



## Recherche d'Information (RI)

**Niveau du cours :** Master (M2) **Parcours(s) :** DSC

**Crédits ECTS :** 4

**Enseignants :** Michel Beigbeder (EMSE, Saint-Étienne)  
Mathias Géry (UJM, Saint-Étienne)

**Période d'enseignement :** 3<sup>ème</sup> semestre **Langue d'enseignement :** Français

**Pré-requis :** Algorithmique, Programmation

### Objectifs :

1. Connaître les principes des modèles théoriques des systèmes de RI.
2. Connaître les algorithmes principaux de mise en œuvre des systèmes de RI.
3. Être capable d'évaluer ces systèmes qualitativement et quantitativement.
4. Être capable d'adapter ces principes et algorithmes à des données structurées.
5. Être capable de mettre en œuvre ces principes et algorithmes dans un Système de Recherche d'Information (SRI) traitant des données hétérogènes du Web.

**Mots clés :** Recherche d'Information (RI), Modèles de RI, Indexation, Pondération, Évaluation, Compétition de RI, Documents Structurés, Hypertextes, Web.

### Syllabus :

1. Le modèle de la Recherche d'Information.
2. L'indexation des documents.
3. Les différents modèles : modèle booléen, vectoriel, probabiliste et leurs extensions.
4. L'évaluation par des mesures quantitatives des Système de Recherche d'Information (SRI).
5. La prise en compte de la structure intra-document (ex. XML).
6. La prise en compte de la structure inter-documents (ex. les liens hypertextes).
7. Projet : de la théorie à la pratique, mise en œuvre d'un SRI sur des données structurées du Web, participation à une mini-compétition de Recherche d'Information.

**Volume horaire :** CM (20h), TD (14h), TP (6h).

**Contrôle des connaissances :** Un contrôle théorique (coefficient 1, durée 2h), un projet (coefficient 1). Rattrapage : contrôle théorique (durée 2h).

### Bibliographie et ressources :

Christopher D. Manning, Prabhakar Raghavan et Hinrich Schütze, Introduction to Information Retrieval, Cambridge University Press, 2008 (disponible en ligne à <http://nlp.stanford.edu/IR-book/>)

Massih-Reza Amini et Eric Gaussier, Recherche d'information, Applications, modèles et algorithmes. Fouille de données, décisionnel et big data, Eyrolles, 2013.

Soumen Chakrabarti, Mining the Web, Discovering Knowledge from Hypertext Data,  
Morgan-Kaufmann Publishers, 2002 (disponible en ligne à  
<https://www.cse.iitb.ac.in/~soumen/mining-the-web/>)

### **Informations complémentaires/contacts :**

Michel Beigbeder  
École des Mines de Saint-Étienne  
E-mail: [mbeig@emse.fr](mailto:mbeig@emse.fr)  
Web : <http://www.emse.fr/~mbeig/>

Mathias Géry  
Université Jean Monnet, Saint-Étienne  
E-mail : [mathias.gery@univ-st-etienne.fr](mailto:mathias.gery@univ-st-etienne.fr)  
Web : <http://perso.univ-st-etienne.fr/gerymath/>