



U S T H B *République Algérienne Démocratique et Populaire* U S T H B

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université des Sciences et de la Technologie Houari Boumediene

Faculté d'Electronique et d'Informatique
Département Informatique

Mémoire de Master

Spécialité:

Ingénierie Logiciel (IL)

Thème :

**Conception et développement d'une plateforme de
recommandation d'articles scientifiques.**

Sujet Proposé par :

M (e)..... :

Soutenu le :.../.../....

Présenté par :

Devant le jury composé de:

M..... Président (e)

M..... Membre

Binôme N° : 005 / 2022

Remerciements

Table des matières

Table des matières.....	3
Liste des figures.....	4
Liste des tableaux.....	6
Liste des abréviations.....	7
Introduction générale	8
Plan du mémoire	9
Etat de l'art	10

Liste des figures

Figure 1 : Capture d'écran du navigateur Mosaic.

Figure 2 : Interface du réseau SixDegrees.

Figure 3 : Interface du réseau LinkedIn.

Figure 4 : Chronologie des médias sociaux.

Figure 5 : Représentation graphique d'un réseau social.

Figure 6 : Evolution du nombre d'utilisateurs de réseaux sociaux dans le monde.

Figure 7 : Interface du réseau Twitter.

Figure 8 : Interface du réseau Viadeo.

Figure 9 : Interface ResearchGate.

Figure 10 : Nouveaux utilisateurs d'Academia par mois.

Figure 11 : Interface Academia.edu.

Figure 12 : Interface MyScienceWork.

Liste des tableaux

Tableau 1:

Liste des abréviations

SR	Système de Recommandation.
CERIST	Centre de Recherche sur L'information Scientifique et Technique.
CACHE	Chicago Area Computer Hobbyist's Exchange
NCSA	National Center for Supercomputing Applications
WWW	World Wide Web

Introduction générale

Avec la montée fulgurante de la technologie et la facilité d'accès à Internet, les utilisateurs sont submergés par un large éventail de choix disponibles et une quantité considérable d'informations [1]. En effet, nous vivons privilégiés dans un monde infiniment riche en information, et dans lequel la plupart, si ce n'est la totalité de l'information dont nous avons besoin est à l'appui d'une touche et est prête à être exploitée. L'essor du Web, conforté par le rapide développement de nouvelles technologies d'information et de communication a conduit à la production d'un volume d'information sans précédent, la quantité d'information manipulée dans le monde est n'as jamais été aussi élevée et sa gestion aussi laborieuse. Ce flux informationnel est ingérable avec les moyens classiques, ainsi il devient nécessaire d'avoir accès à des outils et des techniques efficaces pour filtrer les données avant de les transmettre aux utilisateurs.

Suite à des années de recherches actives et fort fructueuses, les systèmes de recommandations (**SR**) ont vu le jour. On est maintenant capables de filtrer les données et ainsi fournir à l'utilisateur lambda des choix pertinents sur ce qu'il aime lire, acheter, regarder ou manger, etc. cette solution a été implémenté dans différentes plateformes destinées à différents publics, dont : **la communauté scientifique**, qui a vu naître des plateformes académiques visant la recommandation d'articles scientifiques et répondant à leurs besoins de recherche.

Notre projet de master proposé par **CERIST**, consiste à la conception et développement d'un noyau de réseau social de recommandation d'articles scientifiques et l'implémentation de ses différentes fonctionnalités. Il aura pour mission d'encourager le partage des connaissances, d'améliorer la visibilité des chercheurs et de leurs travaux et aussi d'offrir à ses derniers une expérience personnalisée selon leurs domaines d'étude et de recherche.

Plan du mémoire

Ce présent manuscrit détaille les étapes que nous avons suivies pour arriver à terme de notre projet. Il va s'articuler comme suit :

- **Chapitre 1 :** Le premier chapitre de ce mémoire sera consacré à introduire des définitions et généralités sur le domaine d'étude, un état de l'art des plateformes de recommandation en ligne existantes ainsi que la problématique soulevée.
- **Chapitre 2 :** Ce chapitre comportera la phase d'analyse, qui détaillera la modélisation conceptuelle de notre solution pour intégrer les fonctionnalités de base (inscription, stockage, communication, suivre, etc.) ainsi que la recommandation.
- **Chapitre 3 :** La phase d'implémentation, qui montrera le travail réalisé ainsi que l'environnement et les outils utilisés afin d'atteindre ce résultat.

Nous achèverons ce manuscrit par une conclusion qui englobera les objectifs atteints, et les améliorations envisageables.

Chapitre 1

Etat de l'art

1. Etat de l'art

1.1. Introduction

Dans ce chapitre, nous commencerons par un bref historique relatant l'apparition des réseaux sociaux, nous définirons notre domaine de recherche puis présenterons les plateformes de recommandation scientifique en ligne déjà existante. Nous détaillerons par la suite la problématique, pour finalement proposer une solution relative aux problèmes mentionnés.

1.2. Qu'est-ce qu'un réseau social académique ?

1.2.1. Historique

Bien que le terme « réseau social » semble être un terme relativement récent, évoquer son histoire n'est pas exagéré puisque pour comprendre son origine on doit élargir notre vue vers la fin des années 1970. Nous allons ci-dessous parcourir l'histoire de l'apparition et diffusion des réseaux sociaux les plus marquants [2] :

- Durant l'hiver 1978, Ward Christensen et Randy Suess, Deux passionnés d'informatique et membres du Chicago Area Computer Hobbyist's Exchange (**CACHE**) ont cernés un problème au sein de leur club : les membres de ce dernier s'échangent fréquemment des bulletins, des notes et des informations sur leurs recherches. Christensen et Suess ont eu l'idée de concevoir une solution en ligne pour l'archivage et le partage de ses bulletins, et c'est ainsi que le premier Bulletin Board System (**BBS**) a vu le jour : Le Hobbyist Computerized Bulletin Board. Ce fut le premier site de partage et d'échange d'informations par voie informatique [3].
- En 1993, des étudiants du National Center for Supercomputing Applications (**NCSA**) d'Illinois développent *Mosaic*, le premier navigateur web graphique permettant d'afficher le World Wide Web (**WWW**) et feuilleter les informations [4].

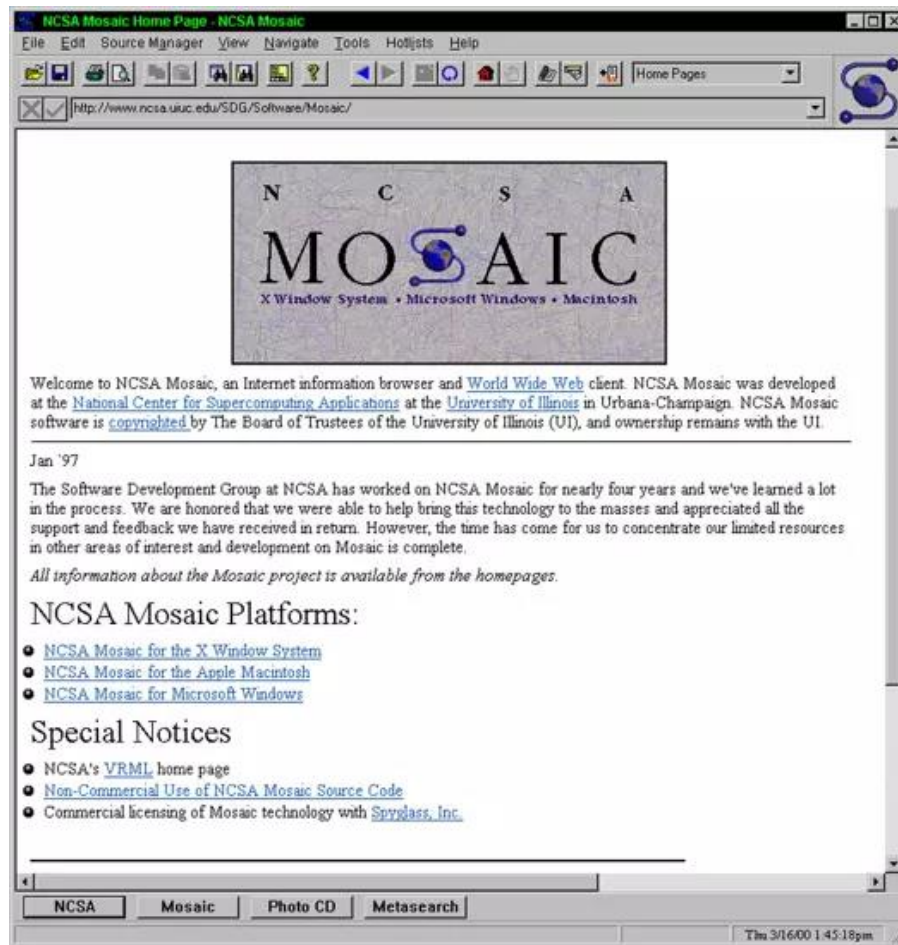


Figure 1 : Capture d'écran du navigateur Mosaic.

- En 1997 naît le tout premier réseau social original tel qu'on pourrait l'imaginer aujourd'hui, *SixDegrees* à la particularité d'avoir des pages de profil et la possibilité de se connecter entre elles. Il proposait aux internautes de créer leur profil, d'entrer en relation avec leurs proches, amis et familles, et publier des informations, envoyer des messages ainsi que recruter de nouveaux membres ce qui était révolutionnaire à l'époque.

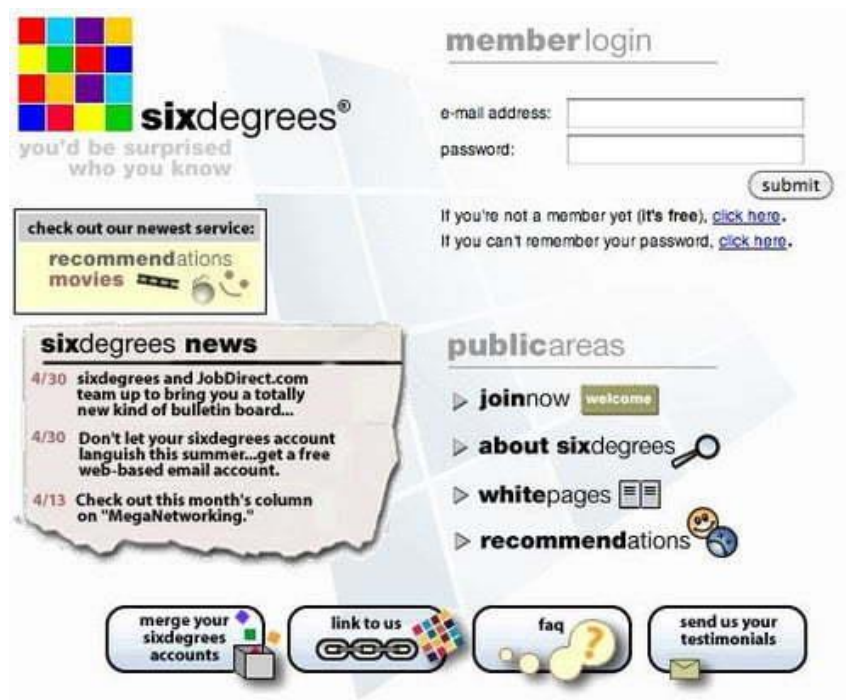


Figure 2 : Interface du réseau SixDegrees.

- L'année 1999, nous voyons le début de ce qui deviendrait finalement « le flux » de contenu avec l'introduction de *LiveJournal*, les utilisateurs pouvaient dorénavant poster des mises à jour et s'exprimer, permettant ainsi à leurs connexions de les suivre.
- Fondée en 2002 par Jonathan Abrams, *Friendster* a été le premier à autoriser les commentaires de profil et à faciliter la recherche et la connexion avec plus de personnes du monde entier.
- Le premier réseau social professionnel a vu le jour en 2003 par Reid Hoffman. *LinkedIn* est le plus ancien réseau social moderne, son objectif est de se démarquer en mettant en avant ses compétences et simplifier la mise en relation et en contact des professionnels dans le but de se constituer un réseau et de l'élargir.

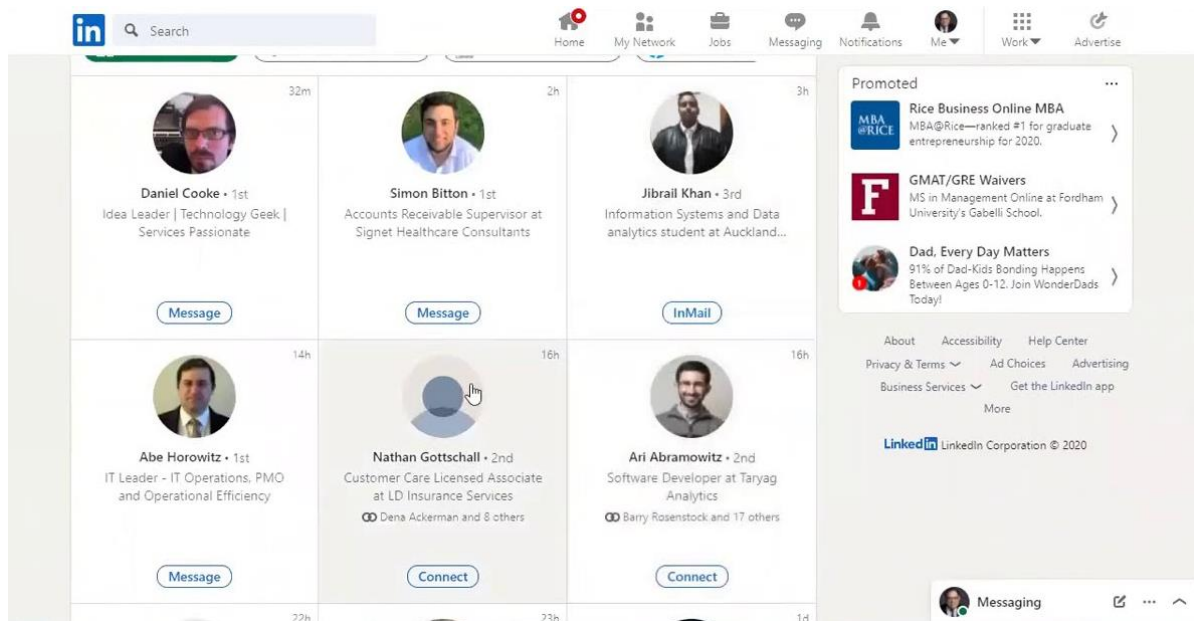


Figure 3 : Interface du réseau LinkedIn.

- C'est en 2004 que *Facebook* voit le jour, au main de Mark Zuckerberg et ses collègues de l'université d'Harvard. Son succès reposait sur l'évolution des messages et des commentaires figés proposés par les réseaux sociaux de l'époque.
- Suite à *Facebook* de nombreux réseaux sociaux sont nés, on en décompte : *Youtube* en 2005 orienté vers le partage de contenu vidéos, *Twitter* en 2006 qui a introduit le fil d'actualité, les tags, les hashtags et les réponses, ainsi que *Google+* en 2011 qui permettait principalement le chat vidéo entre utilisateurs.
- En 2008, avec la venue d'un autre type de réseau social tel que *ResearchGate* et *Academia.edu* un nouveau concept est introduit : le réseautage social pour chercheurs et scientifiques. Nous en parlerons plus en détails dans les prochains chapitres de ce présent manuscrit.

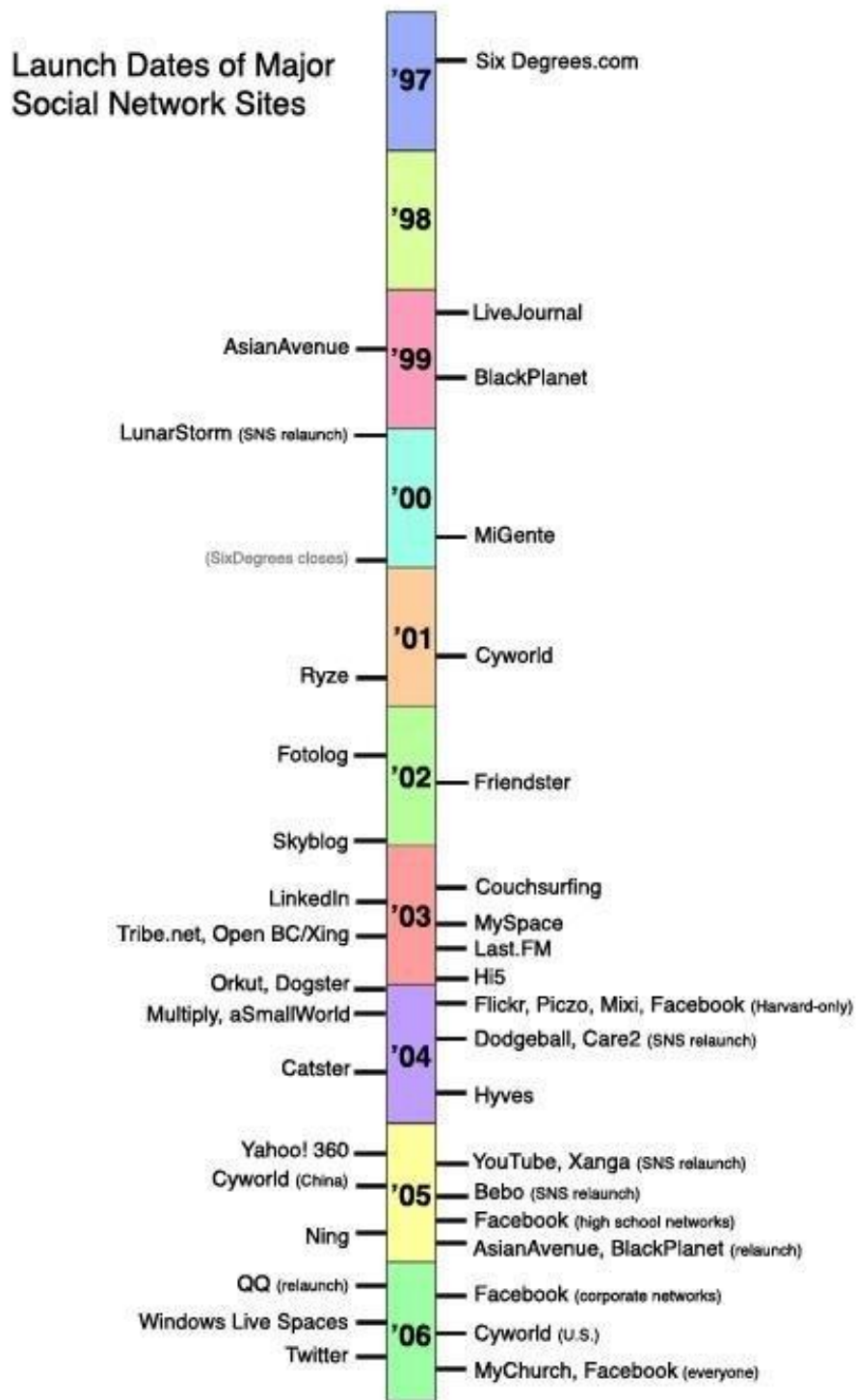


Figure 4 : Chronologie des médias sociaux.

En raison de la présence constante des réseaux sociaux dans la vie des individus, et le très grand impact qu'ils possèdent sur la vie sociale des internautes, on remarque une évolution sans cesse de leur nombre d'utilisateurs. En effet, le nombre d'utilisateurs actifs des réseaux sociaux s'élève à 4,55 milliards. La plupart de ses utilisateurs utilisent les réseaux sociaux en recherche de contenu divertissant, pour passer le temps et/ou rester en contact avec leurs proches. Aussi, 40,4% des socionauts (utilisateurs de réseaux sociaux) les utilisent pour un usage professionnel (*Statista 2021*) [6].

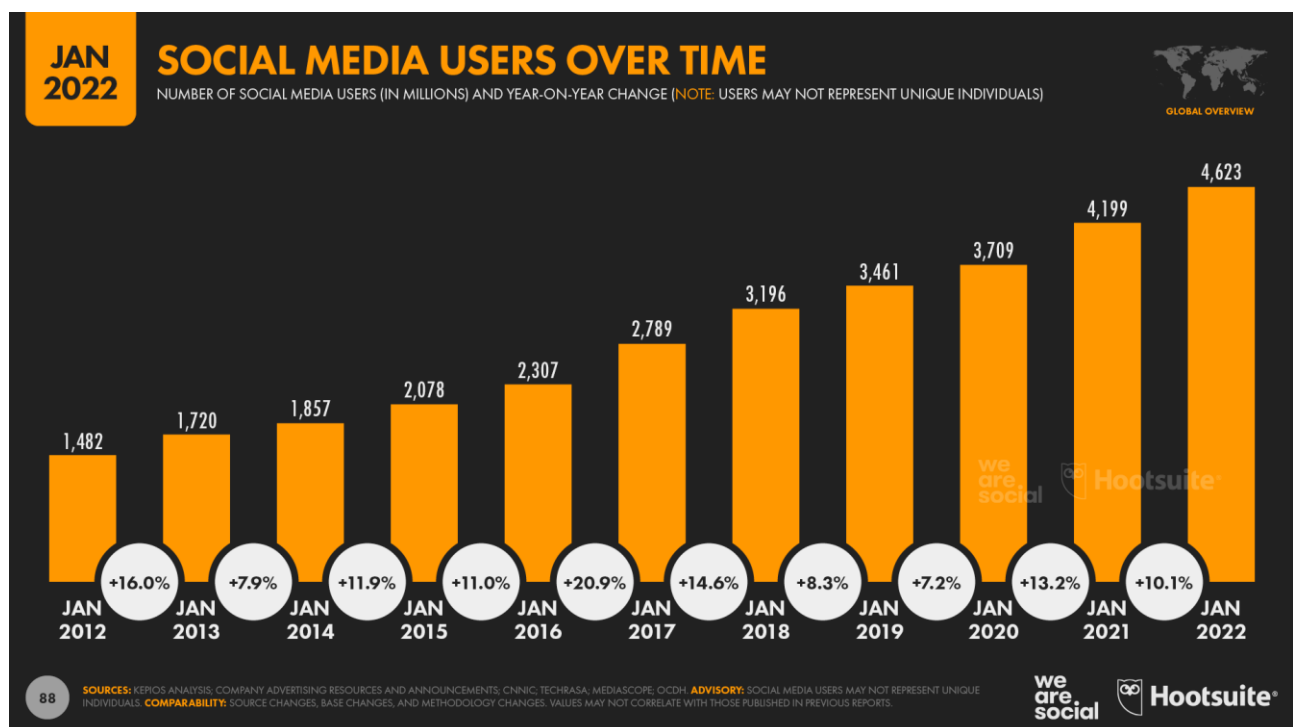


Figure 6 : Evolution du nombre d'utilisateurs de réseaux sociaux dans le monde.

1.2.3. Types de réseaux sociaux

Il est à noter qu'il existe plusieurs types de réseaux sociaux organisés selon plusieurs critères :

- La tranche d'âge ciblée.
- Le type de publication : Article, photographie ou vidéographie tel *Youtube*.
- L'objectif du réseau social : *LinkedIn* par exemple à pour but de permettre à l'utilisateur d'élargir ses opportunités professionnelles.

Suivant ses critères, nous distinguons trois types de réseaux sociaux chacun ciblant une population différentes et offrant des services différents : ceux destinés au grand public, ceux à usage professionnel reposant sur le partage de connaissances, et ceux destinés à la communauté scientifique [7].

A. Réseaux sociaux grand public

Ils ciblent toute la population indépendamment de la tranche d'âge. Ils ont pour principal objectif de divertir l'utilisateur, favoriser les relations sociales et partager les informations. On cite comme exemple les plus connus tels que *Facebook* et *Twitter*.

Ils offrent à l'utilisateur les fonctionnalités suivantes :

- Etablir un réseau regroupant les personnes aux intérêts communs (passions, activités, profession, etc.).
- Rapprocher les distances en gardons en contact l'utilisateur avec ses proches, ses amis, ses collègues, etc.
- Etre à jour en permanence quant à l'actualité à travers le monde entier.
- Favoriser la liberté d'expression : chaque utilisateur peut s'exprimer et échanger ses idées et ses passions au sein d'une communauté et avoir ainsi des retours sur une question ou un avis.

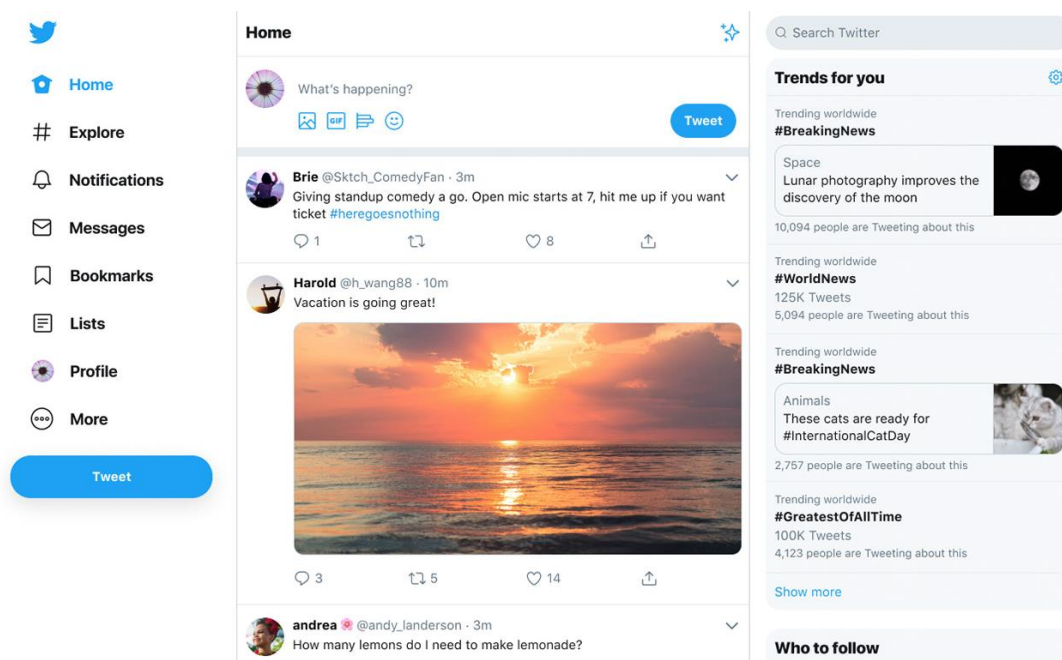


Figure 7 : Interface du réseau Twitter.

B. Réseaux sociaux professionnels

Ils sont destinés uniquement à des usages professionnels orienté vers la mise en valeur du profil professionnel à travers les compétences et expériences, ce qui est fort utile au employés et employeurs. Les plus connus sont *LinkedIn*, *Viadeo* et *Xing*.

Ils offrent à l'utilisateur les fonctionnalités suivantes :

- Faciliter le recrutement aux employeurs et au demandeurs d'emplois.
- Permet à l'utilisateur de remplir son profil tel un curriculum vitae (CV), et de joindre ses contacts à travers la messagerie afin de contrôler son image sur le web.
- Permettre à l'utilisateur de distinguer, parmi toute les identités virtuelles, celles qui seraient susceptibles de l'intéresser, et ce à travers la recommandation automatique.
- Permet aux entreprises de mettre en avant leurs activités et leurs actualités.

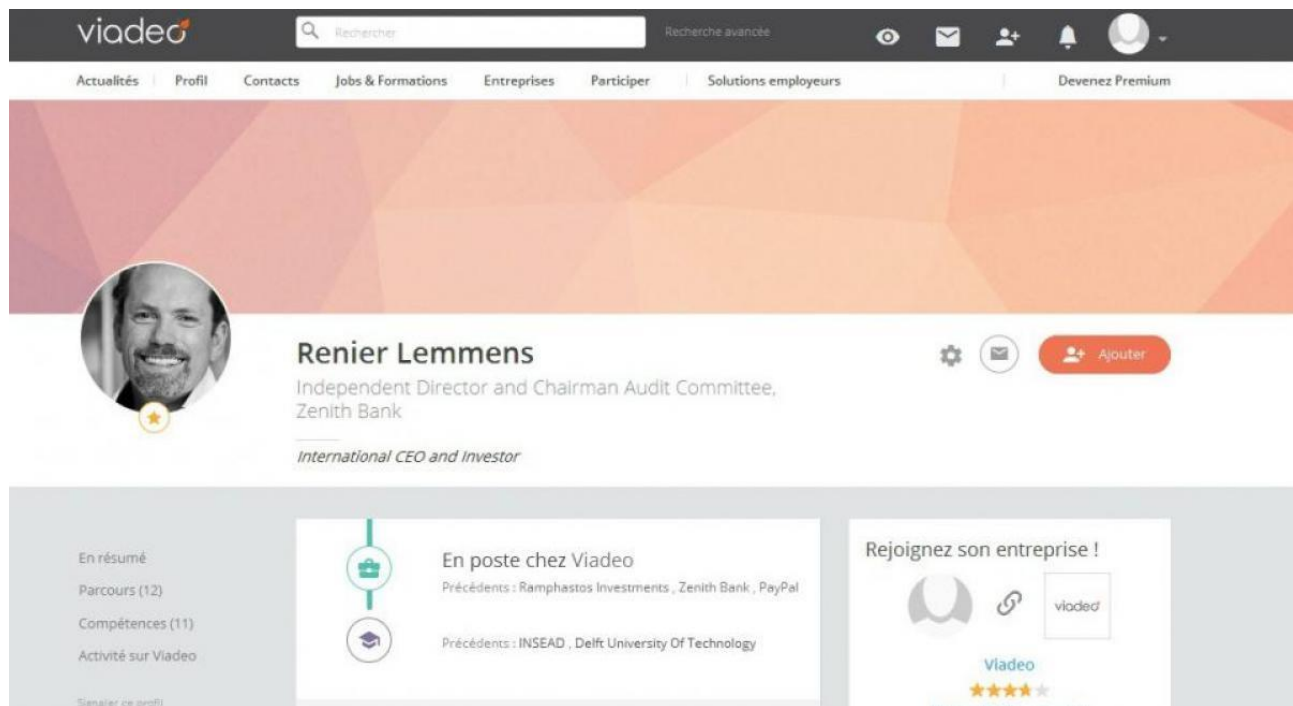


Figure 8 : Interface du réseau Viadeo.

C. Réseaux sociaux scientifique / académique

On arrive ainsi à répondre à la question posée au début de ce chapitre : Qu'est-ce qu'un réseau social académique ?

Un réseau social académique ou aussi appelé scientifique est, comme ses homologues professionnels et grand public, un espace de regroupant une communauté de même intérêt : la recherche. Ils sont devenus incontournables pour les chercheurs comme pour les institutions de recherche, ce sont des outils très efficaces et largement utilisés car ils permettent de :

- Répondre aux besoins des chercheurs en termes de collaborations et d'accès à l'information et à la connaissance, ce qui se traduit par une importante production scientifique.
- Améliorer la visibilité des chercheurs ainsi que celle des travaux dont ils ont besoin pour poursuivre et étendre leur activité en mettant à leurs dispositions des contenus pédagogiques.
- Encourager les utilisateurs au partage des connaissances par le dépôt de leurs travaux de recherche.
- Entrer en contact avec des pairs et échanger sur un sujet et/ou valider et annoter des articles après publication.
- Assurer une grande visibilité des universités et un meilleur classement à l'échelle internationale en évaluant les publications produites par chacune.

Ces derniers sont devenus indispensables à la communauté scientifiques grâce aux fonctionnalités qu'ils offrent et au temps précieux qu'ils fournissent.

1.3. Réseaux sociaux académiques existants

1.3.1. ResearchGate

ResearchGate est un site web proposant un service de réseautage social pour les chercheurs et scientifiques de toutes disciplines. Il a été lancé en 2008 par deux virologues Ijad Madisch et Sören Hofmayer, et un informaticien Horst Fickenscher [8], et ce, dans le but de remédier à la manière dont le contenu scientifique se partage de nos jours. Sa mission est de connecter le monde de la science et le rendre accessible à tout le monde à travers le monde [9].

Qu'on soit chercheur, étudiant, ingénieur ou technicien, *ResearchGate* permet une recherche scientifique sémantique dans une base de 118 millions de publications. Le site propose aussi un serveur de fichiers publics, un blog, des discussions méthodologiques et des groupes d'échanges. *ResearchGate* annonce avoir plus de 15 millions de chercheurs et scientifiques dans 192 pays.

Leur slogan : « *we exist to empower research* » (nous existons pour renforcer la recherche).

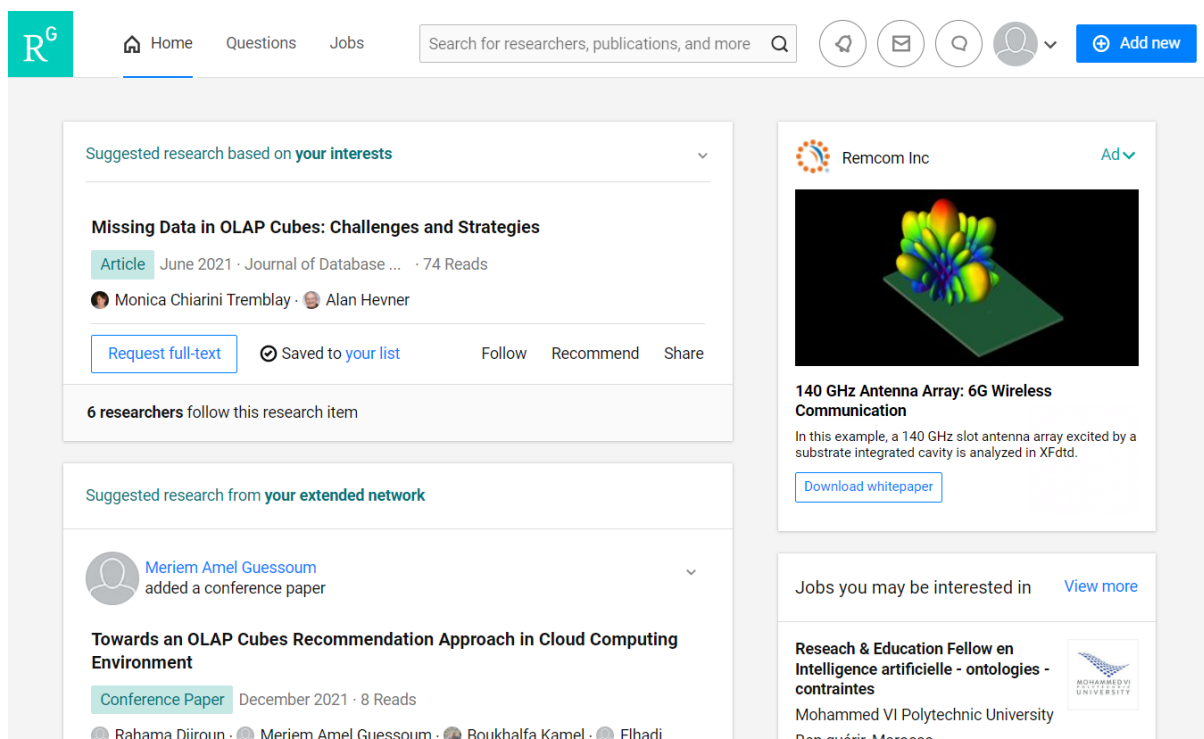


Figure 9 : Interface ResearchGate.

1.3.2. Academia.edu

Comme son homologue, *Academia* est une plateforme de partage de contenu pédagogique. Lancée en 2008 peu après *ResearchGate* par Richard Price, cette plateforme s'adresse aux chercheurs et aux universitaires et leur propose diverses fonctionnalités de réseautage social. Elle compte plus de 180 millions d'utilisateurs et ne cesse de croître exponentiellement [10].

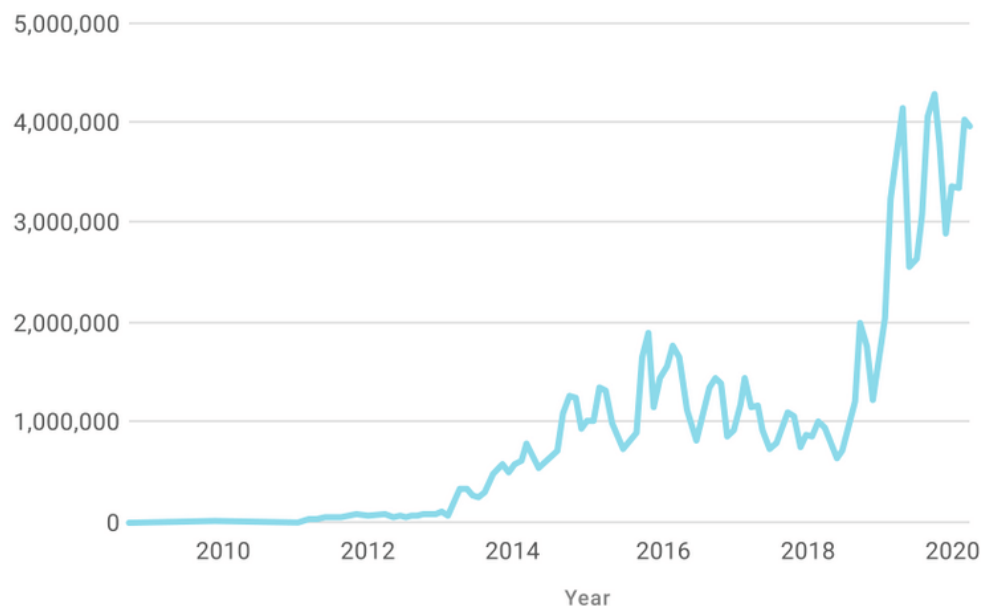


Figure 10 : Nouveaux utilisateurs d'Academia par mois.

La mission principale d'*Academia* est d'accélérer la recherche dans toutes les disciplines et la démocratiser, et ce, afin de trouver des solutions aux problèmes les plus importants de notre ère : le réchauffement climatique, certaines maladies mortelles, l'intelligence artificielle, etc.

Pour se faire elle se repose sur 4 piliers [10] :

- **Libre accès :** Le premier pilier afin d'atteindre sa mission est de s'assurer que chaque article scientifique ait un accès direct et gratuit sur internet. *Academia* compte plus de 22 millions d'articles et travaille afin de multiplier ce nombre.
- **Distribution :** le deuxième pilier d'*Academia* est de fournir le moyen le plus rapide et pertinents de distribution et recommandations d'articles dans le monde grâce à son algorithme qui arrive à recommander jusqu'à 20 millions d'articles par jour.

- **Peer review** : Le troisième pilier est de s'assurer de la fiabilité des articles publiés en son sein. La plupart des utilisateurs ne sont pas des chercheurs ou des experts, d'où l'importance d'implémenter une solution afin de garantir aux lecteurs la validité de chaque article. *Academia* utilise un système nommé PaperRank qui agit au niveau de chaque article.
- **Formats** : Le dernier pilier repose sur le format de l'information partagée. En effet, il existe des instances d'information qui ne peuvent être traduites en document PDF (data-sets, code, vidéos, etc.). *Academia* a lancé la fonctionnalité « Summaries » qui en entraînant un algorithme de Machine Learning, a pu fournir pour chaque document parmi les 22 millions présents sur le site, un résumé ne dépassant pas 700 mots, offrant ainsi aux utilisateurs à qui le temps est précieux, un moyen d'avoir une vue globale sur chaque document.

The screenshot displays the Academia.edu homepage. At the top, there is a search bar and a navigation menu with icons for Mentions, Letters, Analytics, Upload, and Tools. A timer shows 2 hours, 5 minutes, and 0 seconds. A 'Road' button is also visible. Below the navigation, the 'Software Engineering' section is highlighted, showing 815,432 followers and a 'Follow' button. The main content area lists several papers, including 'Reusing empirical knowledge during cloud computing adoption', 'Strategies for Improving MDA-Based Development Processes', and 'Criteria-Based Evaluation Framework for Service-Oriented Methodologies'. Each paper entry includes a 'Save to Library' button, a 'Download' button, the author's name, view count, and a brief abstract. A 'Related Topics' sidebar on the right lists various fields like Computer Science, Empirical Software Engineering, and Software Testing, each with a 'Follow' button.

Figure 11 : Interface Academia.edu.

1.3.3. MyScienceWork

Fondé en 2010 par Virginie SIMON et Tristan DAVAILLE, *MyScienceWork* sert la communauté scientifique internationale en promouvant le libre accès aux publications scientifiques et la diffusion illimitée de l'information en partageons plus de 90 millions de publications scientifiques. Leur but est de construire un futur proche où le temps de recherche est optimisé et la production scientifique est à son apogée.

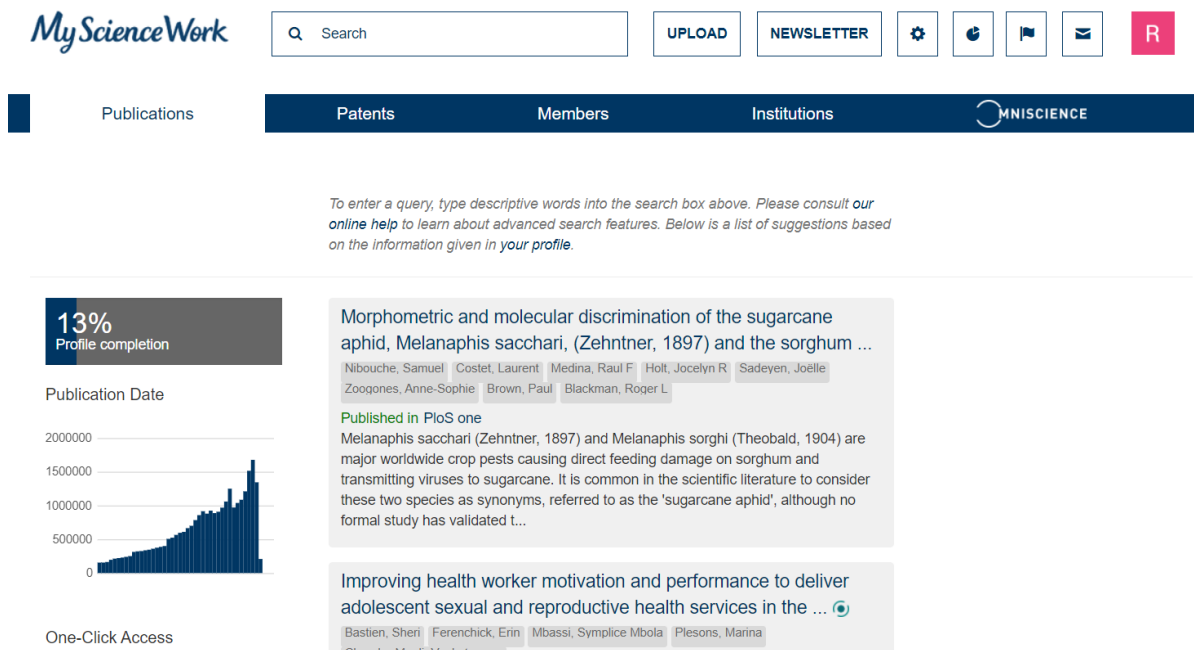


Figure 12 : Interface MyScienceWork.

Il existe un grand nombre d'autres réseaux sociaux scientifiques moins utilisés, tels : *Social Sciences Research Network*, *Nature Network*, *Mendeley*, *etc.* mais qui ont les mêmes buts et travaillent vers l'accomplissement de la même mission : **DEMOCRATISEZ LE SAVOIR.**

1.4. Système de recommandation

1.4.1. Introduction

Avec l'immense masse d'informations disponibles de nos jours sur Internet et le besoin fondamental de l'utilisateur en terme de temps, il est devenu nécessaire de concevoir des mécanismes et développer des outils qui permettent à ces derniers d'accéder à ce qui les intéresse le plus rapidement possible. Afin de faire face à ce problème de surcharge en informations, les chercheurs ont conçus les systèmes de recommandations.

La recommandation peut être comparée à un dialogue entre une personne experte d'un domaine précis et l'autre désireuse d'acquérir des informations dans ce domaine. Plus concrètement, un bibliothécaire va pouvoir, en fonction des goûts d'un de ses clients, proposer une liste d'ouvrages à ce dernier qui ne sera autre qu'une recommandation au sens des systèmes de recommandation [11].

1.4.2. Définition

Robin BURKE, président du département des sciences et informations de l'université de Colorado définit les systèmes de recommandation comme étant :

« Des systèmes capables de fournir des recommandations personnalisées permettant de guider l'utilisateur vers des ressources intéressantes et utiles au sein d'un espace de données important. »

Ils ont pour objectif de fournir à l'utilisateur des ressources pertinentes en fonction de ses préférences et ainsi, réduire grandement le temps de recherche.

Les systèmes de recommandations sont de nos jours présents et implémentés dans toute plateformes offrant du contenu quels que soit son type : *YouTube* utilise un système de recommandations afin d'afficher à l'utilisateur des vidéos similaire à sa recherche et/ou ses visionnages, *Facebook* le long du fil d'actualité, les sites e-commerce afin de promouvoir des articles que l'utilisateur pourrait ajouter à son panier, etc.

1.4.3. Les Systèmes de recommandations dans les réseaux sociaux académiques

Dans le cas d'un réseau social académique, un système de recommandation est implémenté afin de pour plusieurs raisons, parmi eux :

- Recommander à l'utilisateur de profession scientifique (qui souvent est très occupé) des articles et revues susceptible de l'intéresser suivant certaines références qui le caractérisent.
- Lui faciliter la tâche d'accéder, partager et diffuser l'information (article, revues, ouvrage, etc.) vers les parties intéressées.
- Promouvoir les articles qui ont eu le plus de retours et de réactions positives.
- Offrir à l'utilisateur des suggestions auxquelles il n'aurait pas spontanément prêtés attention (qu'il n'aurait pas pensé à rechercher).
- Recenser d'une simple façon les données statistiques scientifiques (nombre d'articles scientifiques publiées, les chercheurs les plus actifs, etc.).

1.5. Problématique

Comme décrit précédemment, dans le contexte actuel de déluge de données et d'informations, les méthodes classiques de recherche d'articles se montrent très couteuse en terme de temps, qui est une ressource fort précieuse chez la communauté scientifique.

Pour ce, notre présent travail consiste à concevoir et développer un noyau de réseau social académique comportant la recommandation d'articles scientifiques. Afin de d'offrir au chercheur une utilisation aisée, minimiser le coût horaire et maximiser le travail établi, cette plateforme devra permettre :

- L'inscription et l'authentification, de sorte que chaque utilisateur ait son propre compte qui regroupera ses informations, notamment sa fonction et son domaine de recherches.
- La publication d'article scientifique, leur stockage et partage.
- Le suivie de certains auteurs et/ou domaines intéressant pour ne manquer aucune nouvelle publication.
- La sauvegarde sur le compte et/ou téléchargement des articles souhaités.
- La recommandation d'articles scientifique selon les préférences de l'utilisateurs.

1.6. Conclusion

Ce chapitre a été consacré à la définition du domaine de recherche et la problématique de surcharge informationnel à laquelle fait face l'internaute en générale et la communauté scientifique en particulier. Nous avons parcouru les dernières décennies afin de comprendre l'origine des réseaux sociaux, leurs types et avancées, nous nous sommes intéressée à un type en particulier : les réseaux sociaux académique, que nous avons défini, puis relaté les plateformes scientifiques les plus utilisées en notre temps. Nous avons ensuite, introduit les systèmes de recommandation et la nécessité de les implémenter dans toute plateformes offrant du contenu.

Nous détaillerons la modélisation conceptuelle de la solution proposée dans le chapitre qui suit.

Chapitre 2

Analyse conceptuelle

2. Analyse conceptuelle

2.1. Introduction

Bibliographie

- [1] « User Modeling and User-Adapted Interaction 10 » Daniel BILLSUS et Michael PAZZANI.
Available :
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.13.4426&rep=rep1&type=pdf>
- [2] « Social Media's Dial-Up Ancestor: The Bulletin Board System » Kevin DRISCOLL.
Available : <https://spectrum.ieee.org/social-medias-dialup-ancestor-the-bulletin-board-system>
- [3] « Mosaic : le premier navigateur était lancé il y a 20 ans » Antoine CROCHET-DAMAIS.
Available : <https://www.journaldunet.com/web-tech/developpeur/1121615-mosaic-le-premier-navigateur-etait-lance-il-y-a-20-ans/>
- [4] « Histoire des réseaux sociaux et de leurs innovations » Gwendoline ENTE.
Available : <https://www.agence90.fr/chronologie-innovations-reseaux-sociaux/>
- [5] « Définir et analyser les réseaux sociaux » Michel FORSE.
Available : <https://www.cairn.info/revue-informations-sociales-2008-3-page-10.htm>
- [6] « Les chiffres 2021 des réseaux sociaux » Florian VEDOVATI.
Available : <https://www.onlyso.fr/les-chiffres-2021-des-reseaux-sociaux.html>
- [7] « Les réseaux sociaux – BTS SIO » J. HAROLD.
Available : <http://reseauxlapie.canalblog.com/>
- [8] « ResearchGate introduces RG Score, Klout for boffins » Jamillah KNIGHT.
Available : <https://thenextweb.com/news/researchgate-introduces-rg-score-like-klout-boffins>
- [9] « About us » Site officiel ResearchGate.
Available : <https://www.researchgate.net/about>
- [10] « The story of Academia.edu » Richard Price.
Available : https://www.academia.edu/43013575/The_Story_of_Academia_edu

[11] « Etat de l'art sur les Systèmes de Recommandation» Nicolas BECHET.

Available : <https://docplayer.fr/1209466-Etat-de-l-art-sur-les-systemes-de-recommandation.html>

