

# Recherche d'Information - Introduction

Recherche et Extraction  
d'Information

Anne-Laure Ligozat/Xavier Tannier



# Qu'est-ce que la Recherche d'Information ?

Ensemble des méthodes et techniques pour l'acquisition, l'organisation, le stockage, la recherche et la sélection d'information pertinente pour un utilisateur

Index



Recherche Google

J'ai de la chance

**A**

- association 12, 405, 406, 410
  - OO 790
  - récurive 414
- association (H3M,T) 426
- associativité 170
- Astrahan M. M. 295, 296, 373
- Athlone M. P. 810, 827
- atomique
  - relations 353
- transactions 436, 441
- valeurs nulles 63, 104, 642
- attribut 89, 97
- authentification
  - voir mot de passe
- automatique locale 703
- automatisation
  - voir sécurité
- auxiliaire 427
- AVG - voir fonction d'agrégation
- axiome 906
- axiome déductif 325
- axiome de base 909, 931
- axiomes de Armstrong 320, 328

**B**

- B-trees 850
- Badal D. Z. 534
- Bancillon F. 823, 944, 945, 947
- Banerjee J. 803
- Barron G. M. 872
- Barnsley M. F. 884
- base de connaissances 819
- base de données 3, 10
  - avantages 16
  - DD2 890
- base de données déductive 910
- base de données distribuée 55, 695
- principe fondamental 69H
- base de données experte 840
- base de données extensible 923
- base de données intentionnelle 926
- base de données logique 840
- base de données relationnelle 110
- base de données statistique 508
- Batey D. S. 313, 874
- Bayer R. 478, 851, 815
- BCNF 337, 354, 361
- BDM 16
- Bockley D. A. 884
- iberech D. 325
- Boell C. 382, 392, 393, 946
- BEGIN DECLARE SECTION (SQL) 283
- BEGIN TRANSACTIONS 439
- Boil D. 739
- Bentley J. L. 884
- Bovounis P. A. 303, 419, 473, 534, 576, 729
- Biton (DD2) 884, 895
- Bitton D. 817
- Blomstedt A. 880

# Les acteurs de la Recherche d'Information

## Collection :

un ensemble de documents



## Utilisateur :

un besoin d'information et/ou une tâche à accomplir



**Système de RI** : l'outil qui doit retrouver les documents **pertinents** pour le besoin de l'utilisateur

# Recherche d'Information



Web Images Videos Wikipedia More ▶

librairie orsay

Search

Advanced Search

Home > Web results 1-10 of 119,601 for language:fr librairie orsay, Page 1 - Next page



## Librairie Orsay, Livre - Recherche adresse

Bonnes adresses > Essonne > Orsay > Culture, Loisirs et Voyages > **Librairie Librairie Orsay** Résultats 1 à 1 sur 1 pour **Librairie Orsay Librairie** Du Lycee Donnez votre avis 0% 57 Rue Paris - 91400

[www.justacote.com/orsay-91400/librairie](http://www.justacote.com/orsay-91400/librairie)

Cached - Bookmark



## Librairie Orsay - librairies

Culture high tech **librairie Orsay** - Toutes les infos sur Culture high tech **librairie Orsay** - Avis des internautes, téléphone, horaires, itinéraires, adresse et plan

[fr.nomao.com/orsay/faire-du-shopping/librairie.html](http://fr.nomao.com/orsay/faire-du-shopping/librairie.html)

Cached - Bookmark



## Achat - Vente **Librairie Orsay** - 91400, Cession **Librairie Orsay** - 91400

Des milliers d'annonces de **Librairie Orsay** - 91400 à vendre ou a ceder avec Vivastreet **Orsay** - 91400, trouvez votre **Librairie** parmi plus de 11 000 ANNONCES 100% GRATUITES Achat - Vente Cession

[fonds-commerce.vivastreet.fr/annonces-commerce-pas-de-porte-orsay-91400/q/librairie](http://fonds-commerce.vivastreet.fr/annonces-commerce-pas-de-porte-orsay-91400/q/librairie)

Cached - Bookmark

### Site type:

- » Blog
- » Forum

### Multimedia:

- » Video

### Filetype:

- » pdf
- » swf
- » text
- » word

### Related terms:

- » Art Contemporain
- » Art Moderne
- » Assemblee nationale
- » Centre Pompidou
- » Grand Palais
- » Musee d'Orsay
- » Musée National

# Recherche d'Information

- Où se trouve la librairie la plus proche de chez moi ?
- Qui est actuellement en tête du Top 14 de rugby ?
- Quels sont les titres mentionnés à la une du journal Le Monde d'aujourd'hui ?
- Que rapporte la une du Monde d'aujourd'hui sur la politique étrangère ?
- Quels sont les films qui passent ce soir sur la TNT ?
- Dans quels films Jean Rochefort et Philippe Noiret ont-ils joué ensemble ?
- Quels sont les logiciels d'installation de logiciels sous Linux/Debian ?
- Comment peut-on installer des logiciels sous Linux/Debian ?
- Quelle est la traduction du mot anglais "ice" en français ?
- Qui était Claude Bernard ?

## Questions

- Quelle est la nature des résultats attendus ?
- Comment évalue-t-on la pertinence des résultats ?
- Sous quelle forme doit-on formuler ses requêtes ?
- ...

# *Information vs. données*

"Les **données** sont reçues, stockées et retrouvées par un endosystème. Les données sont impersonnelles ; elles sont disponibles pour tout utilisateur du système.

L'**information**, en revanche, est un ensemble de données qui correspond à un besoin particulier.

Le concept d'information a des composantes personnelles et temporelles absentes du concept de donnée."

(R. R. Korfhage, 1997)

# Information

EGC 2016   Présentation▼   Conférence▼   e-EGC   Ateliers   Inscription   Informations Pratiques▼

 **EGC 2016  
REIMS**

**16ème Conférence Internationale  
Francophone sur l'Extraction  
et la Gestion des Connaissances**

**18 - 22 Janvier 2016**  
<http://egc2016.univ-reims.fr>



Dans le cadre des 50 ans de l'IUT de Reims, l'Université de Reims Champagne-Ardenne et l'IUT de Reims ont le plaisir d'accueillir EGC2016 sur le site de l'IUT du 18 au 22 janvier 2016.

La conférence Extraction et Gestion des Connaissances (EGC) est un événement annuel important réunissant des chercheurs, des industriels et des étudiants des communautés de l'apprentissage automatique, ingénierie et représentation des connaissances, statistique et analyse de données, fouille de données, systèmes d'information, bases de données afin de discuter des innovations et des résultats scientifiques les plus récents de ces domaines.

# Informations visibles

- 16<sup>èmes</sup> Conférence Francophone "Extraction et Gestion des Connaissances"
- Du 18 au 22 janvier 2016
- Reims, France
- Président d'honneur : Jim Hendler, Université Libre de Bruxelles, Belgium



EGC 2016 Présentation Conférence e-EGC Ateliers Inscription Informations Pratiques

EGC 2016 REIMS

16ème Conférence Internationale Francophone sur l'Extraction et la Gestion des Connaissances

18 - 22 Janvier 2016  
<http://egc2016.univ-reims.fr>

EGC

Institut Universitaire Technologique de Reims-Châlons-Charleville, France  
Centre de Recherches en STIC (CReSTIC EA3804), Université de Reims Champagne-Ardenne

Dans le cadre des 50 ans de l'UIT de Reims, l'Université de Reims Champagne-Ardenne et l'UIT de Reims ont le plaisir d'accueillir EGC2016 sur le site de l'UIT du 18 au 22 janvier 2016.

La conférence Extraction et Gestion des Connaissances (EGC) est un événement annuel important réunissant des chercheurs, des industriels et des étudiants des communautés de l'apprentissage automatique, ingénierie et représentation des connaissances, statistique et analyse de données, fouille de données, systèmes d'information, bases de données afin de discuter des innovations et des résultats scientifiques les plus récents de ces domaines.

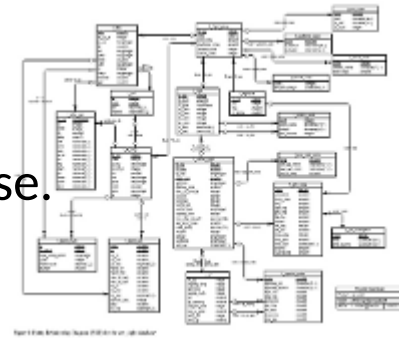


# *Ce qu'une machine voit...*

- 16<sup>èmes</sup> Conférence Francophone "Extraction et Gestion des Connaissances"
- Du 18 au 22 janvier 2016
- Reims, France
- Président d'honneur : Jim Hendler, Université Libre de Bruxelles, Belgium

# Diversité des besoins d'information (1/2)

- Recherche d'un **élément connu**
- L'utilisateur sait exactement quels éléments il recherche.
  - Il sait reconnaître les éléments désirés s'il les voit.
  - *Exemple* : recherche d'une citation bibliographique précise.
    - **Bases de données** (SQL, XQuery, etc.)
- Recherche d'une **information générale**
  - L'utilisateur recherche une information sur un sujet en général. Il existe de nombreuses façons de décrire le sujet.
  - Il est possible que l'information pertinente ne soit pas reconnue
  - Cette information peut ne satisfaire l'utilisateur
  - que de façon partielle.
  - *Exemple* : Les réformes de la recherche en France
    - **Recherche d'information « traditionnelle »**



# Diversité des besoins d'information (2/2)

- Recherche d'une **information précise**  
L'utilisateur recherche une information spécifique mais ignore sous quelle forme elle se présente.
  - Réponse partielle impossible
  - Exemple : À quelle date le président Kennedy a-t-il été assassiné ?

➤ **Extraction d'information** et **systèmes de question-réponse**
- **Exploration**  
Le but n'est pas de répondre à une question en particulier, mais de parcourir l'ensemble des données pour découvrir quels types d'informations concernant un sujet ou un domaine sont présents.
  - **Navigation**



# Diversité des sources d'information



- **Emplacement** des informations
  - Ressources locales ou distantes
  - Problèmes : disponibilité, identification, répartition sur plusieurs sources, variabilité des formats (encodage des caractères et description du contenu)
- **Nature** des fichiers de ressources
  - **Bases de données** : formats bien décrits, langages d'interrogation non ambigus (ex : SQL pour les bases de données relationnelles)
  - **Fichiers annotés** : formats plus ou moins bien décrits, annotations de présentation et/ou de description sémantique du contenu, langages d'interrogation (ex : XSLT/XPath pour les fichiers XML)
  - **Fichiers textes** : formats très peu ou pas décrits, langue(s) connue(s) ou non, plus ou moins de régularité entre les documents d'une même classe, pas d'interprétation générique possible (problème du Traitement Automatique des Langues)

# Diversité des problèmes

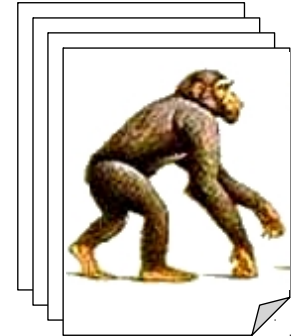
- Difficultés d'**accès, couverture, temps de traitement**
- Les bases documentaires sont **très grandes**, réparties sur de
  - **nombreux supports** dans des **endroits différents**
- Difficultés de définition de la **pertinence**
  - Comment un document remplit-il le **besoin informationnel** d'une personne donnée ?
  - Quelle est sa **pertinence** ? Comment la mesure-t-on ?
- Difficulté d'**exploitation**
  - Les documents pertinents ne sont pas nécessairement dans la **langue** de la requête.
  - L'information recherchée n'est pas nécessairement clairement identifiable dans un document.



# Grandes évolutions de la RI

- Précédemment :

- Bases documentaires structurées et de petite taille
- Accès par des métadonnées et rarement par le texte intégral
- Utilisation de langages documentaires (contraints) par les spécialistes



- Aujourd'hui

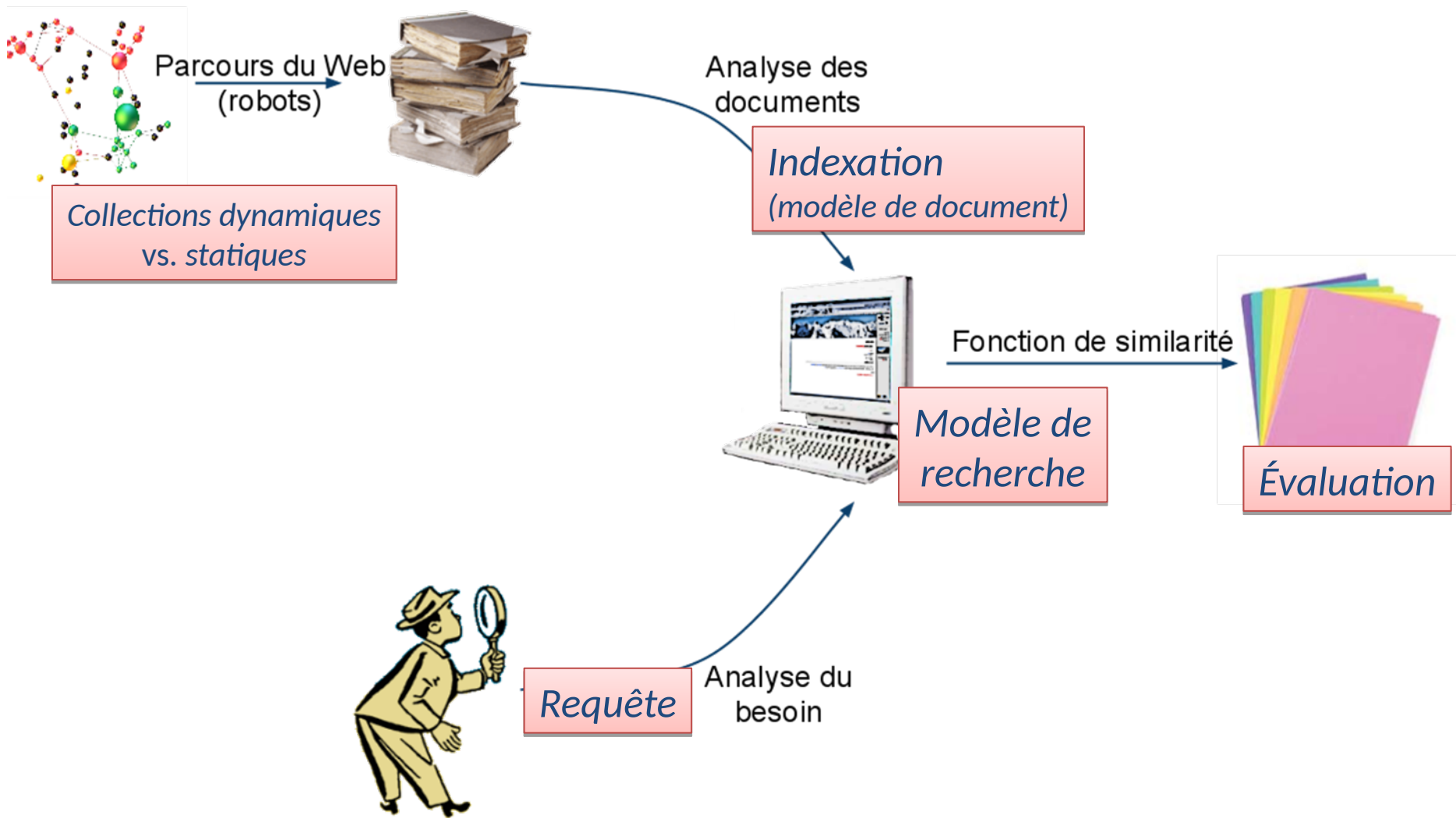
- Documents multimédia sous forme électronique
- Nombreux formats de représentation (texte brut, HTML, XML, PDF, RTF, formats propriétaires...)
- De plus en plus de données non structurées
- Une masse d'information gigantesque (Web...)
- Une composante sociale



# Recherche d'information sur le Web

- Sur Internet : utilisation massive par des **utilisateurs non experts**
  - Domaine d'une importance économique majeure
  - La requête typique est constituée d'au plus quelques mots clés
  - Les utilisateurs s'adaptent aux outils
- Une partie du web n'est pas directement **accessible** (web invisible, dont pages à accès restreint et pages dynamiques)
- L'information présente est fortement **multilingue** : les documents répondant aux requêtes peuvent être dans des langues différentes
- L'information présente n'est pas toujours **fiable**
- La **visualisation** de l'information est particulièrement importante : classement des résultats, présentation d'extraits, extraction de segments pertinents, etc.

# Recherche d'Information





**Les difficultés de la RI :**

**L'humain et le langage**

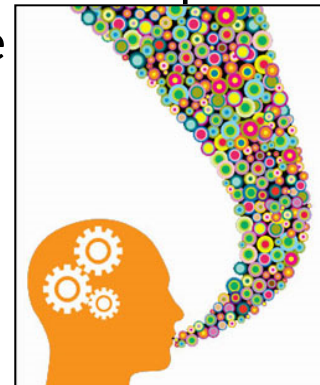
# Les difficultés de la RI : le facteur humain

- Le besoin d'information de l'utilisateur est parfois **vague** et toujours **subjectif**.
  - La **perte d'information** entre la réalité du besoin d'information et son expression peut être importante.
  - La pertinence d'un document pour une requête est une notion variable et très complexe à définir.
- ⇒ Il ne peut pas exister de système de recherche d'information parfait
- L'évaluation d'un système dépasse les aspects habituels de performance informatique
- L'humain est subjectif, versatile, et il utilise un **langage "naturel"** !



# Les difficultés de la RI : le facteur "langage"

- À la différence des langages artificiels, le **langage "naturel"** est :
    - **Implicite** : tout n'est pas dit dans les textes et leur compréhension requiert une importante connaissance sur le contexte et sur le monde
    - **Redondant** : la langue offre de nombreuses façons de formuler le même contenu
    - **Ambigu** : un même énoncé peut souvent être interprété de différentes façons
  - La recherche d'information est encore compliquée par le fait que :
    - Les mots peuvent jouer des rôles différents dans les textes
    - Les atomes de sens peuvent être des mots ou des groupes de mots (termes)
- Il est compliqué de **formuler son besoin d'information**  
(perte d'information entre besoin et requête)



# Caractère implicite de la langue

- Connaissance du langage et des **conventions langagières**

Q : *Le voisin est-il chez lui ?*

R : *Sa voiture est devant le portail*

(implicature conversationnelle)

- Connaissance du **contexte**

*C'est la deuxième fois qu'il reçoit un carton*

(Sport ? Courrier ? Accident ?)

- Connaissance du **monde**

*La Nouvelle-Zélande va tailler la France en pièces.*

(métonymie + langage figuré + actualité du rugby)

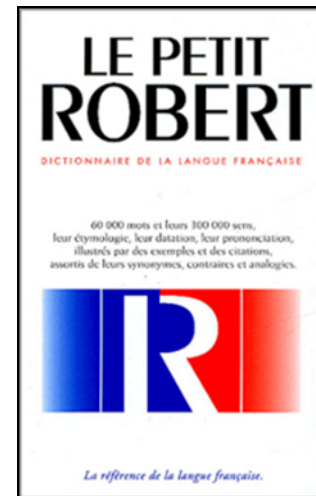
- **Déduction** (présupposition)

*Ravaillac a assassiné Henri IV en 1610.*

⇒ Henri IV est mort en 1610.

# Caractère redondant de la langue

- Au niveau lexical
  - **Synonymie** : vélo et bicyclette
  - **Hyperonymie** et **hyponymie** : véhicule  $\triangleleft$  vélo  $\triangleleft$  VTT
  - **Méronymie** et **holonymie** : pédale  $\diamond$  pédalier  $\diamond$  vélo
- Abréviations et sigles
  - *S'il-vous-plaît* et *SVP*, *VTT* et *Vélo Tout Terrain*
- Entre mots et expressions
  - **Périphrases** : *lave-vaisselle* et *machine à laver la vaisselle*
  - **Définitions** : *selle* et *petit siège, le plus souvent de cuir, d'un cycle ou d'un véhicule à deux roues à moteur*
- Glissements de sens (synonymie contextuelle)
  - *Il a écrit un **papier/article** sur la recherche d'information*
  - *\***Vos papiers/articles** s'il-vous-plaît !*



# Caractère *redondant* de la langue

- La **paraphrase** (synonymie au niveau syntaxique)

- Qui sera élu par le peuple en 2022?
- Qui le peuple choisira-t-il dans 5 ans ?
- Qui sortira vainqueur des urnes pour le prochain quinquennat ?



- Synonymie et paraphrase ne sont pas transitives !

Paul ressort souvent **excité** de la **récréation**.  
Paul profite bien de la **distraction** de la récréation.  
Entre deux cours, les **loisirs** sont bénéfiques à Paul.  
Une période de **détente** entre deux cours ne fait pas de mal à Paul.  
Paul se **repose** entre deux cours.



# Caractère ambigu de la langue

Les **homonymes** sont des mots qui ont une même graphie mais des sens différents



WIKIPÉDIA  
L'encyclopédie libre

Accueil  
Portails thématiques  
Index alphabétique  
Un article au hasard  
Contacter Wikipédia

Contribuer

Aide

Communauté

Modifications récentes

Accueil des nouveaux arrivants

Faire un don

Imprimer / exporter

Boîte à outils

Article [Discussion](#)

Lire [Modifier](#) [Afficher l'hist](#)

## Noyau

 Cette page d'*homonymie* répertorie les différents sujets et articles partageant un même nom.

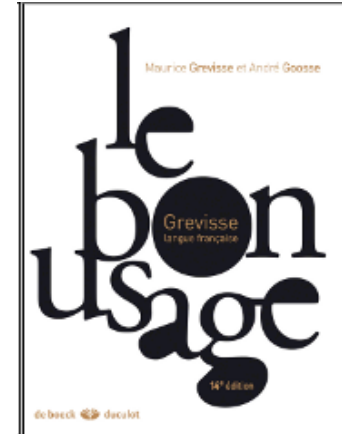
De manière générale, un **noyau** est la partie centrale située au milieu d'un autre objet. Plus particulièrement, le terme peut faire référence à :

- en **biologie**, un **noyau** est un organite qui contient la plupart du matériel génétique ;
- en **linguistique**, un **noyau** est partie fondamentale du **syntagme**, entourée de ses **satellites** ;
- en **botanique**, un **noyau** est la partie centrale, dure, d'une **drupe** ou fruit à noyau ;
- en **électrotechnique**, un **noyau** est la pièce magnétique sur laquelle un fil conducteur est enroulé afin de réaliser une l
- en **fonderie**, un **noyau** est la partie d'un moule permettant la réalisation des parties creuses d'une pièce ;
- en **géologie**, un **noyau** est la partie centrale approximativement sphérique de la **Terre** ou d'une **planète** ;
- en **informatique**, un **noyau** (aussi appelé **kernel**) est la partie fondamentale de certains **systèmes d'exploitation** ;
- en **mathématiques**,
  - en **algèbre**, le **noyau** d'un **morphisme de groupes** est un sous-groupe particulier du groupe de départ,
  - en **analyse fonctionnelle**, un noyau est une fonction permettant de définir un **opérateur intégral** ;
- en **physique**, un **noyau** est la région centrale constituée des **nucléons** d'un **atome** ;
- en **bande dessinée**, **Noyau** est le nom de l'illustrateur **Yves Nussbaum** ;

# Caractère *ambigu* de la langue

- Les **ambiguïtés syntaxiques** :

- Jean vend une tarte **aux pommes**.
- Jean vend une tarte **aux clients**.
- Jean a rapporté un vase **de Chine**.



- Les **anaphores** :

Ségolène trahit Martine. Son ancien mari lui en voulut longtemps.

- Les **ellipses** :

- Quelle est la taille **de la tour Eiffel** ? Le poids ?
- Les Stéphanois portent des **écharpes** vertes et les
- Toulousains des rouges et noires.



# Mots composés

- Les **mots composés** sont beaucoup moins polysémiques
- Les recherches ensemble dans les textes est bénéfique (mais compliqué)
- Ils ont un sens qui n'est pas la composition des sens des atomes
  - *Homme-grenouille*
  - *Pomme de terre*
  - *Traitement de texte*



© M. Heinrich, J. Negra

# Morphologie

La **morphologie** est l'étude de la construction des mots (leurs structures, variations, et similitudes)

- L'**analyse morphologique** permet de décomposer un mot et d'extraire principalement :
  - La **racine**
    - Le **lemme**
    - La **catégorie morphosyntaxique** (catégorie grammaticale)
    - Les **traits morphologiques**

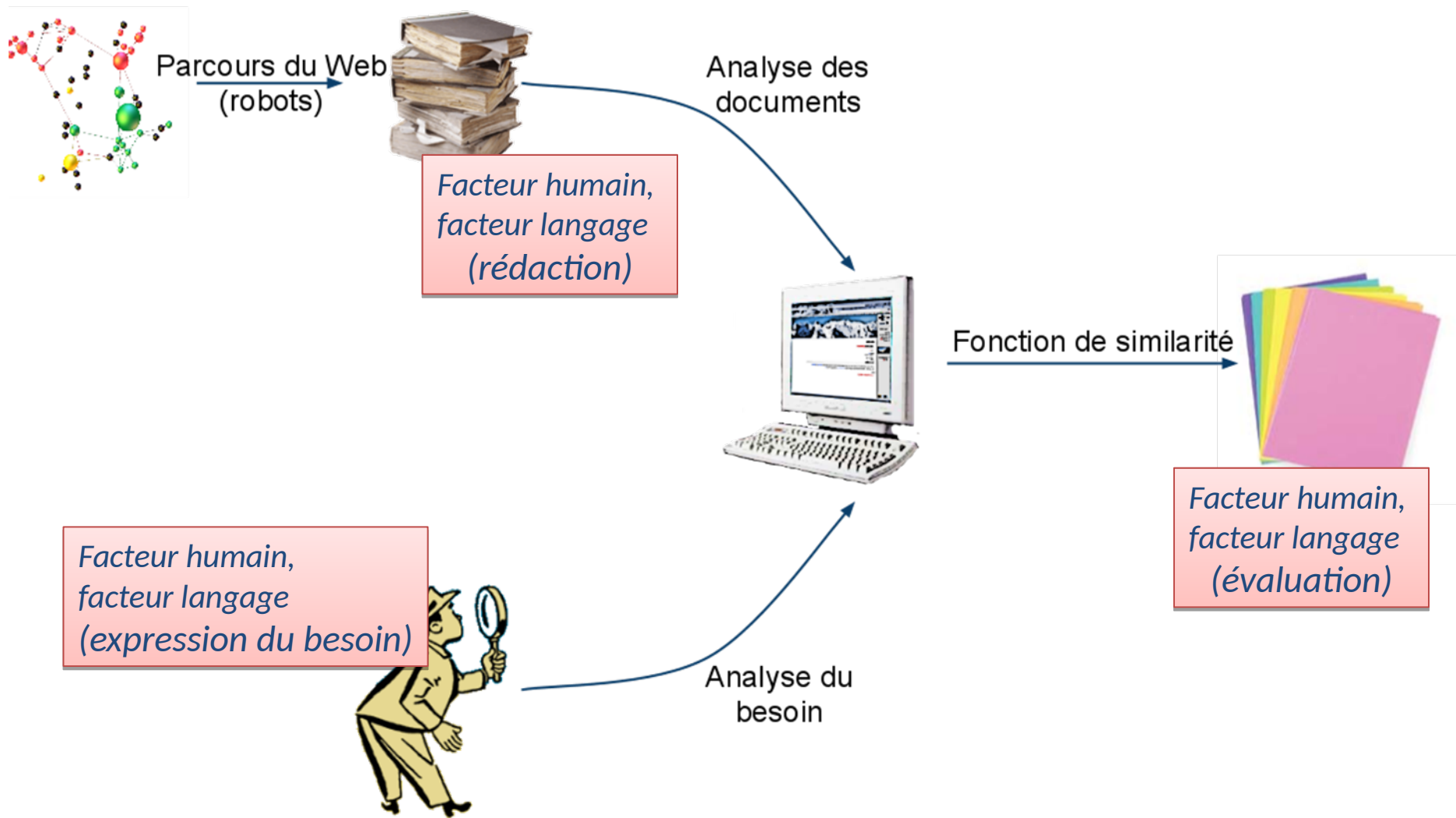
# *Construction de mots*

## **Flexion**

- **Composition**
- **Dérivation** (affixation)
  - **Préfixation**
  - **Suffixation**

(à combiner)

# Recherche d'Information



# Recherche d'Information

La recherche d'information est un **traitement statistique** de **chaînes de caractères**.

- ⊕ Elle permet de traiter de **grandes masses** d'informations.
- ⊕ Elle s'applique à du **texte en langue naturelle**, ne demande pas d'intervention des concepteurs des sites, ni de représentation particulière pour les connaissances manipulées.
- Les machines **ne comprennent rien au sens des informations** traitées.
- Un moteur de recherche ne permet **aucune inférence**, aucun
- ⊖ recoupement d'information.

# Accès au contenu sémantique

*Avec l'aide de Brigitte Safar*

# Accès au contenu sémantique

EGC 2016 Présentation▼ Conférence▼ e-EGC Ateliers Inscription Informations Pratiques▼

 **EGC 2016  
REIMS**

**16ème Conférence Internationale  
Francophone sur l'Extraction  
et la Gestion des Connaissances**

**18 - 22 Janvier 2016**  
<http://egc2016.univ-reims.fr>



Dans le cadre des 50 ans de l'IUT de Reims, l'Université de Reims Champagne-Ardenne et l'IUT de Reims ont le plaisir d'accueillir EGC2016 sur le site de l'IUT du 18 au 22 janvier 2016.

La conférence Extraction et Gestion des Connaissances (EGC) est un événement annuel important réunissant des chercheurs, des industriels et des étudiants des communautés de l'apprentissage automatique, ingénierie et représentation des connaissances, statistique et analyse de données, fouille de données, systèmes d'information, bases de données afin de discuter des innovations et des résultats scientifiques les plus récents de ces domaines.

# Accès au contenu sémantique

Le texte des pages Web est en langage naturel

- En HTML, le **marquage** (ou **balisage**) existe,
- mais concerne :
  - La **mise en forme** (fonte, taille, couleur)
  - Les **hyperliens**
- Le **contenu sémantique** est accessible aux hommes, mais pas (facilement) aux machines...



Comment rendre le Web "compréhensible" pour libérer l'utilisateur du travail de recherche, d'intégration et d'exploitation des données ?



# Ajouter du marquage (XML) ?

**<event>**

**<name>**16<sup>èmes</sup> Conférence Francophone "Extraction et Gestion des Connaissances"**</name>**

**<date>**Du 18 au 22 janvier 2016**</date>**

**<lieu>**Reims, France**</lieu>**

**<president>**Président d'honneur : Jim Hendler, Université Libre de Bruxelles, Belgium**</president>**

**</event>**

# *Mais que voit la machine ?*

**<event>**

**<name>**16<sup>èmes</sup> Conférence Francophone "Extraction et Gestion des Connaissances" **</name>**

**<date>**Du 18 au 22 janvier 2016**</date>**

**<lieu>**Reims, France**</lieu>**

**<president>**Président d'honneur : Jim Hendler, Université Libre de Bruxelles, Belgium**</president>**

**</event>**

# Intégration d'information

Title	Actors	Director
...	...	...
You Will Meet a Tall Dark Stranger	N. Watts A. Banderas	Woody Allen
Vicky Christina Barcelona	S. Johansson J. Bardem	Woody Allen
...	...	...

*Le meilleur plan pour un Woody Allen ce soir, à Paris, suivi d'un resto chinois pas trop loin.*

Nom	Adresse	Spécialité
...	...	...
Saveurs d'Asie	34, rue... 75013 Paris	Chinois
...	...	...

```
<film>
  <titre>Vous allez rencontrer un bel et sombre
    inconnu</titre>
  <salles>
    <adresse>
      <rue>128-162, avenue de France</rue>
      <ville>Paris (75013)</ville>
    </adresse>
  </salles>
  ...
```

# *Intégration d'information*

- Objectif :
  - Répondre à des **requêtes complexes**, de **haut niveau d'abstraction**
  - Interface entre l'utilisateur et les sources lui permettant donnant l'impression d'utiliser un système **homogène** et **centralisé**
  - Développement de serveurs d'informations au-dessus des sources d'informations existantes
- Difficultés :
  - Des sources **dispersées**
  - Des données **hétérogènes**
  - Des organisations et des langues **différentes**
  - Du texte **non structuré**

# Intégration d'information

- Trois points de vue

## Entrepôt de données

- Construction de **bases de données réelles**
- **Regroupement** des informations pertinentes pour les **applications** considérée
- Définition du **schéma**, de son peuplement, de sa mise à jour

## Médiateur

- Laisser les données où elles sont (dans les **sources**)
- Exploiter les **descriptions abstraites** de ces sources
- Description d'un **schéma global**
- Construction d'un **plan de requêtes** pour l'interrogation

## Web sémantique

- **Annotation** des sources
- **Raisonnement** sur les **métadonnées** ainsi obtenues

# W3C (*World Wide Web Consortium*)

- Le W3C, un organisme de **standardisation** à but non lucratif
- Fondé en 1994
- Le but du W3C : promouvoir l'**évolutivité** du Web et l'**interopérabilité** des technologies (HTML, XHTML, CSS, XML, RDF, PNG, SVG, SOAP...).
- Ses **recommandations** :
  - **Ajouter de la sémantique** pour rendre les ressources plus accessibles à des processus automatiques.
  - Attacher aux ressources des **étiquettes** qui les définissent : adjoindre aux données des **métadonnées**.
  - Définir formellement les termes utilisés.
  - Définir des **ontologies** (ensemble de termes et de relations sur un domaine).
  - Développer un **langage** standard pour représenter les ontologies.

# Vers une représentation des connaissances...

## La Joconde de Léonard de Vinci

Entrer sur le site de [La Joconde](#) de leonard de vinci

Google™ Recherche personnalisée

Rechercher

Bienvenue sur mon site qui a pour but de parler de [la Joconde](#) du célèbre peintre [Léonard de Vinci](#). J'essaierai de faire un portrait de l'artiste peintre de Vinci et bien entendu de celui de [Mona Lisa](#). Tout le monde sait que [La Joconde est devenue un tableau](#) que l'on pourrait considérer de Mythique, une peinture à l'huile réalisée vers le début du XVIème siècle ( peut être en 1503 ou quelques années après ). [Ce tableau est exposé aux musée du Louvre à Paris](#) dirigé par Vincent Pomarède, autrefois sous la responsabilité du conservateur Cécile Scamizzi. Il est situé en profondeur à cause de la lumière et de certains photographes qui ne peuvent s'empêcher d'utiliser le Flash, à savoir que la lumière détruit à long terme l'oeuvre...

Cliquez sur le tableau de la Joconde pour entrer dans le site :



# Vers une représentation des connaissances...



WIKIPÉDIA  
L'encyclopédie libre

Accueil  
Portails thématiques  
Index alphabétique  
Un article au hasard  
Contacter Wikipédia

Contribuer  
Aide  
Communauté  
Modifications récentes  
Accueil des nouveaux arrivants  
Faire un don  
Imprimer / exporter  
Boîte à outils

[Créer un compte](#) ou [se connecter](#)

Article [Discussion](#)

Lire [Voir la source](#) [Afficher l'historique](#)

Rechercher



## La Joconde

**La Joconde**, ou *Portrait de Mona Lisa*, est un tableau de **Léonard de Vinci**, réalisé entre 1503 et 1506 qui représente un buste, probablement celui de la florentine **Mona Lisa del Giocondo**. Cette peinture à l'huile sur panneau de bois de peuplier de 77 x 53 cm est exposée au musée du Louvre à Paris. *La Joconde* est l'un des rares tableaux attribués de façon certaine à Léonard de Vinci.

*La Joconde* est devenue un tableau éminemment célèbre car, depuis sa réalisation, nombre d'artistes l'ont prise comme référence. Elle constitue en effet l'aboutissement des recherches du **xv<sup>e</sup> siècle** sur la représentation du portrait. À l'époque romantique, les artistes ont été fascinés par l'énigme de *La Joconde* et ont contribué à développer le mythe qui l'entoure, en faisant de ce tableau l'une des œuvres d'art les plus célèbres du monde, si ce n'est la plus célèbre.

### Sommaire [masquer]

#### 1 Historique

- 1.1 Ancien Régime
- 1.2 XIX<sup>e</sup> siècle
- 1.3 Vol du tableau
- 1.4 La Joconde pendant les deux guerres mondiales
- 1.5 Depuis les années 1960

#### 2 Description

**La Joconde**



[Portail de la peinture](#)



[Portail de l'histoire de l'art](#)



[Portail de la Renaissance](#)

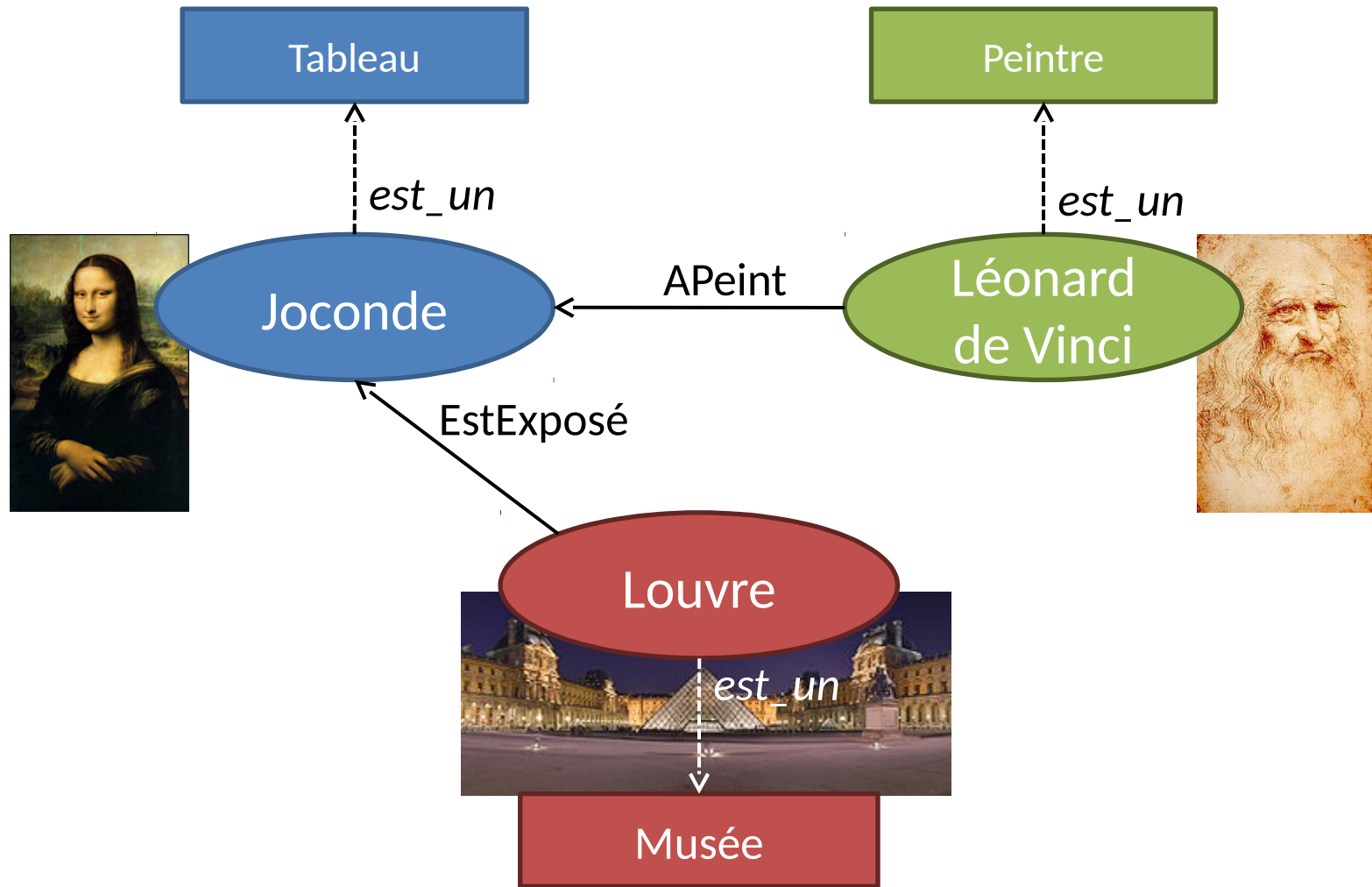


[Portail de l'Italie](#)

Catégories : [Tableau de Léonard de Vinci](#) | [Œuvre conservée au Louvre](#) | [Portrait](#) | [Tableau du XVIe siècle](#) | [+]



# Vers une représentation des connaissances...



# *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

- + Il permet de donner du sens aux informations.
- + Il autorise une intégration d'information et des raisonnements.
- Un gros travail de représentation des connaissances ou de construction d'un schéma global est nécessaire.
- Il manipule des données structurées et non plus du texte.

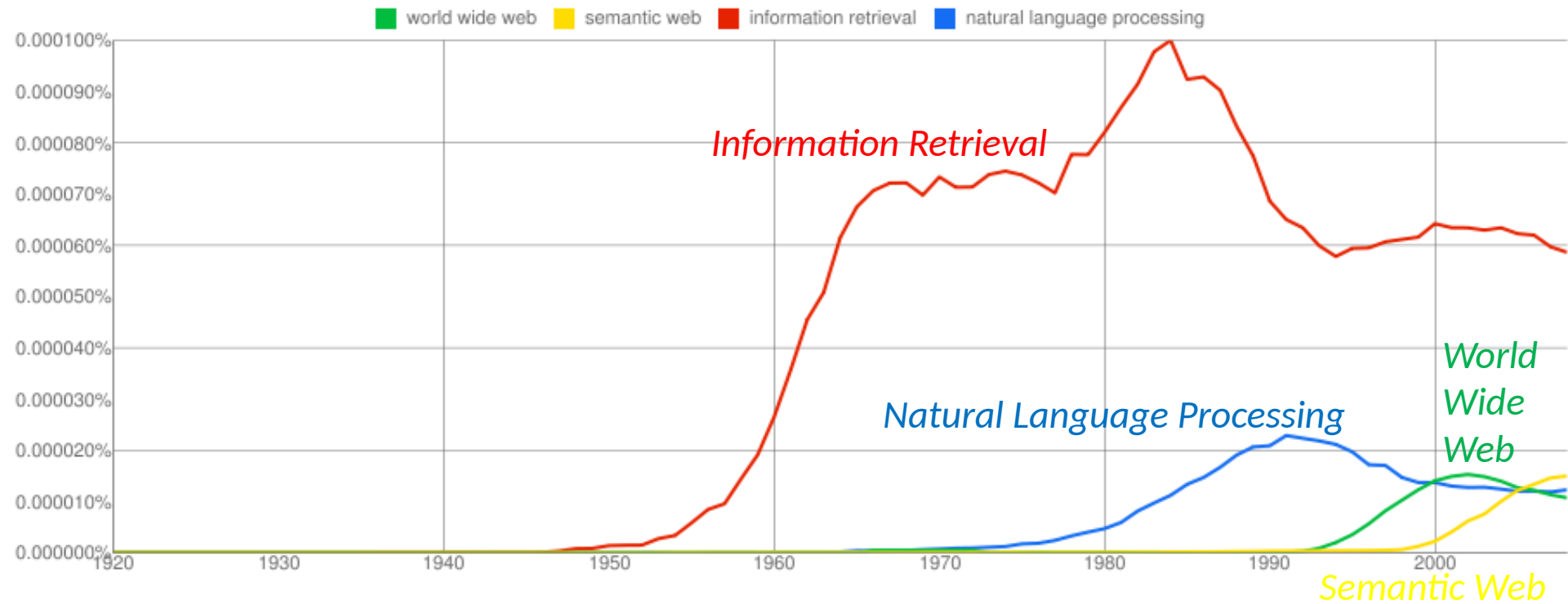
# *Conclusion de l'introduction*

- Les **humains** préfèrent les documents textuels dans leur langue naturelle.
- Les **machines** préfèrent des données structurées et sans surprise.

Il est difficile de construire des documents appropriés à la fois pour les machines et pour les humains.

⇒ *Recherche d'information et web sémantique sont complémentaires.*

# Historique



Google labs

Books Ngram Viewer