ou rouleau impérial et dans les pays anglophones Spring roll, Summer Roll ou Vietnamese roll. Au Vietnam, le rouleau de printemps se mange en apéritif ou en entrée lors des repas. Dans certaines régions du pays, il peut également être préparé à base de poisson blanc de rivière frit, accompagné de feuilles de menthes et d'une fine tranche d'ananas. Ses autres appellations à travers le monde : Vietnamese Roll à Hong Kong, chūnjuǎn (春卷, littéralement rouleau de printemps) en Chine, ou Harumaki (春巻き, littéralement rouleau de printemps) au Japon sont des plats contenant principalement des légumes en julienne tels que le chou, la carotte, ou le champignon oreille de Judas. Certains types incorporent de la viande émincée, habituellement du bœuf ou du poulet; d'autres encore incorporent des fruits de mer. La garniture est enroulée dans une mince feuille faite à base de farine de blé [Laquelle ?]. Les rouleaux à la crevette (Shrimp rolls) que l'on trouve à Singapour sont de très petite taille, avec de la pâte de crevette à l'intérieur. Les petits rouleaux de printemps thaïlandais sont appelés Poh pia tod. Le rouleau de printemps est aussi populaire en Australie, où il est considéré comme un plat traditionnel asiatique occidentalisé, et il a son pendant australien appelé le Chiko Roll. (Voir aussi : Dim sim) Lumpia est le nom des rouleaux de printemps aux Philippines et en Indonésie. Aux Pays-Bas et en Belgique, les rouleaux de printemps sont connus sous le nom de Loempia (traduction phonétique), typiquement introduits par des immigrants venant d'Indonésie. À côté de ces rouleaux de printemps salés bien connus, il y a des rouleaux de printemps sucrés contenant de la pâte de haricot rouge ; on les trouve surtout dans le nord et l'est de la Chine (province du Zhejiang par exemple). Le rouleau de printemps à la chinoise était à l'origine une sorte de « snack » consommé à l'extérieur lors des visites au cimetière pour le nettoyage printanier des tombes. Les ingrédients étaient emballés dans une crêpe cuite d'avance dont la composition variait selon la région, et l'ensemble n'était pas remis à cuire. Le terme en est venu à englober toutes sortes de mets similaires, et de nos jours les rouleaux de printemps sont parfois frits. Dans la cuisine asiatique des États-Unis, on distingue le rouleau de printemps proprement dit, enrobé dans une fine crêpe de riz crue et l'egg roll emballé dans une crêpe plus épaisse, à base d'œuf, qui contient plus de garniture, et qui est frit.

Below the text is a table of properties and values:

Property	Value
<code>dbo:abstract</code>	<ul style="list-style-type: none"> Le rouleau de printemps, appelé Goi Cuôn, Bo Cuôn ou Bi Cuôn selon la farce, est une spécialité culinaire du Vietnam. C'est également le nom d'un plat chinois, nommé "chūnjuǎn" (春卷), très connu en occident. La garniture est enroulée dans une galette de riz crue. Il y a trois variétés vietnamiennes : La farce du Bi Cuôn se compose de viande de porc rôtie émincée, de filaments de couenne de porc, de concombre et de poudre de riz grillé. La farce du Bo Cuôn se compose de viande de bœuf émincée et de légumes. La farce du Goi Cuôn se compose de porc cuit dans un bouillon et émincé, de crevettes, de l'omelette, de la vermicelle de riz et de légumes. Il ne faut pas le confondre avec le Nem ran, ou Cha gio, spécialité du Vietnam, connu en France sous le nom de Nem, de pâté impérial ou rouleau impérial et dans les pays anglophones Spring roll, Summer Roll ou Vietnamese roll. Au Vietnam, le rouleau de printemps se mange en apéritif ou en entrée lors des repas. Dans certaines régions du pays, il peut également être préparé à base de poisson blanc de rivière frit, accompagné de feuilles de menthes et d'une fine tranche d'ananas. Ses autres appellations à travers le monde : Vietnamese Roll à Hong Kong, chūnjuǎn (春卷, littéralement rouleau de printemps) en Chine, ou Harumaki (春巻き, littéralement rouleau de printemps) au Japon sont des plats contenant principalement des légumes en julienne tels que le chou, la carotte, ou le champignon oreille de Judas. Certains types incorporent de la viande émincée, habituellement du bœuf ou du poulet; d'autres encore incorporent des fruits de mer. La garniture est enroulée dans une mince feuille faite à base de farine de blé [Laquelle ?]. Les rouleaux à la crevette (Shrimp rolls) que l'on trouve à Singapour sont de très petite taille, avec de la pâte de crevette à l'intérieur. Les petits rouleaux de printemps thaïlandais sont appelés Poh pia tod. Le rouleau de printemps est aussi populaire en Australie, où il est considéré comme un plat traditionnel asiatique occidentalisé, et il a son pendant australien appelé le Chiko Roll. (Voir aussi : Dim sim) Lumpia est le nom des rouleaux de printemps aux Philippines et en Indonésie. Aux Pays-Bas et en Belgique, les rouleaux de printemps sont connus sous le nom de Loempia (traduction phonétique), typiquement introduits par des immigrants venant d'Indonésie. À côté de ces rouleaux de printemps salés bien connus, il y a des rouleaux de printemps sucrés contenant de la pâte de haricot rouge ; on les trouve surtout dans le nord et l'est de la Chine (province du Zhejiang par exemple). Le rouleau de printemps à la chinoise était à l'origine une sorte de « snack » consommé à l'extérieur lors des visites au cimetière pour le nettoyage printanier des tombes. Les ingrédients étaient emballés dans une crêpe cuite d'avance dont la composition variait selon la région, et l'ensemble n'était pas remis à cuire. Le terme en est venu à englober toutes sortes de mets similaires, et de nos jours les rouleaux de printemps sont parfois frits. Dans la cuisine asiatique des États-Unis, on distingue le rouleau de printemps proprement dit, enrobé dans une fine crêpe de riz crue et l'egg roll emballé dans une crêpe plus épaisse, à base d'œuf, qui contient plus de garniture, et qui est frit. Spring rolls are a large variety of filled, rolled appetizers or Dim Sum found in East Asian and Southeast Asian cuisine. The name is a literal translation of the Chinese chūnjuǎn (春卷 'spring roll'). The kind of wrapper, fillings, and cooking technique used, as well as the name, vary considerably within this large area, depending on the region's culture.
<code>dbo:thumbnail</code>	http://commons.wikimedia.org/wiki/Special:FilePath/Spring_rolls_on_sale.jpg?width=300
<code>dbo:type</code>	dbr:Hors_d'oeuvre
<code>dbo:wikiPageID</code>	1139529 (xsd:integer)
<code>dbo:wikiPageRevisionID</code>	640625403 (xsd:integer)
<code>dbp:c</code>	春卷
<code>dbp:course</code>	dbr:Hors_d'oeuvre



Principe du système : sélection des catégories

La distribution des mots est comparée et les 20% de catégories de DBpedia dont les distributions sont les plus similaires aux distributions de mots de l'ontologie PJ sont conservées.

The screenshot shows a web browser window with the address bar displaying `dbpedia.org/page/Spring_roll`. The page title is "About: Rouleau de printemps". Below the title, it states "An Entity of Type : [Hors d'oeuvre](#), from Named Graph : <http://dbpedia.org>, within Data Space : [dbpedia.org](#)".

The main content area contains a paragraph describing the spring roll (Gỏi Cuốn, Bo Cuốn, Bi Cuốn) as a Vietnamese specialty. It mentions its Chinese equivalent, chūnjuǎn (春卷), and lists various fillings like pork, beef, and vegetables. It also notes regional variations and the use of rice paper wrappers.

Below the text is a table with two columns: "Property" and "Value".

Property	Value
<code>dbo:abstract</code>	<ul style="list-style-type: none"> Le rouleau de printemps, appelé Gỏi Cuốn, Bo Cuốn ou Bi Cuốn selon la farce, est une spécialité culinaire du Vietnam. C'est également le nom d'un plat chinois, nommé "chūnjuǎn" (春卷), très connu en occident. La garniture est enroulée dans une galette de riz crue. Il y a trois variétés vietnamiennes : La farce du Bi Cuốn se compose de viande de porc rôtie émincée, de filaments de couenne de porc, de concombre et de poudre de riz grillé. La farce du Bo Cuốn se compose de viande de bœuf émincée et de légumes. La farce du Gỏi Cuốn se compose de porc cuit dans un bouillon et émincé, de crevettes, de l'omelette, de la vermicelle de riz et de légumes. Il ne faut pas le confondre avec le Nem ran, ou Cha gio, spécialité du Vietnam, connu en France sous le nom de Nem, de pâté impérial ou rouleau impérial et dans les pays anglophones Spring roll, Summer Roll ou Vietnamese roll. Au Vietnam, le rouleau de printemps se mange en apéritif ou en entrée lors des repas. Dans certaines régions du pays, il peut également être préparé à base de poisson blanc de rivière frit, accompagné de feuilles de menthes et d'une fine tranche d'ananas. Ses autres appellations à travers le monde : Vietnamese Roll à Hong Kong, chūnjuǎn (春卷, littéralement rouleau de printemps) en Chine, ou Harumaki (春巻き, littéralement rouleau de printemps) au Japon sont des plats contenant principalement des légumes en julienne tels que le chou, la carotte, ou le champignon oreille de Judas. Certains types incorporent de la viande émincée, habituellement du bœuf ou du poulet; d'autres encore incorporent des fruits de mer. La garniture est enroulée dans une mince feuille faite à base de farine de blé [Laquelle ?]. Les rouleaux à la crevette (Shrimp rolls) que l'on trouve à Singapour sont de très petite taille, avec de la pâte de crevette à l'intérieur. Les petits rouleaux de printemps thaïlandais sont appelés Poh pia tod. Le rouleau de printemps est aussi populaire en Australie, où il est considéré comme un plat traditionnel asiatique occidentalisé, et il a son pendant australien appelé le Chiko Roll. (Voir aussi : Dim sim) Lumpia est le nom des rouleaux de printemps aux Philippines et en Indonésie. Aux Pays-Bas et en Belgique, les rouleaux de printemps sont connus sous le nom de Loempia (traduction phonétique), typiquement introduits par des immigrants venant d'Indonésie. À côté de ces rouleaux de printemps salés bien connus, il y a des rouleaux de printemps sucrés contenant de la pâte de haricot rouge ; on les trouve surtout dans le nord et l'est de la Chine (province du Zhejiang par exemple). Le rouleau de printemps à la chinoise était à l'origine une sorte de « snack » consommé à l'extérieur lors des visites au cimetière pour le nettoyage printanier des tombes. Les ingrédients étaient emballés dans une crêpe cuite d'avance dont la composition variait selon la région, et l'ensemble n'était pas remis à cuire. Le terme en est venu à englober toutes sortes de mets similaires, et de nos jours les rouleaux de printemps sont parfois frits. Dans la cuisine asiatique des États-Unis, on distingue le rouleau de printemps proprement dit, enrobé dans une fine crêpe de riz crue et l'egg roll emballé dans une crêpe plus épaisse, à base d'œuf, qui contient plus de garniture, et qui est frit. Spring rolls are a large variety of filled, rolled appetizers or Dim Sum found in East Asian and Southeast Asian cuisine. The name is a literal translation of the Chinese chūnjuǎn (春卷 'spring roll'). The kind of wrapper, fillings, and cooking technique used, as well as the name, vary considerably within this large area, depending on the region's culture.
<code>dbo:thumbnail</code>	<ul style="list-style-type: none"> http://commons.wikimedia.org/wiki/Special:FilePath/Spring_rolls_on_sale.jpg?width=300
<code>dbo:type</code>	<ul style="list-style-type: none"> <code>dbr:Hors_d'oeuvre</code>
<code>dbo:wikiPageID</code>	<ul style="list-style-type: none"> 1139529 (xsd:integer)
<code>dbo:wikiPageRevisionID</code>	<ul style="list-style-type: none"> 640625403 (xsd:integer)
<code>dbp:c</code>	<ul style="list-style-type: none"> 春卷
<code>dbp:course</code>	<ul style="list-style-type: none"> <code>dbr:Hors_d'oeuvre</code>



Principe du système : sélection des catégories

Align (τ, DBp)

0. Initialize $\kappa_{DBp} \leftarrow \emptyset$
 1. **for each class c in κ do**
 2. Initialize $DBp_c \leftarrow \emptyset$
 3. Find κ_c
 4. Initialize $C_{\kappa_c} \leftarrow \emptyset$
 5. **for each keyword k in κ_c do**
 6. Find ω_k
 7. **if $\omega_k \neq \emptyset$ then**
 8. $C_{\kappa_c} \leftarrow C_{\kappa_c} \cup \{\text{categories from } \omega_k\}$
 9. **end if.**
 10. **end for.**
 11. **for each category ζ in C_{κ_c} do**
 12. Set $kw_\zeta \leftarrow \text{getKeywordsFor}(\zeta)$
 13. **if isAlignable(ζ, kw_ζ) then**
 14. $DBp_c \leftarrow DBp_c \cup \{\text{DBpedia cat related to } \zeta\}$
 15. **end if.**
 16. **end for.**
 17. **end for.**
 18. Return κ_{DBp} .
-

Principe du système d'enrichissement par alignement



Expériences et évaluation

Les alignements sont validés manuellement par un expert du domaine.

Trois configurations expérimentales sont évaluées indépendamment :

1. **CLASSNAME**: seul le nom de la classe de l'ontologie (ex Restaurant) est utilisé dans le sac de mots.
2. **BOW – IDF**: Le sac de mots de l'ontologie est constitué de 50% de mots clefs avec le plus fort poids IDF.
3. **BOW** : Le sac de mots de l'ontologie est complet.

Configuration	Number of aligned classes	Coverage
<i>CLASSNAME</i>	694/2757	0.25
<i>BOW – IDF</i>	1916/2757	0.69
<i>BOW</i>	2485/2757	0.90



Merci

Eric.charton@yp.ca