



LA RECOMMANDATION DE LECTURE

10 juin 2012

Premier site français sur le livre



Première communauté de lecteurs

1

52 000 lecteurs membres

partagent leurs lectures, critiques et citations

2

Site numéro 1 en trafic

1,1 million de visites
830 000 uniques/mois

consultent les notices enrichies du site à la recherche de leur prochaine lecture

The screenshot shows a user profile for 'gtheisseire'. The profile includes navigation tabs for 'Accueil', 'Profil', 'Livres (1373)', 'Amis (118)', and 'Autres'. There are social media icons for RSS and a message icon. The main content area features a section for 'Actualités' with a message about a new sign system. Below this, there are several book recommendations with user avatars and star ratings. For example, 'Bigmammy' has added a review for 'L'Homme de l'art' by Pierre Assouline, and 'LiliGalipette' has appreciated reviews for 'MissAlfie' and 'Les petits ruisseaux' by Pascal Rabat. A 'Quiz' section is also visible, with a question about Albert Camus's 'L'Étranger'.

The screenshot shows a book page for 'Limonov' by Emmanuel Carrère. The page includes a 'Prix Renaudot' badge and a 'Note moyenne' of 3.87/5 based on 151 notes. There are buttons for 'Ajouter à mes livres', 'Lire un extrait', 'Ajouter une critique', and 'Ajouter une citation'. A 'Critiques et avis' section shows a review by 'Par Malaura' from November 22, 2011, with a 5-star rating. The review text discusses the character's complexity and the author's style. There are also social media sharing options and a 'Faire découvrir' section with links to Amazon, Facebook, and a blog.

(Source : Google Analytics, mai 2012)

Quelles formes de recommandation



1

Recommandation sociale



de **totom**, le 24 avril 2012

J'ai sans doute un peu forcé le trait! mais je maintiens ce que je dis sur Canguilhem: même s'il n'a pas de front les thèmes de l'éloge de la fuite (en particulier la liberté), il y aurait beaucoup à dire, je essayer de me condenser un peu: prenons pour exemple 2 idées centrales chez lui qui me paraissent totalement contraires à une vision déterministe de l'individu: la subjectivité et la résistance:

2

Automatisée book centric

♥ Vous aimez ce livre?
Babelio vous suggère

 <p>Musicophilia : La musique, le cerveau et nous Oliver Sacks</p>	 <p>A la recherche de l'intelligence artificielle Daniel Crevier</p>
--	--

3

Automatisée user centric

 **Lecteurs proches** > voir plus

 <p>Flanfl Indice 100</p>	 <p>jerome10 Indice 80</p>	 <p>Mam-Coz Indice 77</p>
		

Quelles difficultés?



1

Pertinence qualitative

Ex : GEB => Musique ? Logique?
Intelligence artificielle?

2

Pertinence quantitative

Ex : Favoriser la longue traîne?
Rousseau => Voltaire ou Bachofen

3

Arbitrage sensibilité/précision

Prendre des risques?

$$F_{\beta} = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{\beta^2 \cdot \text{precision} + \text{recall}}$$

Quelles difficultés?



4

Data sparsity

Qualification d'un nouvel entrant
Livre avec peu d'infos/nouveautés

5

Scalabilité algorithmique

Robustesse aux changements d'échelle

6

Performance

50 000 membres
2 millions de titres
3 millions de métadonnées
...

Des problématiques propres au livre



1

Répétition d'auteur

Titre de Kundera => Titre de Kundera?

2

Répétition de série/collections

Gallimard Noire => Gallimard Noire ?

3

Thématiques semi-structurées

BD/Essai/Roman ...

Besoin de mesurer les présupposés cognitifs du lecteur

Collecte et traitement
de métadonnées

Algorithme de ranking
user/book centric

- **Prime aux réseaux sociaux : captation de données**
- **Nouveaux enjeux : structuration de métadonnées**

Métadonnées qualifiées



1

Métadonnées d'usage

- Lectures
- Notes
- Clics
- Visites
- Taux d'ajout, Etc.

2

Métadonnées descriptives

roman aventure littérature récits fantasy
science-fiction anticipation fantastique historique suspense
manga essai bande dessinée théâtre nouvelles
poésie classique roman historique biographie
autobiographie témoignage jeunesse littérature jeunesse
roman d'amour histoire thriller
romans policiers et polars contes art enfance
enfants drame amitié mort femmes enquêtes adolescents
magie humour religion musique horreur politique peinture
amour guerre seconde guerre mondiale société cuisine voyages
science mythologie famille vampires philosophie psychologie
états-unis littérature américaine Paris (France) chine inde italie
russe afrique japon littérature japonaise littérature française
littérature anglaise angleterre littérature russe littérature italienne
littérature allemande moyen-âge 19ème siècle

3

Métadonnées non structurées

- Critiques
- Extraits
- Résumés
- Titres

4

Métadonnées structurées

- Taxinomie
- Date
- Nationalité
- Ontologie, etc.

Algorithmie classique / Statistique

- Notation explicite/implicite
- Clusters de populations
- Corrélation statistique
- Analyse qualitative prédictive

Traitements sémantiques

- Association des inflexions linguistiques : {roman, romans}
- Association d'erreurs orthographiques : {cuisine, cuisne}
- Association synonymique : {deuxième guerre mondiale, ww2}
- Association thématique de premier ordre : {roman policier, polar et roman noir}

Traitement documentaire assisté



id	titre	999	lecteurs	nb	livres	consult	cluster	no des	de trce	bonus fiche	notaire	thème	seo	periode	redondance	drive	agg	reaj2	A_cluster	dt_creation	pre_vis	pre_reco
122	indie	ind	245	3587	838	<input type="checkbox"/>	a europe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	0000-00-00 00:00:00	4	1						
255	mythologie	ind	339	4783	1455	<input type="checkbox"/>	w mythologie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	0000-00-00 00:00:00	3	1						
134	essais	ind	330	3732	1016	<input type="checkbox"/>	a	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	0000-00-00 00:00:00	3	1						
236	indie	ind	329	2170	537	<input type="checkbox"/>	w axe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	0000-00-00 00:00:00	3	1						
175	roman	ind	320	2848	664	<input type="checkbox"/>	w europe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	0000-00-00 00:00:00	4	1						
64	essais	ind	310	4533	700	<input type="checkbox"/>	a mesure européenne	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	0000-00-00 00:00:00	2	1						
150	essai	ind	309	3887	1111	<input type="checkbox"/>	w	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	0000-00-00 00:00:00	3	1						
286	écrits	ind	304	1786	486	<input type="checkbox"/>	a politique	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	0000-00-00 00:00:00	3	1						
667	roman Aristotele	ind	304	1485	187	<input type="checkbox"/>	a correspondance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0000-00-00 00:00:00	1	1
100	essais	ind	302	4817	1543	<input checked="" type="checkbox"/>	w sciences humaines	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	0000-00-00 00:00:00	3	1						

Algorithmes de filtrage



1

Filtrage collaboratif

Importance des notes, clusters de lecteurs

«philosophie» => Musso

«philosophie» => Schopenhauer

Dimension temporelle

2

Filtrage CB

roman aventure littérature récits fantasy
science-fiction anticipation fantastique historique suspense
manga essai bande dessinée théâtre nouvelles
poésie classique roman historique biographie
autobiographie témoignage jeunesse littérature jeunesse
roman d'amour histoire thriller
romans policiers et polars contes art enfance
enfants drame amitié mort femmes enquêtes adolescents
magie humour religion musique horreur politique peinture
amour guerre seconde guerre mondiale société cuisine voyages
science mythologie famille vampires philosophie psychologie
états-unis littérature américaine Paris (France) chine inde italie
russie afrique japon littérature japonaise littérature française
littérature anglaise angleterre littérature russe littérature italienne
littérature allemande moyen-âge 19ème siècle

3

Réseau neuronal

Feedback pertinence

Identification des effets de bords

Algorithmes autocorrectifs

1 Optimisation algorithmique

- Rapidité
- Performance
- Scalabilité

2 Nouvelles méthodes de requête

- Recherche à facette
- Requête en langage naturel

3 Nouvelles métadonnées

- Détection automatique d'émotions
- Linguistique informatique

Contacts



Pierre Fremaux

pierre@babelio.com

+33 (6) 63 01 56 05

Guillaume Teisseire

guillaume@babelio.com

+33 (6) 60 53 74 99