



La Data au service de l'innovation dans les Services d'Information Documentaires (SID) universitaires nationaux

Amine Sennouni

► To cite this version:

Amine Sennouni. La Data au service de l'innovation dans les Services d'Information Documentaires (SID) universitaires nationaux. *Revue française des sciences de l'information et de la communication*, 2017, 10, <10.4000/rfsic.2759>. <hal-02298522>

HAL Id: hal-02298522

<https://hal.science/hal-02298522v1>

Submitted on 10 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-SA 4.0 - Attribution - Non-commercial use - ShareAlike - International License

La Data au service de l'innovation dans les Services d'Information Documentaires (SID) universitaires nationaux

Sennouni Amine



Édition électronique

URL : <http://journals.openedition.org/rfsic/2759>

DOI : 10.4000/rfsic.2759

ISSN : 2263-0856

Éditeur

Société Française de Sciences de l'Information et de la Communication

Ce document vous est offert par Conservatoire national des arts et métiers (Cnam)

le cnam

Référence électronique

Sennouni Amine, « La Data au service de l'innovation dans les Services d'Information Documentaires (SID) universitaires nationaux », *Revue française des sciences de l'information et de la communication* [En ligne], 10 | 2017, mis en ligne le 01 janvier 2017, consulté le 10 octobre 2019. URL : <http://journals.openedition.org/rfsic/2759> ; DOI : 10.4000/rfsic.2759

Ce document a été généré automatiquement le 10 octobre 2019.



Les contenus de la *Revue française des sciences de l'information et de la communication* sont mis à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.

La Data au service de l'innovation dans les Services d'Information Documentaires (SID) universitaires nationaux

Sennouni Amine

Introduction

- 1 La révolution de la donnée sera-elle la troisième étape de la révolution numérique après l'informatique puis l'Internet ? Le « Big data » est une mine de données dont l'exploitation consiste à sélectionner et croiser les données pertinentes en vue de les déployer dans la gestion des ressources humaines et dans les outils techniques marketing. Quant aux industries culturelles et créatives, les données personnelles revêtent une importance particulière, puisqu'elles permettent entre autres d'optimiser le temps de visite sur un site web ou sur un site réel, de comprendre les comportements et prédire les attentes des utilisateurs. Une façon pour les secteurs de la presse, du cinéma ou du jeu vidéo d'enrichir leur offre tout en optimisant les coûts.

Aperçu sur l'importance de la data dans les services documentaires

- 2 L'apport réel du numérique et du Big Data à l'accès à la connaissance a permis de prime abord l'évolution d'un accès arborescent et hiérarchique à la connaissance à un accès fragmentaire et multidimensionnel dans les SID. Le Big Data a permis aux services d'information documentaire de faciliter l'accès à l'information par le biais des moteurs de recherches et les systèmes de recommandation, ainsi que la visualisation, permettant une exploration optimisée des connaissances avec la possibilité de dévoiler des liens entre des domaines de connaissance pourtant auparavant éloignés. De même, une réflexion sur les

modèles de bibliothèque montre ce passage des modèles d'organisation et de classification des connaissances à des modèles de visualisation de ces organisations et, en dernier lieu, aux modèles de navigation dans les collections, virtuelles et physiques.

- 3 Un service d'information documentaire et ses (méta) données sont ainsi devenus un objet politique, soit le reflet de l'évolution des normes et des valeurs avec lesquelles elles interagissent continuellement dans son environnement au sens le plus large à travers les Humanités Numériques, qui ont voulu élaborer un cadre critique pour la visualisation des données et ont fait de la subjectivité pour cerner la connaissance de la bibliothèque et de son histoire. Il en est de même pour les algorithmes qui sont, d'une part, déployés à cet effet, et d'autre part, la visualisation repose largement sur ces algorithmes.
- 4 En effet, à partir des données structurées et non structurées des SID, plusieurs services ont vu les jours et certains produits ont été améliorés et optimisés de manière à satisfaire les usagers :

Une meilleure interopérabilité des données

- 5 Les données des SID sont actuellement enregistrées sous forme de notices en format MARC. Il s'agit d'un standard métier établi dans le monde entier depuis plus de 50 ans. Toutefois, plusieurs autres acteurs de l'information utilisent des formats XML ou sur d'autres formats plus modernes.
- 6 Les LOD se basent par contre sur RDF (Ressource Description Framework), qui est un standard du web flexible et extensible. Ainsi, des données bibliographiques de structures très diverses (MARC, Dublin Core, EAD, etc.) peuvent être converties en un format unique basé sur RDF, qui aide à effectuer une recherche fédérée bien plus performante que lors de l'interrogation de plusieurs bases de données séparées. Par exemple, la Bibliothèque nationale de France a introduit un service data.bnf.fr (<http://data.bnf.fr/>), qui permet puiser des résultats d'une requête de plusieurs sources simultanément (le catalogue des imprimés, la bibliothèque digitale et d'autres données non issues des catalogues traditionnels etc.)¹.

De meilleures possibilités de recherche

- 7 Les données, notamment celle structurées ont l'avantage d'être relatives à d'autres jeux de données de provenance externe, ce qui permet d'opérer requêtes plus pointues. D'ailleurs, c'est le cas des données géographiques à l'image des latitudes et longitudes pour la recherche, et la liaison entre deux personnes², comme en témoigne l'illustration qui suit :

Figure 1. Recherche de relation entre deux personnes sur Le portail finlandais Kulttuurisampo³

Person 1 Rubens, Peter Paul (flaamilainen taiteilija, 1577-1640)

Person 2 Miró, Joan (espanjalainen taiteilija, 1893-1983)

Rubens, Peter Paul (Flemish painter and draftsman, 1577-1640)
patron was
Louis XIV, King of France (French monarch, collector, and patron, 1638-1715)
knows
Louis XV, King of France (French ruler, patron, and collector, 1710-1774)
grandparent of
Charles X, King of France (French ruler, patron, and collector, 1757-1836)
patron of
Lhuillier, Charles (French painter and sculptor, born 1824, died 1898 or 1899)
teacher of
Dufy, Raoul (French painter, designer, and ceramicist, 1877-1953)
collaborated with
Llorens Artigas, Josep (Spanish ceramicist, 1892-1980)
collaborated with
Miró, Joan (Spanish painter, printmaker, and ceramicist, 1893-1983)

- 8 L'amélioration des possibilités de recherche peut élargir le rôle actuel du catalogue d'un dispositif de localisation vers un véritable outil de recherche.

Une meilleure sérendipité

- 9 Les données structurées et non structurées permettent aussi de donner à l'utilisateur un accès à des ressources complémentaires, comme des notices d'autorités d'autres bibliothèques ou des articles Wikipédia, comme le rapporte l'exemple ci-après :

Figure 2. Une seule interface pour les ressources variées du Centre Pompidou, sur Max Ernst⁴

MAX ERNST

Mots-clés : Modifier/ajouter RETOUR À LA RECHERCHE INITIALE

À propos de la personne ★ ✎ 🔗

+ lire la suite

À propos de la personne

SES ŒUVRES

SES DOCUMENTS

EST LE SUJET DE

LES ÉVÉNEMENTS

LES DOSSIERS PÉDAGOGIQUES

- 10 Ces applications sont sous format RDF, et aident l'utilisateur à trouver des informations qu'il ne cherchait pas dans leurs requêtes initiales.

Une meilleure visibilité des données sur le web

- 11 L'utilisation d'identifiants uniques et pérennes pour les données contribue à créer des relations manquantes entre les ressources, de manière à améliorer la visibilité des données des SID. A titre d'illustration, presque 80 % des visiteurs du service data.bnf.fr proviennent des moteurs de recherche, alors même que le catalogue traditionnel n'y est absolument pas visible⁵.

Matériel et méthodes

- 12 Pour répondre à la problématique soulevée dans le présent article, nous avons fait appel à la méthode descriptive à travers l'instrument de recherche du questionnaire dont le choix de l'échantillon s'explique par le fait qu'il permet la collecte de données auprès d'un grand nombre d'individus, à un coût relativement peu élevé, en termes de déplacements et de personnel.
- 13 Par ailleurs, il contribue à la fiabilité en favorisant une meilleure uniformité, et diminue aussi du degré du biais, puisque l'intervieweur ne peut pas influencer sur les réponses de façon intentionnelle ou non.
- 14 Pour atteindre les objectifs de cette recherche, nous avons administré un questionnaire auprès d'une population de 3 responsables de service d'information documentaire, destinés principalement aux chercheurs et aux étudiants, malgré le niveau de confidentialité élevé, il s'agit des entités documentaires marocaines ci-après : l'Institut Marocain de l'Information Scientifique et Technique (IMIST), la Bibliothèque Nationale du Royaume du Maroc (BNRM), et la Bibliothèque de l'Institut National des Statistiques et d'Economie Appliquée et qui fournissent tous des produits et des services documentaires orientés vers la recherche et le cadrage des activités d'enseignement pour une population composée principalement des professeurs universitaire, des chercheurs et des étudiants.

RÉSULTATS ET DISCUSSIONS

Les types de documents

- 15 Les collections des services d'informations documentaires ayant répondu à notre questionnaire regorgent d'une myriade de ressources diverses, allant des monographies et périodiques jusqu'aux ressources électroniques et multimédia, avec une prédominance de la première catégorie, comme le relate la figure suivante :

Figure 3. Les types de documents auxquels donnent accès le SID

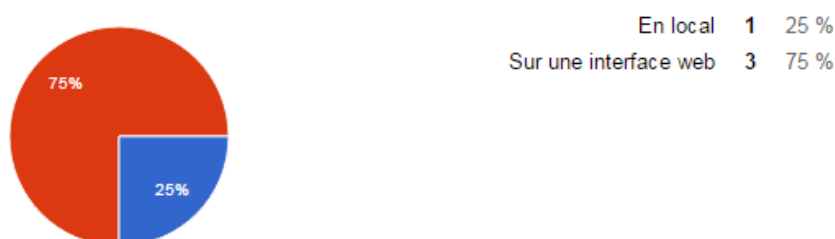


- 16 Les résultats indiqués dans la figure ci-dessus montrent que ce passage des SID au Maroc, du papier au numérique a déjà commencé avec le vent des TIC qui a soufflé sur la sphère documentaire, en se penchant sur des projets de numérisation et en élargissant les acquisitions de façon à ce qu'elles comportent aussi d'autres supports d'information que le papier.

L'accès au catalogue en ligne

- 17 Les services d'information documentaires cibles de notre étude utilisent un catalogues en ligne pour leur permettre d'accéder aux collections et de fouiller dans les ressources disponibles, à travers l'interface utilisateur, composante d'une plateforme de gestion intégré d'une bibliothèque, à savoir un catalogue en ligne public. Il est à signaler, que tous les responsables interrogés ont déclaré que le logiciel PMB est leur Systèmes Intégré de Gestion de Bibliothèque utilisés. Cette question du catalogue en ligne est explicitée par la figure qui suit :

- 18 Figure 4. L'accès au catalogue en ligne du SID⁶

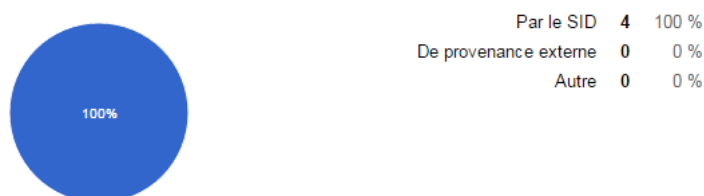


- 19 Quant à la nature du catalogue en ligne, il s'avère que la majorité des SID questionnés, soit 75 %, ont tendance à l'héberger sur le web, et ainsi il est accessible depuis le navigateur de l'utilisateur indépendamment du temps et de l'espace, sans les restrictions liées à un réseau local et qui représentent 25 %, ce qui revient à dire que les services d'information documentaire ont plus de chances de puiser des données surtout sur les usagers et leurs comportements informationnels, dans la mesure où la bonne part des catalogues en ligne hébergés permettent aux entités documentaires de colliger plus d'informations sur tout type d'usagers (réels et potentiels) comparé aux réseaux locaux mis sur place.

La provenance des données du SID

- 20 Les services d'information documentaires interrogés dans notre enquête se penchent à produire et extraire les données en interne avec un taux de 100 %, plutôt que de les avoir de l'extérieur :

Figure 5. Les sources de données du SID⁷



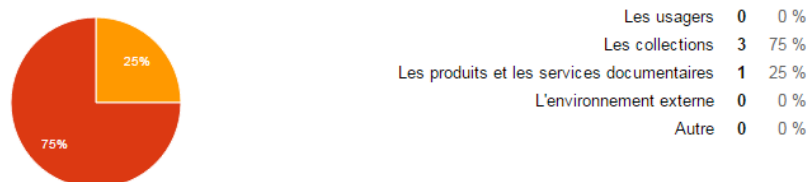
- 21 Nous concluons que les professionnels de l'information et de la documentation se cantonnent dans les données produites en interne, sans utilisation des données externes fiables et crédibles et qui peuvent intéresser l'activité du SID (rapports et études gouvernementaux, livres blancs, données du prêt interbibliothèques etc.)
- 22 Les données internes et externes produites par le SID :
- 23 Les services d'information documentaires produisent régulièrement des données en interne, comme ils les reçoivent de provenance externe dans le cadre de ses activités quotidiennes couvrant toute la chaîne documentaire (l'acquisition, le traitement et la diffusion) :
- Données sur les ressources documentaires ;
 - Ressources matérielles ;
 - Produits et services ;
 - Conventions et partenariats avec des organismes du même secteur ;
 - Liste des documents disponibles dans le SID ;
 - Informations sur les usagers et description des collections ;
 - Notices bibliographiques.
- 24 Une première observation sur le type de données produites par les SID, il est évident qu'elles sont basiques et hétérogènes, c'est-à-dire des données issues de ce qui est produit ou généré automatiquement, au point que les produits ou les services sont considérés comme des données, alors que la question a porté sur les données internes et externes sur les collections et les activités du SID.
- 25 Une autre remarque et non la moindre concerne la quasi-absence des données sur les usagers, qui sont la finalité et la raison d'être de tout système info-documentaire, sauf d'avoir été lancée le brièvement possible dans une seule réponse sur les quatre recueillies. Or, l'utilisateur peut être approché de son besoin jusqu'à son comportement et feed-back, voire le suivre continuellement pour pouvoir rencontrer sa satisfaction.

La nature données des SID :

- 26 Les données sur les SID couvrent les différents maillons de la chaîne documentaire et interviennent toujours en amont et en aval de chaque maillon à part entier pour se compléter, ainsi et sur six éléments présentées dans la question, les réponses des

professionnelles de l'information et de la documentation interrogés ont donné les résultats ci-dessous :

Figure 6. La nature des données du SID^a



- 27 Une lecture dans la figure ressort un constat frappant, nous remarquons bien que les données colligées par les SID cibles de notre enquête ont plus de relation avec les collections dans un premier lieu et les produits et les services documentaires dans un second temps, alors que les usagers sont omis, de même pour l'environnement externe qui est un champ normalement des usagers potentiels.

L'usage fait des données produites par le SID :

- 28 Les données produites ou reçues par les services d'information documentaire sont destinées à plusieurs usages et cumulées sur une période allant de 2 ans à 10 ans, pourtant les collections s'accroissent de 2 % à 5 %, voici les usages cités dans les réponses collectées :

- La recherche bibliographique et fulltext ;
- Mettre à disposition du public et des chercheurs le fonds documentaire ;
- Réunir la documentation existante en relation avec les sciences de l'information ;
- Encourager la recherche dans les domaines liés aux sciences de l'information ;
- Etablir des coopérations avec les institutions et organisations nationales et internationales dans le même domaine ;
- Mieux cerner les besoins ;
- Les données produites servent à orienter les usagers dans leur recherche d'information.

- 29 Nous déduisons de cette panoplie d'usages de données dans les SID que, malgré que les données évoquées en haut soient jugées insuffisantes surtout concernant les usagers, elles sont déployées pour les orienter et cerner leurs besoins, et leur acheminer de nouveaux services, sachant qu'avec les mutations technologiques aujourd'hui la période sur laquelle sont colligées les données semble longue (2 ans au moins), alors que ces données sont mises à jour chaque jour dans d'autres SID de par le monde.

La conservation des données des SID dans le temps :

- 30 Les services d'information documentaires cumulent les données dont ils disposent sur les années :

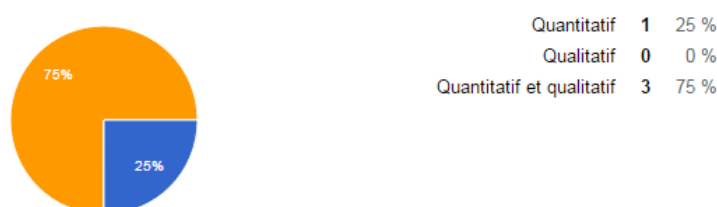
Figure 7. La conservation des données dans le SID⁹

- 31 Il est évident que, la majorité des services d'information documentaires tendent à conserver les données relatives au SID dans le temps pour déduire des tendances, établir des benchmarking, et faire des prévisions. Dans le temps, les données cumulées même au point d'une minute permettent d'avoir un socle solide de Data qui pourrait être exploité dans le sens d'aider les SID à améliorer leurs services aux usagers.

Le type de données collectées par les SID :

- 32 Les SID gèrent les données produites et reçues aussi bien d'ordre quantitatif que qualitatif suivant les proportions ci-après :

Figure 8. Le type de données dans le SID

Figure 9. La manière de collecte de données dans les SID¹⁰

- 33 L'analyse de la figure au-dessus ressort un équilibre en termes du type des données produites et reçues, de façon à ce que les SID font parler le quantitatif pour le rendre aussi qualitatif, comme elles sont colligées de manière directe et indirecte, ce qui permettrait à faire bon usage de ces données.

L'importance d'un système de recommandation pour les usagers d'un SID

- 34 Les services d'information documentaires interrogés classent l'importance voire la nécessité d'un système de recommandation pour faciliter l'accès aux ressources documentaire de manière plus personnalisée de la manière suivante :

Figure 10. La nécessité d'utilisation d'un système de recommandation dans les SID.



- 35 Il est clair que, la majorité des professionnels d'information et de documentation (soit 75 %) considèrent que le système de recommandation est requis pour aider les usagers à exploiter le fonds du SID, ce qui donne une idée sur l'adaptation moyenne des instruments d'accès et d'utilisation des ressources aux attentes des usagers, et leur niveau de satisfaction par voie de conséquence.

CONCLUSION

- 36 L'exploitation de la Data dans la spirale des industries culturelles d'une manière générale, et les services d'information documentaires plus particulièrement, permet d'identifier et d'appréhender les opportunités de création de valeur. Dans ce cadre, la Data aux SID aide à créer de nouveaux services, et améliore la qualité de services existants, tout en tenant compte de analyse et la visualisation de données en temps réel. L'objectif est ici d'améliorer les prises de décision et d'approfondir la connaissance d'un sujet (par exemple via des tableaux de bord ou du datamining).
- 37 Une entreprise agissant dans la sphère culturelle peut créer de la valeur en étant créatrice du flux de données, soit à dessein, soit en tant que produit secondaire d'une autre activité. Le développement des applications pour les dispositifs mobiles permet de plus en plus à ces entreprises de générer des flux de données en temps réel sur les usages et leurs usagers. Ces données peuvent être déployées dans le contexte des SID pour créer des services innovants à l'image des systèmes de recommandation ou encore des outils de découverte, en vue d'optimiser et personnaliser l'accès des utilisateurs aux ressources disponibles.

BIBLIOGRAPHIE

Alexander, Peter J., Peer-to-Peer File Sharing : The Case of the Music Recording Industry». *Review of Industrial Organization*. 2002, vol. 20, n° 2, p. 151-161.

Auray, Nicolas, La consommation en régime d'abondance. La confrontation aux offres culturelles dites illimitées. *Revue française de socio-économie*. 2011. vol. 2, n° 8, 2011, p. 85-102.

Beer, David. Making Friends with Jarvis Cocker : Music Culture in the Context of Web 2.0. *Cultural Sociology*. 2008, vol. 2, n° 2, p. 222-241.

Bisaillon, Jean-Robert, *Métadonnées et répertoire musical québécois : un essai de mobilisation des connaissances dans le nouvel environnement numérique*. Mémoire de maîtrise. Montréal : Institut national de la recherche scientifique, 2013.

Bouquillon, Philippe. *Les industries de la culture et de la communication. Les stratégies du capitalisme*. Presses Universitaires de Grenoble, 2008.

Bouquillon, Philipp. Industries, économie créatives et technologies d'information et de communication. *Tic & Société*. 2010, vol. 4, n° 2, p. 7-40.

Chantepie, Philippe et Alain, Le Diberder. *Révolution numérique et industries culturelles*. Paris : La Découverte, 2010, 111 pages.

Éric, Géorge. *À propos des stratégies industrielles des entreprises médiatiques, relecture de quelques travaux en économie politique de la communication*. Observatoire des mutations des industries culturelles, 2007.

Hugi, Jasmin. Le virage linked open data en bibliothèque : étude des pratiques, mise en œuvre, compétences des professionnels. *Revue Electronique Suisse des Sciences de l'Information (RESSI)*. 2014, n° .14, 2014.

KOP,R. The Challenges to Connectivist Learning on Open Online Networks : Learning Experiences during a Massive Open Online Course. *European Distance and E-learning Network annual Conference*. 2010, p. 4- 32.

MATO, D. All industries are cultural. *Cultural Studies*, 2009, 70 - 87P.

SENNOUNI, Amine., SBIHI, Boubker., GOUARTI, Rachid., BELGHIT, Imad., EL HAIL, Badr. The role of open data in strategic decision making in the telecommunication sector in Morocco. *International Journal of Research Studies in Science, Engineering and Technology (IJRSSET)*. 2015, vol. 2, Issue.6, Juin 2015, pp. 42-51.

NOTES

1. Bermes, Emmanuelle. *Le web sémantique en bibliothèque* [en ligne]. Ed. du Cercle de la librairie, 2011. Collection bibliothèques. Disponible sur : <<http://www.electrelaboutique.com/ProduitECL.aspx?ean=9782765414179>> (Consulté le 24 Octobre 2015).

2. Wenz, Romain. Linked Open Library Data in practice : lessons learned and opportunities for data. *SWIB 2012* [en ligne]. Cologne. 27 novembre 2012. Disponible sur : <http://swib.org/swib12/slides/Wenz_SWIB12_108.ppt> (Consulté le 20 Octobre 2015).

3. Hugi, Jasmin. Le virage linked open data en bibliothèque : étude des pratiques, mise en œuvre, compétences des professionnels. *Revue Electronique Suisse des Sciences de l'Information (RESSI)*. 2014, n° .14, 2014.

4. *Ibid.*, 3.

5. Gandon, Fabien ; Corby, Olivier. *Le web sémantique : Comment lier les données et les schémas sur le web ?*. Dunod, 2012.

6. Résultats de l'enquête de l'étude.

7. *Ibid.*, 5.

8. Résultats de l'enquête de l'étude.

9. *Ibid.*, 7.

10. Résultats de l'enquête de l'étude.

RÉSUMÉS

Appréhender la place de la Data dans les services d'information documentaires (SID) a été abordé en portant un aperçu sur la littérature traitant de la question. Toutefois l'usage de la data dans les SID doit aussi être ancré dans la pratique quotidienne des structures info-documentaires dans le contexte marocain, en tentant de mettre à la loupe l'usage que fait les services d'information documentaires des données qu'ils produisent ou ils reçoivent dans leur fonctionnement et l'accomplissement de leur mission. Ainsi, quatre responsables de bibliothèques et de centres de documentation ont été interrogés sur ces éléments à travers un questionnaire qui leur a été acheminé (un centre de documentation spécialisé, une bibliothèque publique, et deux bibliothèques universitaires).

Understand the role of Data in documentary information services (CIS) was discussed by an overview of the literature on the issue. However the use of data in the SID must also be anchored in the daily practice of info- documentary structures in the Moroccan context, attempting to shed light on the use made by the documentary information services data that they produce or they get in their operation and the fulfillment of their mission. Thus, four officials of libraries and documentation centers were asked about these elements through a questionnaire that was sent to them (a specialized documentation center, a public library, and two university libraries).

INDEX

Keywords : cultural and creative industries, information service, Data, content, recommender system

Mots-clés : industries culturelles et créatives, service d'information documentaire, Data, contenus, système de recommandation

AUTEUR

SENNOUNI AMINE

Amine Sennouni, 26 ans, Doctorant en troisième année de thèse au Conservatoire National des Arts et Métiers (CNAM- Paris) en co-encadrement international avec l'École des Sciences de l'Information (ESI-Rabat), dans la spécialité des sciences de l'information, de la data et de la communication (SIC), au laboratoire Dicen-IDF (Dispositifs d'Information et de Communication à l'Ère Numérique – Paris, Ile-de-France). Courriel : sennouni.amine@gmail.com