



---

# Mise en relation automatique de documents scientifiques et recommandation de lectures

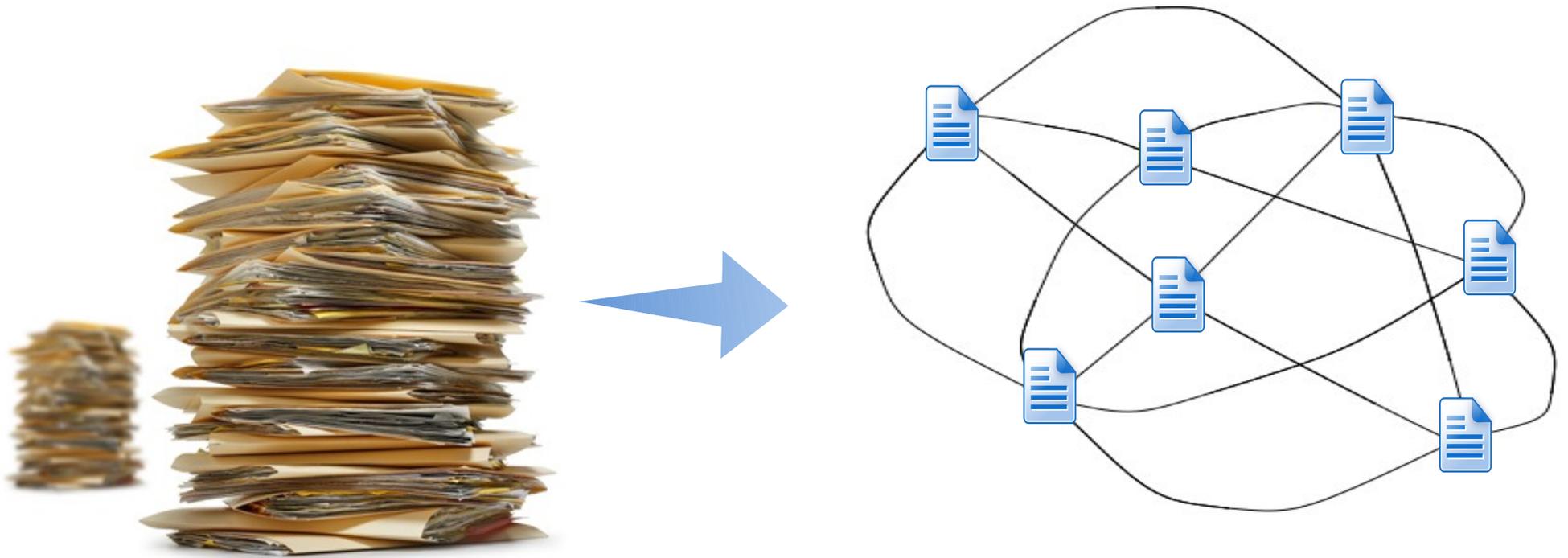
---

Chahinez BENKOUSSAS  
Directeur de thèse : Pr. Patrice BELLOT

# INTER-TEXTES

---

- Valoriser des ensembles de documents de nature scientifique ou économique.
- Développer un panel de méthodes facilitant leur mise en relation



# Plate-formes d'OpenEdition

**OpenEdition** : [OpenEdition Books](#) [Revues.org](#) [Calenda](#) [Hypothèses](#) [Lettre & alertes](#) [OpenEdition Freemium](#)

Ressources électroniques et communication scientifique [Twitter](#) [Facebook](#) [Google+](#)

Accueil FR EN

## openedition

REVUES.ORG CALENDAL HYPOTHESES.ORG

Qui sommes-nous ?  
OpenEdition Freemium  
Adhérer et se former  
S'abonner à la Lettre  
Flux du portail

### CATALOGUES

1303 LIVRES	25437 ÉVÉNEMENTS
413 REVUES	854 CARNETS

### RECHERCHE

[ACTUALITÉS DE REVUES.ORG](#) [ACTUALITÉS DE CALENDAL](#) [ACTUALITÉS D'HYPOTHESES](#)

OpenEdition est un portail de ressources électroniques en sciences humaines et sociales (OpenEdition Books, Revues.org, Hypothèses, Calenda). Si vous souhaitez que votre établissement s'abonne à ses services et vous donne accès aux formats détachables (PDF, ePub) du bouquet *Open Access Freemium*, rendez-vous sur la page de présentation d'OpenEdition Freemium.

#### LECTEURS

OpenEdition propose un vaste catalogue de publications scientifiques en sciences humaines et sociales, principalement en libre accès. Des services complémentaires sont proposés via les bibliothèques et institutions abonnées au programme OpenEdition Freemium : formats de lecture, outils de recherche, service d'alertes, etc.

#### ÉDITEURS

OpenEdition est une infrastructure complète d'édition électronique au service de la communication scientifique en sciences humaines et sociales. Elle propose aux équipes éditoriales et aux éditeurs un ensemble de solutions adaptées à la publication de livres, revues, carnet de recherches et annonces d'événements scientifiques.

#### BIBLIOTHÈQUES

Le programme OpenEdition Freemium offre un éventail de services aux bibliothécaires et à leurs lecteurs. Ces services sont destinés à aider les bibliothèques à gérer leur abonnement à OpenEdition et à faciliter son utilisation par les usagers.

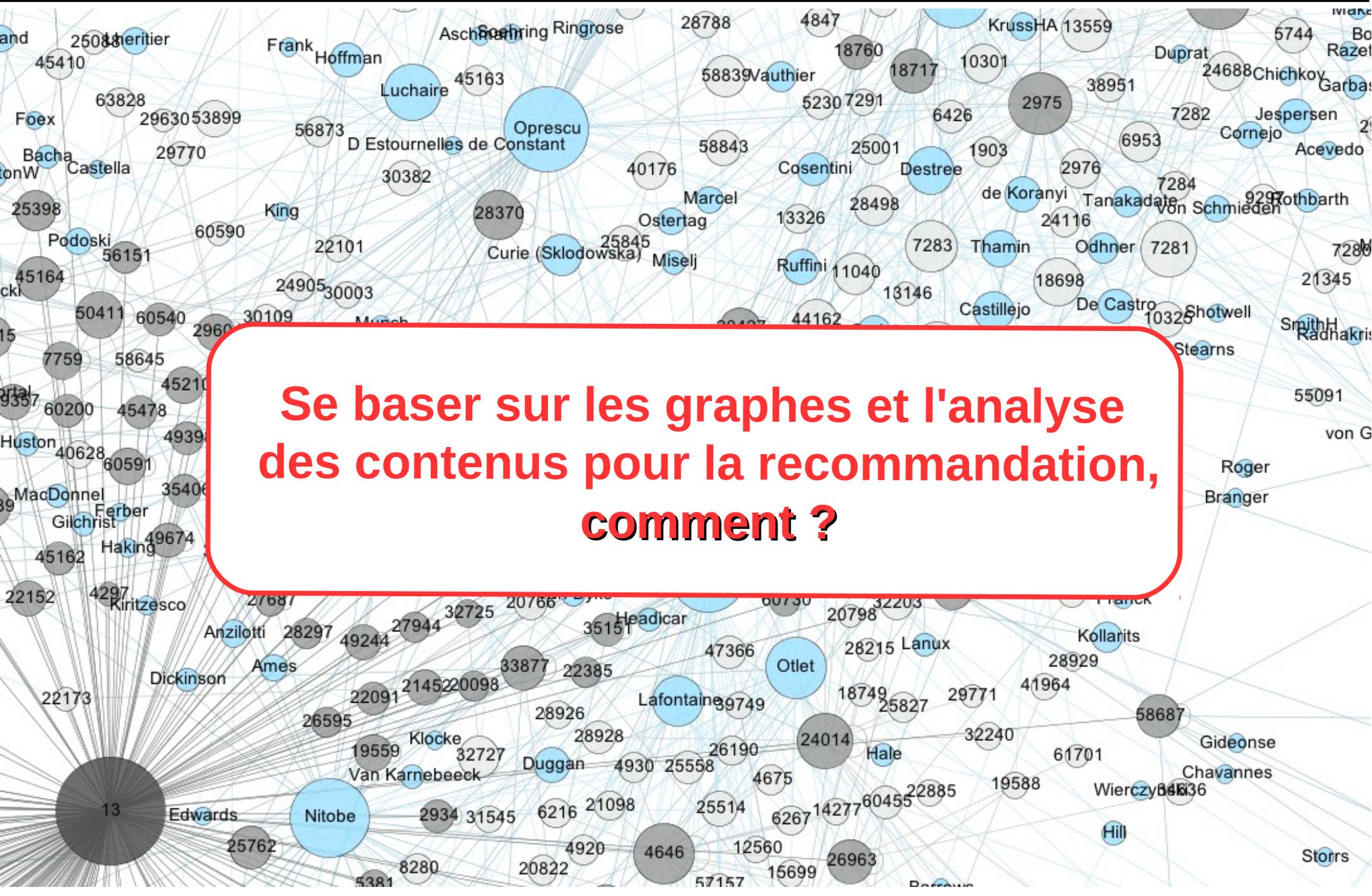
**revues.org**  
+400 journals

**hypothèses**  
+800 blogs

**calenda**  
+25000 events

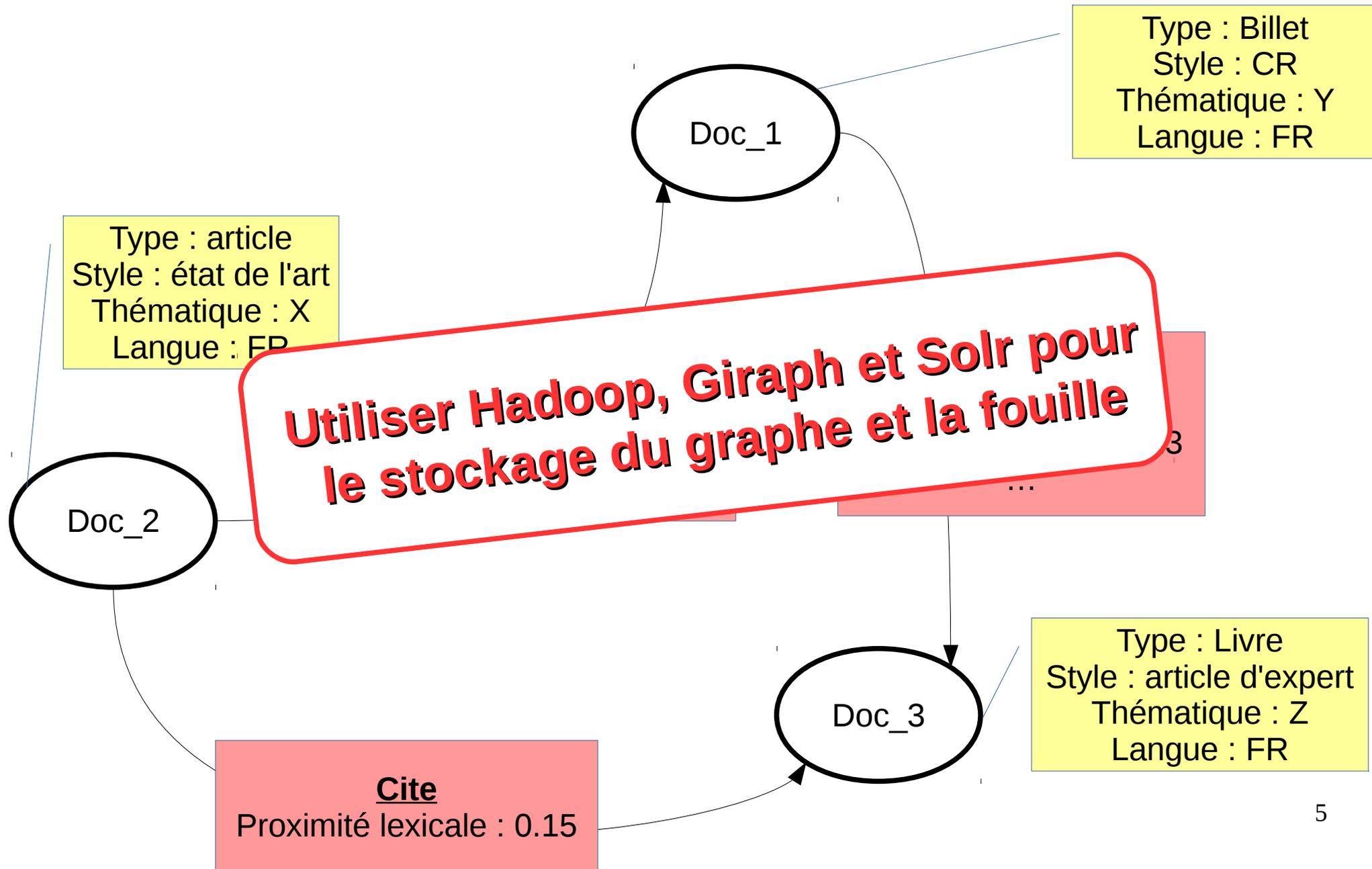
**OpenEdition books**  
+1300 books

# Modélisation des données



**Se baser sur les graphes et l'analyse des contenus pour la recommandation, comment ?**

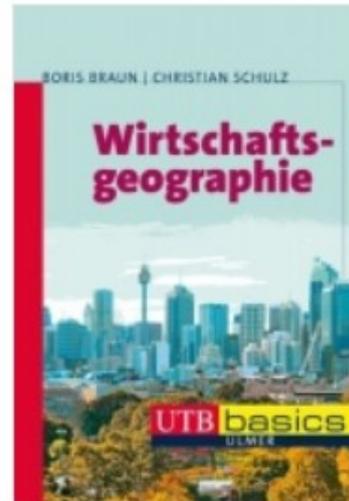
# Modélisation des données



# Systeme de Recommandation

- Notre cas :

Pas de votes mais des Compte Rendu de lecture (CR) :



Full size image  
Credits : © UTB GmbH

To come straight to the point, the new textbook *Wirtschaftsgeographie* [Economic Geography] by Boris Braun and Christian Schulz is a highly recommendable read. In pursuit of its goal to outline “a very dynamic and multifarious sub-domain of geography” [author’s translation] (p. 249), this book is mainly directed at current and prospective students of the field of economic geography (p. 6). It keeps both these promises in quite an appealing manner. Alongside an insightful introduction to the myriad of theoretical structures applied in economic geography, its strength lies mainly in its engaging and stylistically coherent review as well as in the systematic cross-linkage of major topics. Valuable didactic tricks – such as info boxes giving a quick overview at the beginning of each chapter, additional text boxes that provide examples to underline the theoretical structures discussed, highlighted keywords in the body of the text, a rich stock of illustrations and charts, as well as review questions and recommendations for further reading – make this book a helpful companion for students. Topics are successfully cross-linked through clearly arranged and colour coded references of lexical precision as well as by repeatedly addressing, relating, and embedding major theoretical approaches from a variety of perception angles. The book embraces the exacting standards of both the authors, who explain them as follows: “Currently relevant basic assumptions, theories, and models of economic geography, their differences as well as their manifold linkages shall be presented more vividly than in comparable textbooks” [author’s translation] (p. 6).

# Systeme de Recommandation

---

## Exploitation des CR pour la recommandation

- Recherche de documents porteurs d'opinion
- Détection automatique des Comptes Rendus de lecture (méthodes de classification thématique pour la classification en Genre) [*Benkoussas et al., LREC 2014*]
- Construction d'un corpus de critiques.
- Liaison automatique (Bilbo) de chaque CR avec son livre, article, ...

# Systeme de Recommandation

## (INEX Social Book Search 2014)

!-- version 1.0 / 2009-11-04T13:44:03+01:00 -->

```
-<book>
  <isbn>0001360000</isbn>
  <title>Mog's Kittens</title>
  <ean>9780001360006</ean>
  <binding>Board book</binding>
  <label>HarperCollins UK</label>
  <listprice>$6.99</listprice>
  <manufacturer>HarperCollins UK</manufacturer>
  <publisher>HarperCollins UK</publisher>
  <readinglevel>Ages 4-8</readinglevel>
  <releasedate/>
  <publicationdate>1994-09-01</publicationdate>
</book>
<topic id="1116">
  <title>Which LISP?</title>
  <mediated_query>introduction book to Lisp</mediated_query>
  <group>Purely Programmers</group>
  <narrative> It'll be time for me to shake things up and learn a new language soon. I had started on Erlang a while back and getting back to it might be fun. But I'm starting to lean toward Lisp--probably Common Lisp rather than Scheme. Anyone care to recommend a good first Lisp book? Would I be crazy to hope that there's one out there with an emphasis on using Lisp in a web development and/or system administration context? Not that I'm unhappy with PHP and Perl, but the best way for me to find the time to learn a new language is to use it for my work... </narrative>
  <catalog>
    <book>
      <LT_id>859035</LT_id>
      <entry_date>2006-04</entry_date>
      <rating>0.0</rating>
      <tags>livingry, hunger, poverty, policy</tags>
    </book>
    <book>
      <LT_id>569191</LT_id>
      <entry_date>2005-11</entry_date>
      <rating>0.0</rating>
      <tags>law, the firm, economics</tags>
    </book>
  </catalog>
  <content>
    In a tearful heartwarming story, we see that Mog has a next of kin. With her hands full with two little kittens, Mog is unable to leave her cat box (with the kittens) to sleep with Nicky. Two of Nicky's friends come by to visit Mog and the kittens, wishing that they themselves had a kitten. The kittens actually want homes of their own. Nicky decides to give the two kittens to the very same (two young girls) friends who were admiring them. The two girls' families are shown, happy, with Mog's kittens. While Nicky and Mog are sad at first, they both realize that it's what's best for all involved, and Mog resumes sleeping with Nicky again.
  </content>
  <rating>5</rating>
  <totalvotes>0</totalvotes>
  <helpfulvotes>0</helpfulvotes>
</review>
```

# Systeme de Recommandation

(INEX Social Book Search 2014)

---

- Modèles de Recherche d'Information :
  - InL2
  - SDM (Sequentiel Dependance Model)
  - Pseudo Relevance Feedback
    - Query Expansion with Tags
    - Query Expansion with Concepts
  - Combinaison du score des modèles de RI (SDM et InL2) avec la moyenne des votes des utilisateurs

# Systeme de Recommandation

(INEX Social Book Search 2014)

✓ **Classés 2ème** sur INEX SBS 2014

Run	nDCG@10	Recip Rank	MAP	Recall@1000
Best_Run_2014	0.142	0.275	0.107	0.426
<b>InL2</b>	<b>0.128</b>	<b>0.236</b>	<b>0.101</b>	<b>0.441</b>
InL2Feedback	0.114	0.230	0.094	0.434
InL2tagFeedback	0.102	0.212	0.075	0.388
SDM_Rating	0.062	0.120	0.047	0.314
SDM_Concept	0.056	0.118	0.039	0.253
SDM_Tag_Feedback	0.055	0.112	0.040	0.267

Apport non significatifs des extensions rajoutées au modèle InL2

# Conclusion

---

- Projet INTER-TEXTES
- Mise en relation des documents scientifiques pour la recommandation
- Détection automatique des CR par la classification des documents par genre (CR & NCR)
- Modélisation des données en graphe
- Exploiter le graphe pour la recommandation



---

Merci pour votre attention

---

Pour nous suivre → URL du carnet du Lab : <http://lab.hypotheses.org>