

Le traitement des données navigationnelles dans un contexte d'analyse de pratiques informationnelles

Safa Khezami, Valérie Larroche

▶ **To cite this version:**

Safa Khezami, Valérie Larroche. Le traitement des données navigationnelles dans un contexte d'analyse de pratiques informationnelles. Colloque Éducation 4.1 ! Distances, médiations des savoirs et des formations, Jan 2019, Poitiers, France. hal-02148082

HAL Id: hal-02148082
https://hal.science/hal-02148082

Submitted on 11 Jun 2020

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.

Le traitement des données navigationnelles dans un contexte d'analyse de pratiques informationnelles

Safa Khezami*, Valérie Larroche**

* Safa.khezami@univ-lyon3.fr, ELICO, Université Jean Moulin Lyon 3

** Valerie.larroche@enssib.fr, ELICO, ENSSIB

Résumé :

Cet article répond à la question suivante : Comment analyser l'influence du dispositif institutionnel d'apprentissage sur les pratiques numériques informationnelles des étudiants par le biais de traces navigationnelles ?

Le dispositif institutionnel d'apprentissage de l'IUT Jean Moulin Lyon 3 est constitué d'un centre de ressources, d'une bibliothèque numérique, d'un Intranet, d'enseignements liés à l'amélioration de la culture informationnelle des étudiants, essentiellement en présentiel, de documents. On y trouve aussi des étudiants possédant des pratiques diverses et des enseignants de toutes les spécialités enseignées à l'IUT ayant eux aussi des pratiques informationnelles diverses.

Nous avons mis en place une méthodologie basée sur la collecte et le traitement des traces navigationnelles grâce aux services d'AGIMUS-NG. Ce travail a nécessité une étroite collaboration avec des ingénieurs informaticiens assortie de négociations pour sélectionner et préparer les données pertinentes. C'est dans cette optique que l'épaisseur du travail de traduction d'une trace en une donnée analysable par des chercheurs en SIC est expliqué.

Nous ajoutons quelques visualisations de l'activité numérique des étudiants produite dans AGIMUS-NG, qui sont, dans l'interprétation, enrichies par l'analyse de questionnaires et d'informations administratives et pédagogiques. Les résultats illustrent le potentiel des analyses des pratiques numériques informationnelles envisageables par le biais des traces navigationnelles.

Summary :

The main subject of this article is to respond to the following question: How to analyze, using navigational traces, the influence of a learning institutional environment on digital information practices of students ?

The institutional learning environment of IUT Jean Moulin Lyon 3 is made of a resource center, a digital library, an Intranet, classes related to the improvement of the information culture of students and documentations. It also includes students, who have several ways to access the information and each uses them according to their own knowledge. The professors, also part of this environment, have various backgrounds and specializations.

A methodology based on the collection and processing of navigation traces through the services of AGIMUS-NG has been implemented. This work required close collaboration with computer engineers which included negotiations to select and prepare the relevant data.

It is in this context that the importance of working on a mean to translate the navigational traces into analyzable data by researchers in the information and communication sciences field is explained.

The analysis of the visualizations of the digital student activity produced in AGIMUS-NG has been enriched by the interpretation and the analysis of questionnaires as well as administrative and pedagogical information. The results illustrate the potential of analyzes of informational digital practices that can be considered through navigation traces.

Mots-clés : Trace navigationnelle ; pratique informationnelle ; AGIMUS-NG ; bibliothèque numérique ; visualisation ; dispositif institutionnel d'apprentissage.

Introduction

De nos jours, il est indéniable que tout citoyen possède des comportements informationnels influencés par son environnement personnel et professionnel qui peuvent évoluer au cours du temps. A ces comportements peuvent s'ajouter des connaissances et des savoir-faire acquis entre autres par la constitution de patrimoines éducatifs et informationnels. Des préoccupations éducatives sont alors nécessaires pour accompagner les individus dans l'acquisition des compétences minimales en matière informationnelle, compatibles et adaptées à la multitude d'outils et de plateformes mis à leur disposition.

C'est dans ce cadre que s'inscrit notre projet de recherche. Nous souhaitons analyser les pratiques numériques des acteurs du dispositif informationnel de l'IUT Jean Moulin Lyon 3. Dans ce contexte, l'usage de la bibliothèque numérique et du centre de ressources de l'IUT peuvent modifier et faire évoluer les pratiques informationnelles (Chaudiron & Ihadjadene, 2010) des étudiants ou n'avoir aucune incidence. De ce fait, nous cherchons à évaluer l'usage autonome et durable du dispositif IUT, pour à terme, proposer des outils pédagogiques qui favorisent l'usage de la bibliothèque numérique.

Pour évaluer l'évolution des pratiques des étudiants, nous utilisons dans notre recherche une méthodologie plurielle basée sur l'analyse de données déclaratives et navigationnelles. Nous sommes en train de recueillir à la fois des données quantitatives et des traces d'activités numériques (Thibaud, 2010) des étudiants et des enseignants de l'IUT.

Pour les données quantitatives, nous avons administré des questionnaires (en octobre 2017, en mai 2018, puis en novembre 2018) auprès de tous les étudiants de l'IUT. Nous ré-administrerons ces questionnaires en mars 2019, date de fin des enseignements de deuxième année et en mai 2019 pour les premières années.

Parallèlement, nous avons effectué des entretiens qualitatifs auprès d'une vingtaine d'enseignants de l'IUT pendant l'année 2017-2018 pour établir le portrait des pratiques informationnelles des potentiels prescripteurs à l'usage du centre de ressources et de la bibliothèque numérique.

Les résultats issus de ces données déclaratives seront confrontés à l'analyse des traces de navigation des étudiants et des enseignants. En effet, pour discuter les résultats des questionnaires et des entretiens qui restent à caractère déclaratif, il est nécessaire de les croiser à l'analyse des « ombres informationnelles » (Rieder, 2010) que les étudiants et les enseignants laissent lors de leurs passages par la bibliothèque numérique. C'est le croisement

de ces deux types de données qui nous assurent des résultats finaux représentatifs de la réalité des pratiques informationnelles de notre population.

Pour cet article, nous nous concentrons sur les traces navigationnelles que nous considérons comme indicateurs des pratiques numériques informationnelles. Nous pouvons formuler notre questionnement de la manière suivante :

Comment analyser l'influence du dispositif institutionnel d'apprentissage sur les pratiques numériques informationnelles des étudiants par le biais de traces navigationnelles ?

La première partie définit la notion de pratique, de dispositif institutionnel d'apprentissage et de trace navigationnelle, termes utilisés dans la problématique.

La deuxième partie présente notre méthodologie de collecte et de traitement des traces navigationnelles. Pour ce faire, nous utilisons le service numérique AGIMUS-NG[1].

Enfin la dernière partie présente quelques visualisations de l'activité numérique des étudiants qui illustrent le potentiel des analyses des pratiques numériques informationnelles envisageables qui intègrent des traces navigationnelles.

I. Approches théoriques

1. Usage/Pratique informationnelle

Les notions d'usage et de pratique sont souvent convoquées quand on s'intéresse aux activités instrumentées (Rabardel, 1995). Lorsque l'on étudie une plateforme ou un outil, la notion d'usage est employée pour désigner la manière, mais aussi la fréquence avec laquelle un groupe social les utilise. Pour désigner ces manières de faire par le terme d'usage, celles-ci doivent être « *relativement stabilisées sur une période historique plus ou moins longue, à l'échelle d'ensembles sociaux plus ou moins larges (groupes, communautés, sociétés, civilisations)* » (Proulx, 2005). L'étude est alors centrée sur l'outil ou la plateforme. C'est l'objet qui est au centre des observations car c'est lui qui offre des fonctions et des services qui permettent des usages. La notion de pratique informationnelle permet, quant à elle de focaliser l'attention sur la manière de faire d'un individu lorsqu'il réalise une action. Il choisit les outils qui lui semblent les plus en adéquation avec ses compétences et avec la tâche qu'il a à effectuer. La pratique définit « l*a manière dont un ensemble de sources formelles ou non, d'outils, de compétences cognitives sont effectivement mobilisés, par un individu ou un groupe d'individus, dans les différentes situations de production, de recherche, d'organisation, de traitement, d'usage, de partage et de communication de l'information* » (Chaudiron & Ihadjadene, 2010, p. 14). Nous préférons le terme de pratique au terme d'usage car dans cette étude, nous centrons nos observations sur les comportements informationnels de l'étudiant et sur leurs représentations.

2. Dispositif institutionnel d'apprentissage

Dans le cadre de cette recherche, nous limitons notre dispositif à « *l'environnement institutionnel d'apprentissage »* (Peraya & Bonfils, 2014, p. 5) qui inclut la bibliothèque numérique de Lyon3[2]. Ce dispositif est constitué d'un centre de ressources, d'une bibliothèque numérique, d'un Intranet, d'enseignements liés à l'amélioration de la culture informationnelle des étudiants, essentiellement en présentiel, de documents. On y trouve aussi

[1] Nous présentons AGIMUS-NG dans la partie 3 de cet article

[2] La bibliothèque numérique de l'université Jean Moulin Lyon 3 comporte un accès gratuit par Intranet à plusieurs banques et bases de données en SHS. http://bu.univ-lyon3.fr/bibliotheque-numerique-acces-aux-ressources-322745.kjsp?RH=SCD-ACTUS

des étudiants possédant des pratiques diverses et des enseignants de toutes les spécialités enseignées à l'IUT ayant eux aussi des pratiques informationnelles diverses.

La notion de dispositif institutionnel d'apprentissage nous permet de considérer notre objet de recherche comme un espace d'activités collectives distribuées entre des actants humains et non humains en interaction (Larroche, 2018). Cet espace est aussi traversé par des discours institutionnels incitant à l'usage du numérique, matérialisés par certains enseignements, notamment ceux portant sur la méthodologie de recherche d'informations. En effet, le discours institutionnel et les moyens mis à disposition (notamment les enseignements, l'accès gratuit aux banques de données documentaires) visent à encourager l'usage des ressources numériques mis à disposition par l'IUT et par l'université Lyon 3.

En plus des éléments cités ci-dessus pour cerner ce dispositif institutionnel d'apprentissage, nous soulignons qu'il concilie « *une effectivité plurielle, celle des gestionnaires qui, à travers le dispositif, comptent bien atteindre les objectifs qu'ils se sont fixés, et celle des multiples usagers* » (Paquelin, 2011, p. 569) ; les étudiants sont donc au centre de nos préoccupations tout autant que les enseignants et les ingénieurs pédagogiques du centre de ressources documentaires numériques (CRDN).

Si certains chercheurs en SIC s'intéressent aux environnements informationnels que se construisent les étudiants (Bonfils & Peraya, 2011; Peraya & Bonfils, 2012), nous nous intéressons dans le cadre du projet aux évolutions de leurs pratiques informationnelles.

3. Traces navigationnelles

Avec le développement des innovations techniques et numériques dans la sphère informationnelle, des besoins de consultation et d'analyse des pratiques se multiplient. À titre d'illustration citons « *l'analyse des logs de consultation d'Internet en accès libre à la Bpi* » (Amar, 2018). Dans ce contexte, les concepteurs de logiciels, de plateformes et de solutions numériques offrent aujourd'hui des outils de mesures d'activités enrichies par des graphiques statistiques qui sont censés aider à la décision et à la recherche scientifique. Ces traqueurs de traces, nous les trouvons par exemple joints aux plateformes de réseaux sociaux et aux *back offices* des sites web. Ces actions couplées à des données d'identification, s'élèvent à des centaines voire des milliers de lignes d'informations.

Se posent alors maintes questions sur la traçabilité de l'action humaine, le stockage de ces traces, leur gestion et leurs usages.

Toutes ces questions sont à l'origine de plusieurs débats. Pour cet article nous nous limitons à la question de la gestion et de l'utilisation de ces données enregistrées dans le cadre précis de notre terrain de recherche, à savoir : le dispositif institutionnel d'apprentissage de l'IUT Jean Moulin. Dans ce contexte, les traces sont essentiellement liées à l'usage des banques de données de la bibliothèque numérique de l'université.

L'administrateur de la bibliothèque numérique peut alors enregistrer les actions exécutées par un utilisateur : ouvrir une page, cliquer sur un lien, télécharger un article, se connecter à une base de données, etc. Ces traces seront nommées *Logs* de navigation dans cet article.

Pour nous, une action de l'usager (une connexion à une base comme *Cairn* par exemple) a donc une valeur de trace (Ginsburg, 1989) quand un analyste lui donne une valeur d'indice, créant un lien de cause à effet entre les deux. En effet, les pratiques des étudiants provoquent la production de *logging* sur les banques de données disponibles dans la bibliothèque numérique, construisant une « *ombre informationnelle* » (Rieder, 2010) qui se déploie derrière chaque usager. Lors d'analyse de *Logs*, cette ombre est considérée comme « *une source de connaissance, une représentation potentiellement authentique de groupes et*

d'individus, de leurs comportements, idées, besoins, valeurs, relations, antagonismes et ainsi de suite » (Rieder, 2010).

II. AGIMUS-NG

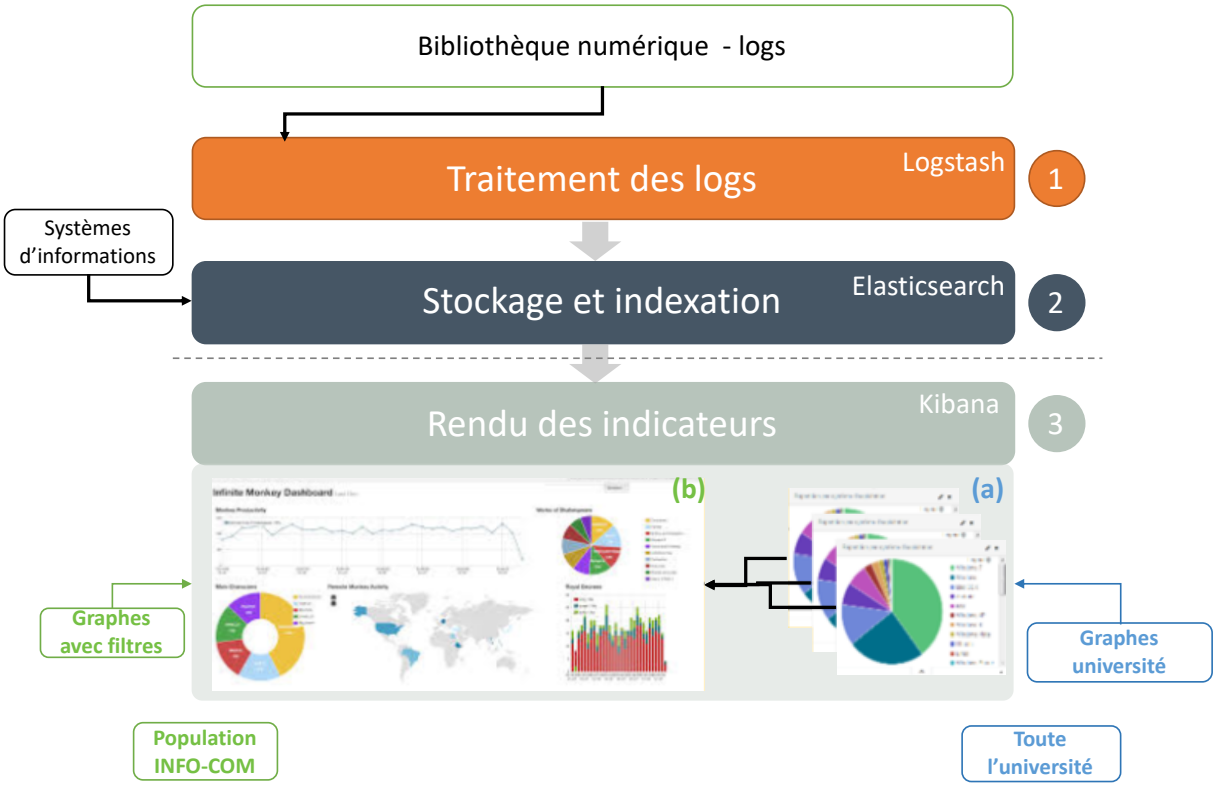
Pour parvenir à l'analyse de l'influence du dispositif institutionnel d'apprentissage sur les pratiques numériques informationnelles, comme nous l'avons évoqué en introduction, nous réduisons pour le moment les pratiques informationnelles aux traces navigationnelles. Cette partie montre l'épaisseur du travail de traduction d'une trace en une donnée analysable par des chercheurs en SIC. C'est pourquoi, nous avons choisi de nous intégrer au projet AGIMUS-NG qui propose des services logiciels spécialisés dans le traitement et la visualisation des traces navigationnelles dans les universités. Ce processus nécessite aussi une collaboration indispensable avec les services informatiques de l'université Jean Moulin Lyon 3 pour bénéficier de compétences logicielles, de sécurité informatique et d'expertise AGIMUS-NG.

1. *AGIMUS-NG pour la Collecte et le traitement navigationnel*

Le projet AGIMUS-NG a été créé en 2014 et lancé par la direction générale de l'enseignement supérieur et de l'insertion professionnelle (DGESIP). Les objectifs de ce projet (Can, Colson, & Marchal, 2015) qui correspondent aux nôtres sont :

- de faciliter la création par les établissements d'enseignement supérieur d'indicateurs d'usages de leurs services numériques,
- de comprendre et analyser les pratiques informationnelles numériques des usagers.

Fig. 1 : Représentation d'AGIMUS-NG par des chercheurs en SIC



La figure 1 présente les 3 couches logiciels simplifiées du service AGIMUS-NG. La première couche, par le biais du logiciel *logstash*[3] traite les journaux de traces des différentes bases et applications utilisées par l'université (dans notre cas, elles sont représentées par la bibliothèque numérique). La deuxième couche, par le biais du logiciel *Elasticsearch*[4] représente le stockage de ces traces, leur indexation et leur transformation en visuel consultable dans *Kibana*, logiciel de visualisation représentant la troisième couche. L'enrichissement avec des informations d'usage issues du système d'informations de l'établissement en question (dans notre cas, l'IUT Jean Moulin Lyon 3) est aussi réalisable dans cette deuxième couche[5].

Kibana est donc l'interface d'*ElasticSearch* qui traduit les *Logs* stockés en visuels compréhensibles par l'homme. Comme le service AGIMUS-NG s'applique à l'ensemble de l'activité numérique de l'université (Environnement numérique du travail, Intranet, etc. (voir par exemple Fig.3)), le service informatique est le seul à être administrateur et à pouvoir créer des visuels de représentation limités à notre terrain (bibliothèque numérique de l'université). Ces représentations génériques forment ensuite le tableau de bord que nous, chercheurs en SIC, pouvons manipuler. *Kibana* propose des filtres prédéfinis pour limiter les visuels à notre population (étudiants du département information communication pour cet article) et à nos questions de recherche[6].

En d'autres termes dans notre projet, AGIMUS-NG permet de traduire les *Logs* de navigation de la bibliothèque numérique sous forme de tableaux de bord et rend ainsi compréhensibles les métadonnées récupérées par le serveur informatique *EZproxy*[7] croisées aux données d'identification des étudiants de l'université. Ces précisions montrent la nécessité d'une collaboration étroite entre chercheurs en SIC et informaticiens de l'université Lyon 3 et de la bibliothèque, certains étant experts et administrateurs d'AGIMUS-NG.

2. *Le processus de négociation entre chercheurs en SIC et ingénieurs en informatique*

Pour analyser les *Logs* de navigation de la bibliothèque numérique, nous nous sommes donc rapprochés d'ingénieurs informaticiens de l'université Jean Moulin Lyon 3 qui travaillent sur la plateforme AGIMUS-NG. Jusqu'ici, ils manipulaient les *Logs* de navigation de *Moodle*, plateforme pédagogique et collaborative accessible de l'Intranet de Lyon 3 pour le Pôle d'Accompagnement à la Pédagogie Numérique (PAPN).

Pour l'analyse de nos questions par les services d'AGIMUS-NG, un dialogue est nécessaire pour adapter nos préoccupations de chercheurs aux contraintes logistiques.

Lors de cette négociation de données (Denis & Goëta, 2013), nous avons produit plusieurs documents qui ont été le support des réunions entre ingénieurs informaticiens, chercheurs en SIC, responsables pédagogiques et enseignants :

- une liste de variables classées par catégorie (Etudiants, Enseignants, Emploi du temps, *Logs* de navigation) indiquant leur fichier d'origine au sein de l'IUT.
- Des cartographies représentant les questions de recherche, les différentes hypothèses associées, et les variables à interroger.

---

[3] https://www.elastic.co/fr/products/logstash
[4] https://www.elastic.co/fr/products/elasticsearch
[5] Cette étape ne sera pas traitée dans cet article.
[6] Les questions de recherche qui serviront d'illustration pour cet article seront expliquées dans la partie 3.
[7] EZproxy est un logiciel de type proxy principalement utilisé par les bibliothèques universitaires afin de fournir un accès distant aux ressources numériques auxquelles elles sont abonnées.

Les ingénieurs nous ont demandé de remplacer les noms des variables par leurs correspondants dans les annuaires pour l'enseignement supérieur (SupAnn) utilisés par AGIMUS-NG.

Si l'on suit les étapes proposées par Denis et Goëta (2013), le travail préalable que nous venons d'évoquer correspond à « *la détermination des données et de la négociation* ».

Le simple recensement des données, ne les rend ni manipulables, ni visualisables dans AGIMUS-NG.

Pour (Denis & Goëta, 2013), l'étape suivante d'« *opérationalisation de l'identification des données »* est nécessaire.

Cette étape implique l'accès aux différentes plateformes, bases de données et annuaires, leur extraction et leur transformation en format unifié. Seule cette procédure rend possible leur visualisation dans *Kibana*. A l'issue de cette étape, nous parviendrons à construire une interface qui convertit les traces en des données intelligibles.

Recenser les données et les rendre disponibles présentent les deux premières étapes de ce processus. Vient en troisième et dernier lieu l'étape de brutification. Il s'agit de produire un formatage qui les unifie selon un même langage, et ce, en les nettoyant afin d'augmenter leurs qualités. « *Des données brutes sont des données jugées bonnes, au bon format, « propre » et intelligibles* » (Denis & Goëta, 2013). Ces données sont celle que nous pouvons visualiser dans les graphes de notre tableau de bord *Kibana*.

Contrairement à la première étape à laquelle nous avons collaboré, les deux suivantes sont à la charge des informaticiens d'autant plus lorsque nous ne possédons pas un accès administrateur. Le statut d'administrateur est réservé aux personnes habilitées à manipuler des données sensibles. Un des intérêts d'AGIMUS-NG est de mettre à disposition des chercheurs et des décideurs, « *exploitants de l'information* » (Denis & Goëta, 2013) des données anonymisées, permettant de respecter le Règlement Général sur la Protection des Données (RGPD)[8]. Le statut d'administrateur fait perdre cette anonymisation d'où la difficulté pour des chercheurs d'obtenir ce statut.

III. Méthodologie pour le rendu de nos indicateurs

C'est donc uniquement la partie *graphe avec filtre* de la figure 1 qui est manipulable par les chercheurs. Nous l'illustrerons en partie (III.2). La limitation de l'accès pour les chercheurs nous oblige à définir avec l'administrateur la construction de notre tableau de bord qui contient les graphes génériques que nous avons négociés. Cette opération est représentée par la partie droite du troisième niveau de la figure 1 (a). Elle sera illustrée dans la partie suivante par la construction de graphes génériques.

1. Co-création des représentations génériques constitutives du tableau de bord

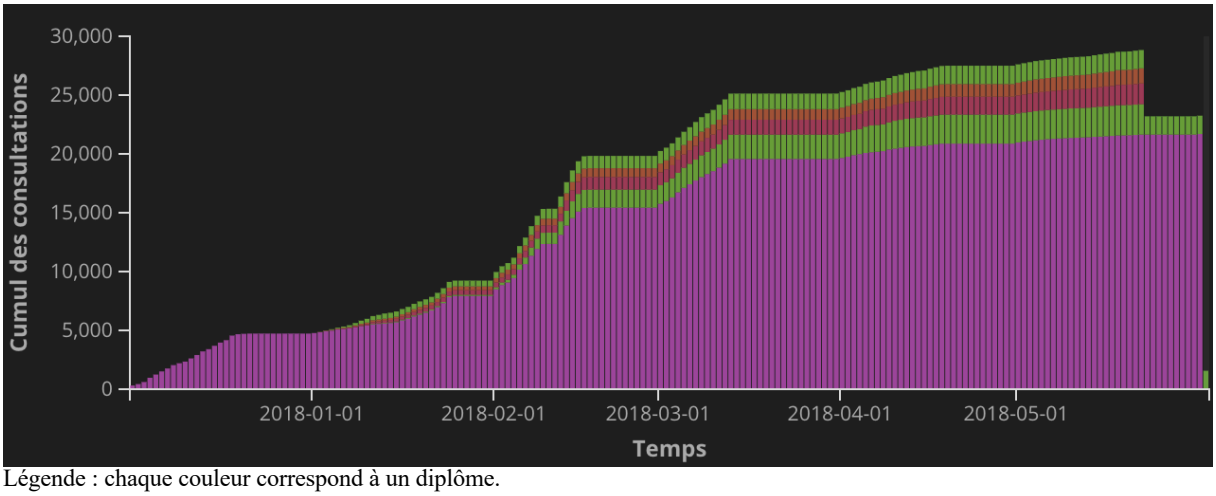
C'est grâce à notre observation du comportement des étudiants, des attentes des enseignants relevées lors d'entretiens et des questionnaires administrés auprès des étudiants que nous avons pu sélectionner les graphes génériques qui nous intéressaient.

[8] https://www.economie.gouv.fr/entreprises/reglement-general-sur-protection-des-donnees-rgpd

L'objectif de notre recherche étant d'analyser le comportement informationnel numérique des étudiants de l'IUT, nous avons opté avec le gestionnaire *Kibana* de l'université Jean Moulin Lyon 3 pour la construction des graphes qui englobent l'ensemble de la population de l'université (graphes université dans la Fig. 1) suivants :

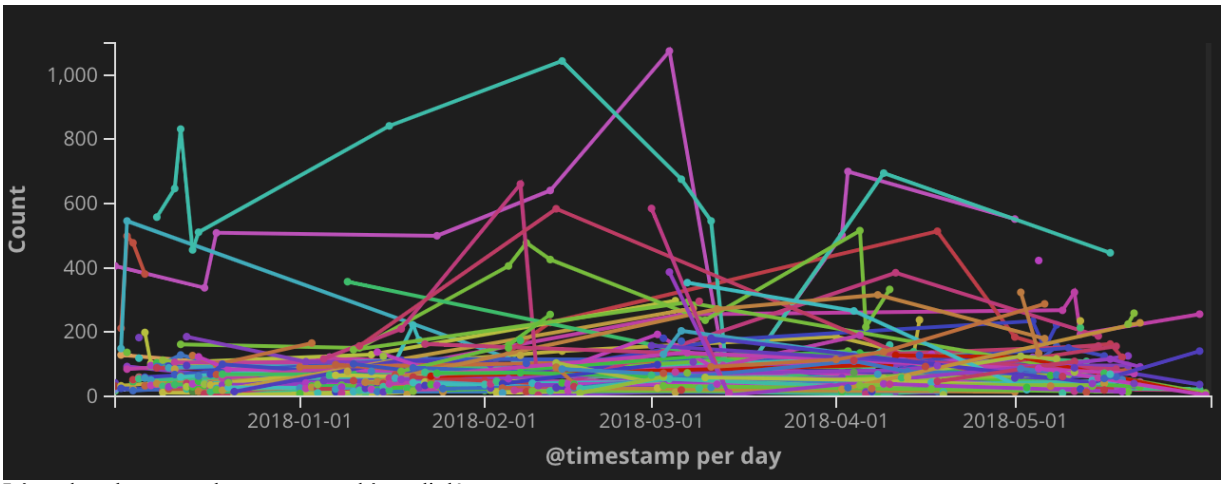
- Cumul des consultations (Fig. 2)
- Consultations par diplôme (Fig. 3)
- Répartition des accès par utilisateur (Fig. 4)
- Publications par diplôme (Fig. 5)

Fig. 2 : Cumul des consultations



Légende : chaque couleur correspond à un diplôme.

Fig. 3 : Consultations par diplôme



Légende : chaque couleur correspond à un diplôme.

Fig. 4 : Répartition des accès par utilisateur

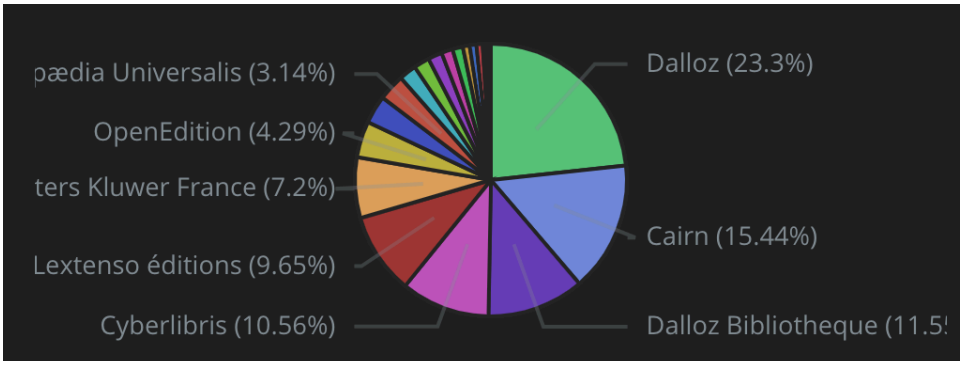
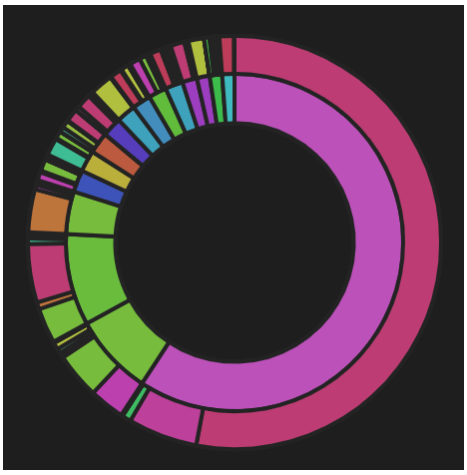


Fig. 5 : Publications par diplôme

Légende : La couronne de l'intérieur correspond aux différents diplômes.

La couronne de l'extérieur correspond aux banques de données utilisées



Ces graphes université, une fois intégrées dans le tableau de bord, pourront avec certaines manipulations (applications de filtres) nous aider à visualiser nos hypothèses ou nos questions.

2. *La traduction des questions de recherche dans le tableau de bord Kibana*

Avant de pouvoir apposer des filtres spécifiques, nous avons dû comprendre le fonctionnement du tableau de bord de *Kibana* et rendre en main l'index *EZPaarse*. Cet index comprend l'ensemble des champs disponibles pour affiner les graphes université et décliner des graphes avec filtres qui pourront apporter des réponses à nos questions de recherche. Cet index comporte une quarantaine d'entrée qui sont en réalité les filtres. Par exemple l'entrée ou le filtre *department* permet de réduire les *Logs* récupérés par l'UFR ou service d'appartenance de la personne. Le filtre *platform* permet quant à lui de sélectionner uniquement les *Logs* de navigation qui concernent une ou plusieurs plateforme(s) documentaire(s) consultée(s).

C'est en combinant plusieurs filtres que l'on peut traduire nos questions de recherche en requêtes compréhensibles par *Kibana*.

Pour cet article nous nous focaliserons que sur les deux questions suivantes :

- Q1 : Quelles évolutions de consultation de la bibliothèque numérique Lyon 3 entre décembre 2017 et mai 2018 pour les premières et deuxièmes années des options

Information numérique des organisations (INO) et communication dans les organisations (Com) ?

- Q2 : Quelles banques de données de la bibliothèque numérique sont les plus consultées entre décembre 2017 et mai 2018 pour les deuxièmes années des options INO et Com ?

Leur traduction pour *Kibana* a demandé l'application de plusieurs filtres dont *date, département, code diplôme* pour permettre d'obtenir sur le graphe uniquement les *Logs* de navigation de cette population.

Les filtres n'ont pas été apposés sur les mêmes graphes génériques pour les 2 questions. Les filtres ont été apposés sur le graphe générique *consultations par diplôme (Fig.3)* et *cumul des consultations (Fig.2)* pour la question 1. Pour la question 2, les filtres ont été apposés sur les graphes génériques *répartition des accès par utilisateu*r et *publications par diplôme*

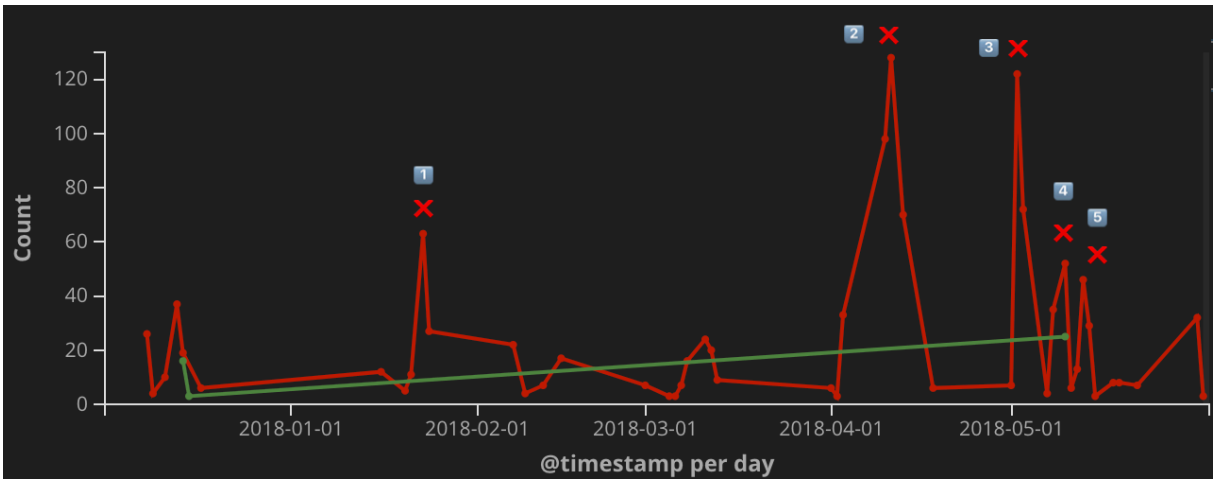
3. *Une ébauche d'analyse des pratiques informationnelles numériques des étudiants de l'IUT Lyon3*

Rappelons que notre projet de recherche concerne l'IUT Jean Moulin Lyon 3. Nous souhaitons croiser à terme des données déclaratives avec des données navigationnelles visualisables sur notre tableau de bord. Il nous faut donc obtenir des graphes pour la population de l'IUT que nous réduisons pour cet article au département information et communication. La période de l'administration des questionnaires étant octobre 2017, mai 2018 et octobre-novembre 2018, nous avons choisi de ne retenir que la période 2017-2018 pour l'analyse des *Logs* de navigation. Le but de cette partie est de montrer le potentiel de *Kibana* et les difficultés en matière d'analyses des traces navigationnelles à partir des 2 questions mentionnées en partie 3.2.

Pour la question 1 portant sur l'évolution des consultations de la bibliothèque numérique de Lyon 3, nous avons choisi de distinguer sur un même graphe, les *Logs* de navigation des étudiants de première et de deuxième année. Comme nous souhaitons comparer les pratiques des INO (25 étudiants en première année, 22 en deuxième année) et des COM (52 étudiants en première année, 47 en deuxième année), nous avons produit deux graphes en déclinant le graphe université[9] *consultation par diplôme* (Fig. 3) en deux graphes, le premier (Fig. 6) filtré par *diplôme=INO* et le deuxième (Fig. 7) filtré par *diplôme=COM.*

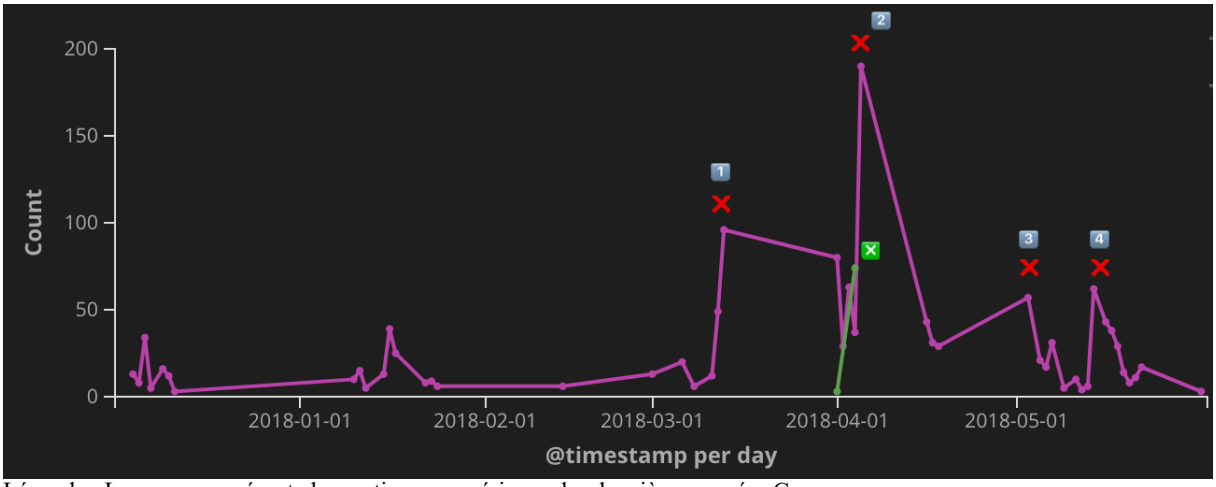
[9] La valeur par défaut est l'ensemble de l'université.

Fig. 6 : Évolution de la consultation de la bibliothèque numérique des INO entre décembre 2017 et mai 2018



Légende : Le rouge représente les pratiques numériques des deuxièmes années INO
Le vert représente les pratiques numériques des première années INO

Fig. 7 : Évolution de la consultation de la bibliothèque numérique des COM entre décembre 2017 et mai 2018



Légende : Le mauve représente les pratiques numériques des deuxièmes années Com
Le vert représente les pratiques numériques des première années Com

Pour l'analyse, nous cherchons essentiellement à expliquer les pics des courbes et à comparer les deux populations INO (Fig. 6) et COM (Fig. 7).

Nous définissons un pic de consultation comme une fracture dans la continuité de la courbe qui donne lieu à une augmentation subite du nombre de consultation supérieure au nombre d'étudiants[10]. Pour la figure 6, nous comptabilisons pour les deuxièmes années INO (courbe rouge) cinq pics (chaque pic est indiqué par une croix). Pour les premières années, il n'y a qu'une faible augmentation continue du nombre de consultation. Pour la figure 7 qui représente la population COM, nous comptabilisons quatre pics (chaque pic est indiqué par une croix) pour les deuxièmes années et un pic le 1er avril pour les premières années.

Les pics seront interprétés par le biais d'informations complémentaires que nous avons récupérées auprès des services administratifs et des enseignants. Il s'agit notamment de

[10] Nous ignorons les pics qui indiquent un nombre de consultation inférieur au nombre d'étudiants.

l'emploi du temps des étudiants, du contenu des séances de cours à la date du pic, des consignes données pour les projets et les devoirs et les dates de rendus.

En INO, le pic 1 de la figure 6 apparaît en deuxième année en semaine 4. Cette semaine est celle liée au *challenge de la veille* qui met en concurrence des équipes d'étudiants des départements INO des IUT participants. Le sujet a été donné par V33[11], le commanditaire de 2018. Ce pic est lié à un enseignant qui les a incités lors de son cours à consulter la bibliothèque numérique pour cerner le sujet. Nous pouvons associer ce premier pic à une injonction. Les COM 2 n'ont pas sollicité les banques documentaires mises à leur disposition à cette même période. Il n'y avait pas de cours lié à des recherches documentaires à cette période pour ce groupe.

Le deuxième pic d'utilisation de la bibliothèque numérique est, quant à lui, présent pour les deux populations de deuxième année. Il s'agit d'une utilisation intensive de la bibliothèque numérique qui a atteint son maximum entre le 1er et le 2 avril 2018. Cela s'explique, selon nous, par le fait que tous les deuxièmes années sont amenés à rendre le 2 avril deux dossiers dans le cadre du cours *préparation au mémoire de stage*. Chaque étudiant devait formuler une problématique liée à son stage et rédiger, à l'aide de recherche de documents scientifiques, un document où il présentait l'état de l'art sur la question. Une analyse institutionnelle de leur organisme d'accueil assortie d'une bibliographie leur était aussi demandée. Tout au long du mois mars, l'enseignant de cette matière incitait les étudiants à l'usage de la bibliothèque numérique. Nous observons un pallier assez élevé sur le graphique 7 pendant le mois de Mars. Ce constat peut s'expliquer par le besoin d'explication, voire de démonstration des COM, peu familiers avec l'usage de la bibliothèque numérique. Quant aux INO, le pallier est moins élevé en figure 7 au mois de mars. Ce fait peut s'expliquer par une meilleure maîtrise de la bibliothèque numérique. Le pic 2 coïncide avec la date de rendu, qui n'est pas associée à un TD. On entrevoit un usage autonome de la bibliothèque numérique qui se confirme pour les pics 3, 4, 5 de la figure 6, et les pics 3 et 4 de la figure 7. En effet, les étudiants sont en stage à cette période et les dates correspondent à des jours fériés (1er mai, ascension, week-end). A l'issue de leurs stages, les étudiants devaient en effet rédiger un mémoire.

L'analyse des graphiques fournis par *Kibana* sont riches d'enseignement à la condition d'être associés à d'autres informations.

Pour la question 2, nous allons détailler le type de banques de données consultées par les étudiants INO (Fig. 8) et COM (Fig. 9) de deuxième année entre décembre 2017 et mai 2018. Les deux graphes sont produits à partir du graphe université de la figure 4. Nous avons omis de détailler les banques de données des premières années car le nombre de consultation visualisable sur les figures 6 et 7 sont très faibles.

[11] Le sujet du challenge est disponible sur : http://challenge-veille-2018.fr/sujet/

Fig. 8 : Banques de données documentaires les plus consultées par les étudiants INO 2

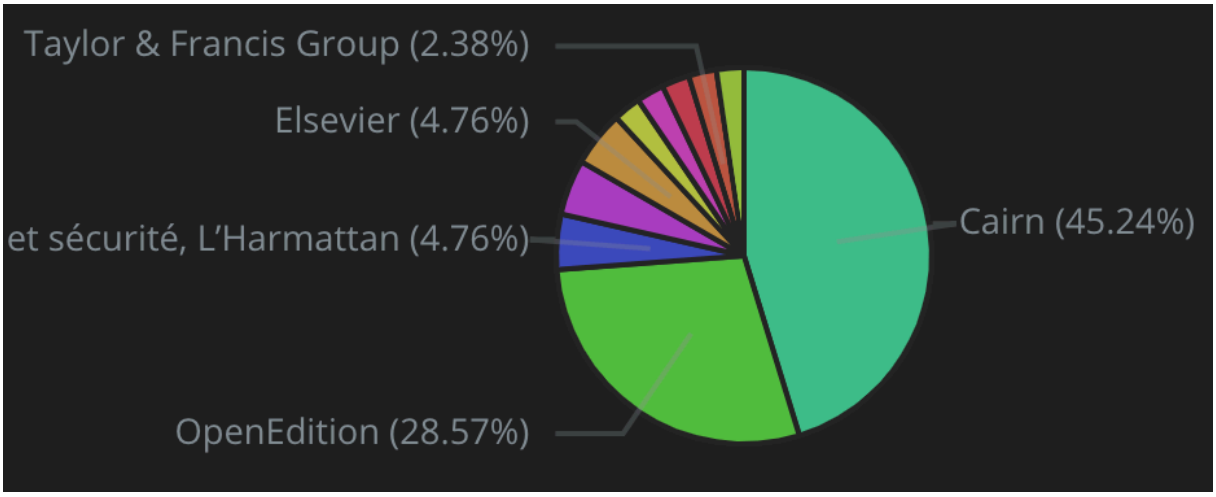
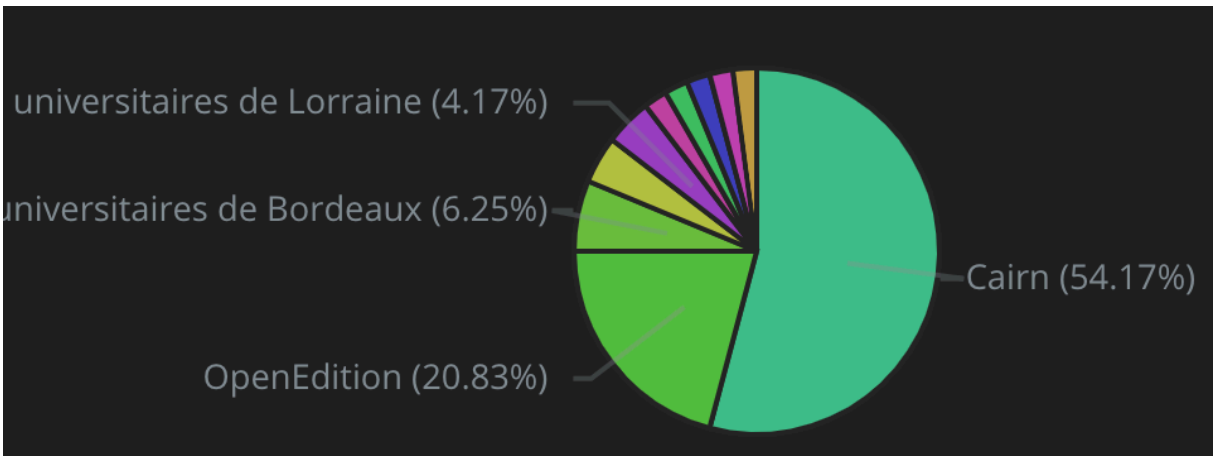


Fig. 9 : Banques de données documentaires les plus consultées par les étudiants COM2



Nous pouvons remarquer que les banques de données *Cairn* et *OpenEdition* sont les plus consultées à la fois par les COM et par les INO. En effet, 45 % des consultations des INO et 54 % des consultations des COM sont pour *Cairn*. Et 28% des consultations des INO et 20 % des consultations des COM sont pour *OpenEdition*.

La consultation de ces deux banques de données qui abritent principalement des documents et des articles scientifiques peut s'expliquer par les consignes, par le type de cours que les étudiants suivent et par la nature des projets réalisés pendant l'année. C'est surement les traces navigationnelles de ces banques de données qui sont représentées par les pics de consultation des figures 6 et 7. Ce constat pourrait être confirmé par des entretiens avec des étudiants.

En plus de cette consigne, tous les étudiants du département Information Communication de l'IUT ont bénéficié d'au moins un cours de *recherche d'informations* pendant leur cursus. Dans ce cours, il était question de présenter et de manipuler les bases et catalogues de données suivants : *Sudoc, Cairn, Persée, Revues.org, Généralis, Factiva, Europresse, Worldcat*. Pourquoi nous ne retrouvons pas toutes ces bases dans les deux figures 8 et 9 alors que les étudiants affirment[12] les utiliser ?

[12] Cette information est indiquée dans les différents questionnaires administrés.

Il s'agit en effet, d'une des limites du projet AGIMUS-NG. Celui-ci ne pouvant récupérer que des informations tracées suivant une logique d'authentification, nous ne pouvons pas visualiser, par exemple, l'activité de nos étudiants sur le portail *Sudoc* et *Persée*.

Nous avons pu observer un manque de précisions et parfois de cohérence dans les représentations graphiques produites par AGIMUS-NG, ce que nous considérons comme une autre limite. En effet, comme il est affiché sur les deux graphes 8 et 9, nous pouvons compter des maisons d'éditions (*Presse universitaire de Bordeaux, Presse universitaire de Lorraine, L'Harmattan, etc*.) présentées comme banques de données documentaires. Cela est selon nous impossible car l'accès à ces maisons d'édition n'est pas répertorié dans la bibliothèque numérique[13]. Ainsi, *Kibana* afficherait deux indicateurs de consultation, celui de *Cairn* si un utilisateur repère un document par cette banque et, celui du catalogue de la *presse universitaire de Bordeaux* s'il le consulte sur ce site ? Cette question reste aujourd'hui sans réponse. Nous continuons à utiliser *Kibana*, à négocier et à échanger avec les services informatiques de l'université afin d'apporter une réponse et clarifier nos doutes.

Conclusion

L'objectif final de notre projet de recherche est de présenter des solutions pratiques aux utilisateurs de la bibliothèque numérique de l'IUT Jean Moulin Lyon 3. Ces solutions viseront à mieux utiliser ces ressources numériques et à inculquer aux étudiants mais aussi aux enseignants, des réflexes de recherche d'informations pensés, construits et pertinents. Avant de proposer des solutions adéquates, nous pensons qu'une analyse fine des pratiques est nécessaire.

Dans cet article, nous avons amorcé cette analyse en présentant la méthodologie liée au traitement des traces navigationnelles. Nous avons illustré la nécessité de croiser les données de navigation à d'autres informations (emploi du temps, contenu des enseignements, etc.) afin d'interpréter les graphes. Ce constat confirme la nécessité d'appréhender les pratiques informationnelles par un dispositif institutionnel d'apprentissage. Il reste néanmoins à approfondir nos analyses de questionnaires et d'entretiens pour les croiser à d'autres questions de recherche et à élargir le panel à l'ensemble des départements de l'IUT (Gestion Administrative et Commerciales des Organisations et Carrières juridiques).

Les premiers résultats limités à deux questions dans cet article peuvent être résumés ainsi :

- La consultation est liée à une injonction d'un enseignant,
- une consultation autonome des banques de données apparaît pendant le stage,
- les banques de données les plus utilisées sont celles qui abritent des documents scientifiques en texte intégral, montrant ainsi le poids de l'injonction des enseignants.

Des doutes persistent quant à l'affichage de certaines données qui sont en relation avec les banques de données utilisées.

La méthodologie présentée dans cet article et celle du projet de recherche global pourront s'appliquer à d'autres formations. AGIMUS-NG étant installé à l'université Jean Moulin Lyon 3, l'investissement sera moindre que pour d'autres universités ou écoles. A terme, ce projet de recherche pourra aider au pilotage du numérique et pourra s'élargir aux traces

[13] Les banques de données documentaires de la bibliothèque numérique sont disponibles sur : http://bu.univ-lyon3.fr/bibliotheque-numerique-acces-aux-ressources-322745.kjsp?RH=SCD-ACCUEIL&RF=SCD-ELECtest

d'activités numériques informationnelles à l'échelle nationale en créant une dynamique d'échange.

Bibliographie :

Amar, M. (2018). *Analyse des logs de consultation d'Internet en accès libre à la Bpi : qu'apporte le Big Data ?* (report). Bibliothèque Publique d'Information. Retrieved from https://archivesic.ccsd.cnrs.fr/sic_01721032/document

Bonfils, P., & Peraya, D. (2011). Environnements de travail personnels ou institutionnels ? Les choix d'étudiants en ingénierie multimédia à Toulon. In *Le numérique au coeur des partenariats : enjeux et usages des technologies de l'information et de la communication* (Presse universitaire de Dakar, pp. 13–28). Dakar.

Can, N., Colson, G., & Marchal, J. (2015). Agimus NG ou les indicateurs d'usage de nos services numériques.

Chaudiron, S., & Ihadjadene, M. (2010). De la recherche de l'information aux pratiques informationnelles. *Études de communication. langages, information, médiations*, (35), 13–30. https://doi.org/10.4000/edc.2257

Denis, J., & Goëta, S. (2013). La fabrique des données brutes. Le travail en coulisses de l'open data. Presented at the Penser l'écosystème des données. Les enjeux scientifiques et politiques des données numériques. Retrieved from https://halshs.archives-ouvertes.fr/halshs-00990771/document

Ginsburg, C. (1989). Trace. Racines d'un paradigme indiciaire. In *Mythes, emblèmes, traces* (pp. 139–180). Paris: Flammarion.

Larroche, V. (2018). *Le dispositif*. ISTE Editions.

Paquelin, D. (2011). La distance : questions de proximités. *Distances et savoirs*, *9*(4), 565–590.

Peraya, D., & Bonfils, P. (2012). Nouveaux dispositifs médiatiques : comportements et usages émergents. Le cas d'étudiants toulonnais en formation à l'UFR Ingémédia. *Distances et médiations des savoirs. Distance and Mediation of Knowledge*, *1*(1). https://doi.org/10.4000/dms.126

Peraya, D., & Bonfils, P. (2014). Détournements d'sages et nouvelles pratiques numériques : l'expérience des étudiants d'Ingémédia à l'Université de Toulon. *Revue des sciences et techniques de l'information et de la communication pour l'éducation et la formation*, (21). Retrieved from https://archive-ouverte.unige.ch/unige:46581

Proulx, S. (2005). Penser les usages des TIC aujourd'hui : enjeux, modèles, tendances. *Presse universitaire de Bordeaux*, *t. 1*, 7–20.

Rabardel, P. (1995). Les hommes et les technologies; approche cognitive des instruments contemporains, 239.

Rieder, B. (2010). Pratiques informationnelles et analyse des traces numériques : de la représentation à l'intervention. *Études de communication. langages, information, médiations*, (35), 91–104. https://doi.org/10.4000/edc.2249

Thibaud, H. (2010). Représenter et modéliser l'activité de recherche d'information experte avec des traces d'activité pour l'apprentissage, Abstract, Resumen. *Les Enjeux de l'information et de la communication*, *me 2010*(1), 83–96.