

Publications en libre accès des universités du Burkina Faso: analyse d'impact et visibilité internationale

Pier Luigi Rossi¹, Minata Traore², and Fatoumata Maïga Diallo³

¹*Institut de Recherche pour le Développement, France*
Email: rossi@ird.fr

²*Université de Bobo Dioulasso, Burkina Faso*

³*Université de Ouagadougou 1, Burkina Faso*

Résumé

Nous avons rendu disponibles, en libre accès sur internet, les collections de mémoires et de thèses soutenues à l'Université de Bobo Dioulasso et à l'Université de Ouagadougou 1 (Burkina Faso). L'analyse des données de consultation (« fichiers de log » présents sur le serveur hôte) permet de produire plusieurs indicateurs concernant l'impact et la visibilité internationale des documents constituant les collections. Nous présentons la répartition temporelle et géographique des consultations, les fréquences des accès spécifiques à chaque document, la nature des questions, l'impact des moteurs de recherche. Ces résultats montrent qu'en mettant à disposition sur Internet des documents scientifiques produits par des institutions africaines, l'importance de leur visibilité et de leur impact peuvent être démontrés en utilisant les données de consultation disponibles sur le serveur. La comparaison de ces données avec celles disponibles pour les documents en libre accès de l'Institut français pour le développement (IRD) permet de conclure que les niveaux de consultation pour les publications scientifiques des pays du « Nord » et du « Sud » sont assez similaires.

Two major collections of dissertations and theses defended at the University of Bobo Dioulasso and the University of Ouagadougou 1 have been made available open access on the Internet. The data analysis of accesses available via the host server makes it possible to produce several indicators concerning the impact and the international visibility of the available documents. We will present the temporal and geographical distribution of the consultations, the specific frequencies for each document, the nature of the questions, the impact of the search engines. These results seem to show that by making available on the Internet scientific documents produced by African institutions their visibility and impact are demonstrable with the data collected by the server. A comparison of these data with those available for the open access documents of the French Institute for Development (IRD) show that the levels of consultation for scientific publications from "Northern" and "Southern" countries are quite similar.

1 Introduction

La collaboration entre l'IRD (Institut de Recherche pour le Développement) et les bibliothèques de l'Université de Bobo Dioulasso et de l'Université de Ouagadougou s'est concrétisée à la suite du projet SIST (Système d'information scientifique et technique), en 2008–2009, et du projet « Thèses en Afrique », en 2014 (ROSSI 2013; ROSSI 2017). Grâce aux financements de ces projets, l'IRD a pu réaliser in situ, au bénéfice des personnels directement

concernés, un transfert de compétences concernant la numérisation des documents et la production de fichiers électroniques au format pdf tout en assurant l'installation des ateliers de numérisation (scanners de productions, logiciels professionnels et chaîne de productions intégrée).

En 2010, pour rendre accessibles sur internet les collections des documents numérisés par plusieurs institutions ayant bénéficié du projet SIST, l'IRD a mis en service le serveur collaboratif BEEP (bibliothèques électroniques en partenariat, www.beep.ird.fr) (ROSSI 2011). Ce serveur vise à valoriser les productions scientifiques des institutions partenaires de l'IRD en essayant de contourner les difficultés qui limitent le fonctionnement de serveurs internet en Afrique (faible bande passante, temps d'arrêt du serveur, coupures de courant fréquentes, mise en œuvre complexe de projets informatiques), en attendant des solutions locales efficaces.

Le site BEEP héberge en libre accès la collection des mémoires de l'Université de Bobo Dioulasso et la collection des thèses de l'Université de Ouagadougou, les deux principales Universités du Burkina Faso.

La collection de l'Université de Bobo Dioulasso est constituée de 922 mémoires de fin d'études dont 626 relèvent de l'Institut du développement rural et 84 de l'École supérieure d'informatique. La collection de l'Université de Ouagadougou est constituée de 493 thèses concernant principalement la Médecine, la Chimie, la Pharmacie.

Produire des statistiques de consultations des collections de documents en libre accès nous paraît essentiel pour mesurer l'impact des documents à l'échelle de l'internet et pour produire des analyses stratégiques utiles aux décideurs des institutions qui produisent ces documents scientifiques (MERCEUR 2015 ; SCHÖPFEL et PROST 2010).

L'analyse des fichiers d'historique des consultations (fichiers de log)¹ du serveur BEEP permet d'élaborer différents indicateurs concernant les consultations des documents disponibles et la localisation géographique des utilisateurs internet. Dans cet article nous allons analyser plus particulièrement les données et les indicateurs des deux collections des Universités du Burkina Faso.

2 Historique des consultations et préparation des données

Le site internet BEEP fonctionne avec le serveur Apache HTTP². Ce logiciel stocke de façon détaillée l'historique des consultations concernant les échanges de fichiers qui se réalisent entre le serveur et le client (la machine qui consulte les fichiers disponibles sur le serveur). Le format du fichier d'historique, en mode « combiné »³, contient plusieurs informations (fig. 1) dont : l'adresse IP du client (41.73.108.18), l'heure à laquelle la requête a été reçue (10/Mar/2016 :22 :58 :07 +0100), l'url du fichier consulté (/collect/.../IDR-2011-YOD-ETU.pdf), le code du statut retourné par le serveur (200), le site depuis lequel le client a lancé

1. Voir la définition d'historique informatique, journal et log à : [fr.wikipedia.org/wiki/Historique_\(informatique\)](http://fr.wikipedia.org/wiki/Historique_(informatique))

2. La documentation et le logiciel « Apache HTTP Server » sont disponibles à : httpd.apache.org

3. Voir : <http://httpd.apache.org/docs/2.4/logs.html>

sa requête (<http://www.google.ml> ...), l'éventuelle question posée par le client (l'étude de fabrication d'une couveuse oeuf), le navigateur utilisé par le client (...Safari ...).

```
41.73.108.18 - - [10/Mar/2016:22:58:07 +0100] "GET /collect/upb/index/assoc/IDR-2011-YOD-ETU/IDR-2011-YOD-ETU.pdf HTTP/1.1" 200 3926290
"http://www.google.ml/url?sa=t&source=web&cd=4&rct=j&q=l'etude de fabrication d'une couveuse oeuf &ved=0ahUKEwirhKPobjfLahWKSROKHZA0AwYQFgggMAM &url=http://www.beep.ird.fr/collect/upb/index/assoc/IDR-2011-YOD-ETU/IDR-2011-YOD-ETU.pdf &usg=AFQjCNEell9pi1qREc9esLaRZKUI3z0sXA" "Mozilla/5.0 (Linux; U; Android 4.1.2; fr-fr; itel it1451 Build/JZO57K) AppleWebKit/534.30 (KHTML, like Gecko) Version/4.0 Mobile Safari/534.30"
```

FIGURE 1 – Exemple d'une ligne de transaction du fichier de journalisation du serveur Apache, format combiné

Pour réaliser notre étude nous avons géolocalisé les adresses IP en utilisant la base de données GeoLite Country de la société MaxMind⁴.

Le fichier d'historique est filtré en ne prenant en considération que les lignes qui comportent des accès aux fichiers **pdf** avec le statut « 200 »⁵. Le fichier qui en résulte est ensuite décomposé en plusieurs sous fichiers. Le champ qui contient le type de navigateur intègre également la signature des robots de recherche. Ceci permet de générer des fichiers de log relatifs à l'indexation qui est faite par les moteurs de recherche, avec en particulier le moteur de recherche de Google (fig. 2).

```
66.249.69.91 - - [01/Jan/2016:21:03:10 +0100] "GET /collect/uouaga/index/assoc/M08549.dir/M08549.pdf HTTP/1.1" 200 2367867 "-"
"Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
```

FIGURE 2 – Exemple d'une ligne de transaction concernant l'indexation d'un fichier pdf par le robot de Google (Googlebot/2.1)

Sont ensuite extraites toutes les lignes présentes dans le fichier de log qui comportent le champ du site depuis lequel le client a lancé sa requête (referer), ou le champ du navigateur utilisé, remplis par un tiret (« - ») : nous présumons que les requêtes qui cachent ces deux champs peuvent être assimilées à des DoS attack⁶ ou des spam referer⁷. Selon la même logique, nous écartons également les accès venant de la même adresse IP pour consulter le même fichier qui se réalisent dans un intervalle inférieur à 5 minutes.

Après ces différentes étapes de filtrage (KAUR et AGGARWAL 2017), nous disposons d'un fichier qui se compose (pour la plus grande partie) de « human access ». Il est prêt pour produire les statistiques et les indicateurs de consultations. Le fichier de travail filtré qui contient tous les accès vers toutes les collections du serveur BEEP se compose de 629.196

4. Les versions gratuites de GeoLite country et de GeoLite city sont en libre accès à : dev.maxmind.com/geoip/legacy/geoip/

5. Voir : tools.ietf.org/html/rfc2616#page-58. Le statut « 200 » indique le succès de la requête. Avec la méthode GET une entité correspondant à la ressource demandée est envoyée avec la réponse.

6. Voir à ce propos : fr.wikipedia.org/wiki/Attaque_par_déni_de_service

7. Voir à ce propos : fr.wikipedia.org/wiki/Referer_spam

lignes soit 39 % du fichier brut de départ. Le fichier des robots représente 42 % et celui des dos-spam 19 % par rapport au fichier brut de départ.

3 Distribution temporelle des consultations

A partir du fichier des « human access », nous avons réalisé l'extraction des consultations pour chaque collection hébergée sur le serveur. Les consultations annuelles des deux collections des Universités burkinabé (2BKF) sont présentées dans le tableau 1.

Collection	Documents	Accès 2016	Moyenne/mois
eismv	1246	104626	8719
upb	922	75675	6306
uouaga	493	68697	5725
inseps	974	56755	4730

TABLE 1 – Consultations 2016 et moyennes mensuelles

Pour mieux situer ces données, nous présentons également les statistiques d'accès de la collection des thèses et mémoire de l'École Inter-Etats des Sciences et Médecine Vétérinaires de Dakar (eismv) et de l'Institut National Supérieur de l'Éducation Populaire et du Sport de Dakar (inseps)⁸.

Pour les collections 2BKF le nombre de consultations annuelles varient entre 68.697 et 75.675 avec des moyennes mensuelles allant de 5.725 à 6.306.

La répartition mensuelle des consultations est présentée dans la figure 3. On remarque une allure assez caractéristique de la distribution avec un creux au mois de janvier, une progression sensible au cours des mois d'avril et mai, un creux pour les mois de l'été et une forte progression pour le mois de novembre. Cette distribution des consultations au cours de l'année est assez emblématique des activités des étudiants avec la préparation de leurs mémoire au printemps, une pause estivale et une reprise avec la rentrée de l'automne. Il est intéressant de remarquer que la distribution des consultations des documents de la collection de l'Université de Bobo Dioulasso ne semble pas vraiment suivre cette distribution. En effet, les consultations progressent au printemps et sont constantes au cours de l'été. Ainsi pour la collection upb la distribution des consultations est globalement assez constante au cours de l'année.

Il est intéressant d'analyser la distribution des consultations ayant comme « referer » les sites de Google. Le referer est le site depuis lequel le client a lancé sa requête. Sur la figure 4 on remarque que les taux de consultations ayant comme origine une requête dans Google varient entre 85 et 95 %. Plus particulièrement, les collections inseps et uouaga ont des pourcentages des requêtes ayant comme referer Google de l'ordre de 95 %. Par contre les collections upb et eismv ont des pourcentages des requêtes ayant comme referer Google de l'ordre de 85 %. Cette différence est probablement due au fait que les utilisateurs des collections upb et eismv utilisent, relativement souvent, un autre referer. Il s'agit de l'interface de consultation

8. Pour ces institutions voir : www.eismv.org/ et inseps.ucad.sn/

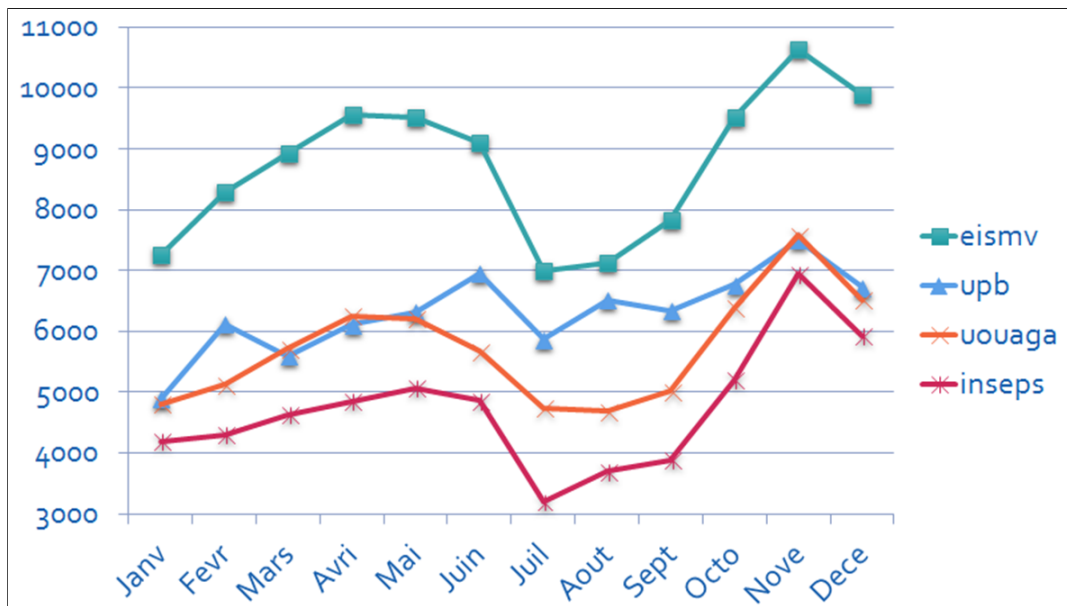


FIGURE 3 – Réparation mensuelle des consultations pour l'année 2016

du site BEEP qui leur est conseillée par les bibliothécaires des institutions dont relèvent les documents.

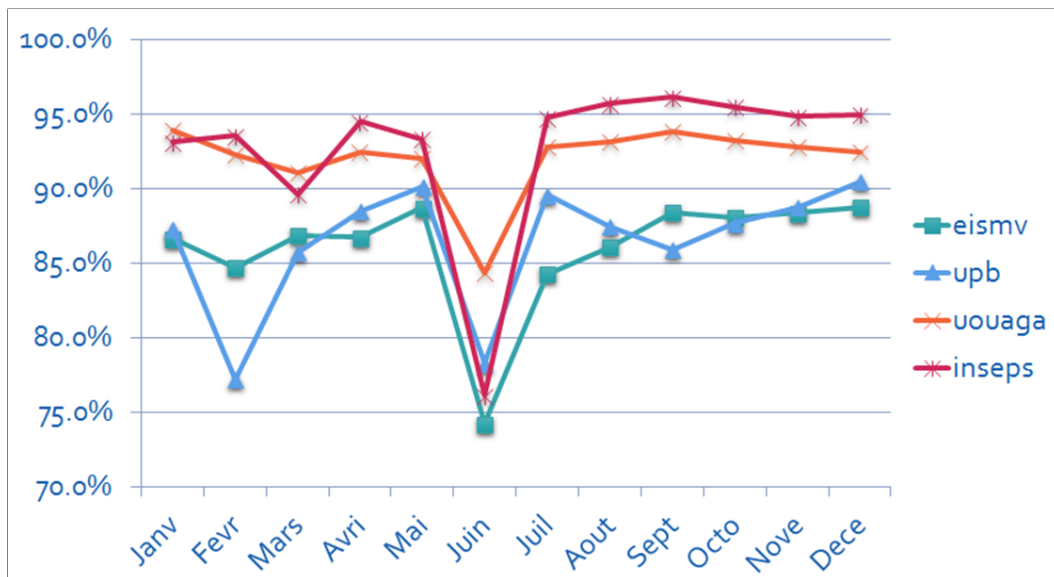


FIGURE 4 – Répartition mensuelle du referer Google

Sur le graphique de la figure 4 on remarque un creux au mois de juin pour toutes les collections. L'analyse détaillée des données montre que le 3 juin un utilisateur, dont l'adresse IP est localisée à Madagascar, a téléchargé tous les documents du site BEEP (aspiration complète du site, fichier par fichier). Le creux de la collection uouaga est moins accentué par rapport à celui des autres collections puisque le nombre de documents de cette collection est deux fois moins important que celui des autres collections. Le creux de la collection upb au mois

de février est du a un déchargement de 700 fichiers de cette collection par une adresse IP localisée au Burkina Faso.

4 Distribution géographique des consultations

Dans la mesure où tous les accès ont été géolocalisés, il est possible de produire des statistiques des consultations en fonction de la distribution géographique.

Pour la collection upb il y a eu en 2016 75.674 consultations qui sont localisées dans 155 pays différents (Fig. 5). Les pays de l'Afrique représentent 83 % des consultations. Le Burkina Faso est le premier pays où sont localisées les consultations (16 %) ce qui s'explique par rapport au contenu des documents (qui traitaient de problématiques scientifiques localisées au Burkina Faso) et par le fait que beaucoup d'utilisateurs de l'Université de Bobo Dioulasso sont invités à consulter les documents de cette collection par les bibliothécaires de cette institution, via l'interface du site BEEP. Pour les consultations africaines, les pays représentés sont essentiellement les pays francophones puisque la plus grande partie des documents de la collection sont écrits en français.

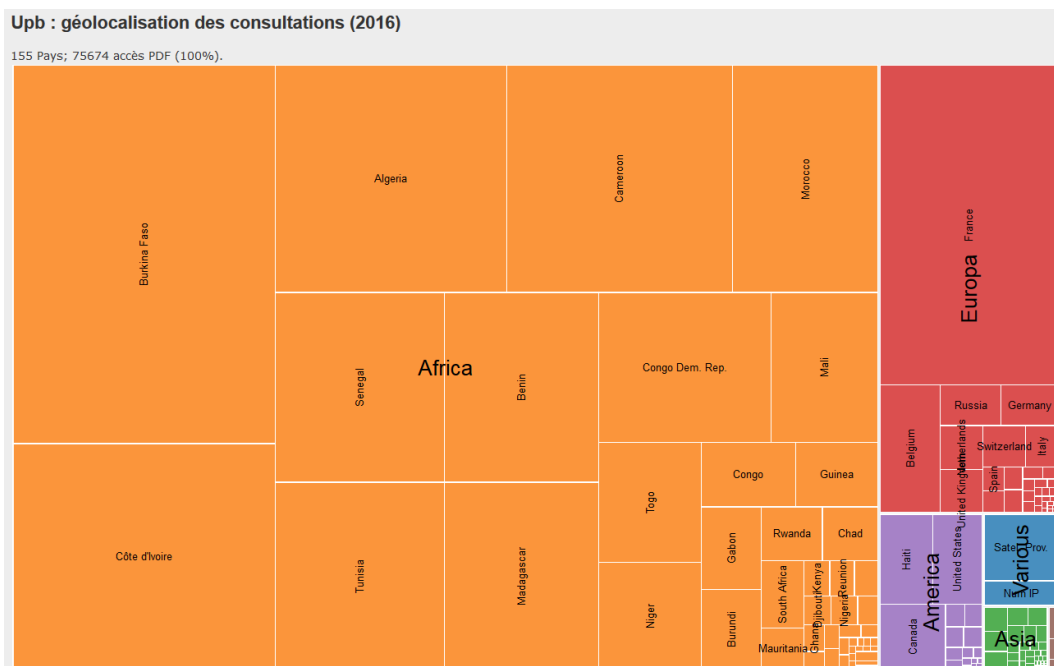


FIGURE 5 – Géolocalisation des consultations de la collection upb

L'Afrique du Sud est le premier pays anglophone qui consulte les documents de la collection upb (0,5 %). Les consultations localisées en France sont de l'ordre de 9 % (premier pays de l'Europe) et celles localisées à Haïti sont de l'ordre de 0,8 % (premier pays des Amériques).

Pour la collection uouaga il y a eu en 2016 68.697 consultations qui sont localisées dans 154 pays différents (Fig. 6). Les pays de l'Afrique représentent 77 % des consultations. L'Algérie est le premier pays où sont localisées les consultations (15 %). Comme pour la collection upb, les pays représentés sont essentiellement les pays francophones puisque la plus grande

partie des documents de la collection sont écrits en français.

L'Afrique du Sud est le premier pays anglophone qui consulte les documents de la collection uouaga (1,1 %). Les consultations localisées en France sont de l'ordre de 13 % (premier pays de l'Europe) et celles localisées à Haïti sont de l'ordre de 1 % (premier pays des Amériques).

Pour l'Asie (1,7 %) on observe que les pays les plus représentés sur le graphe sont des pays francophones (Cambodge, Liban, Vietnam).

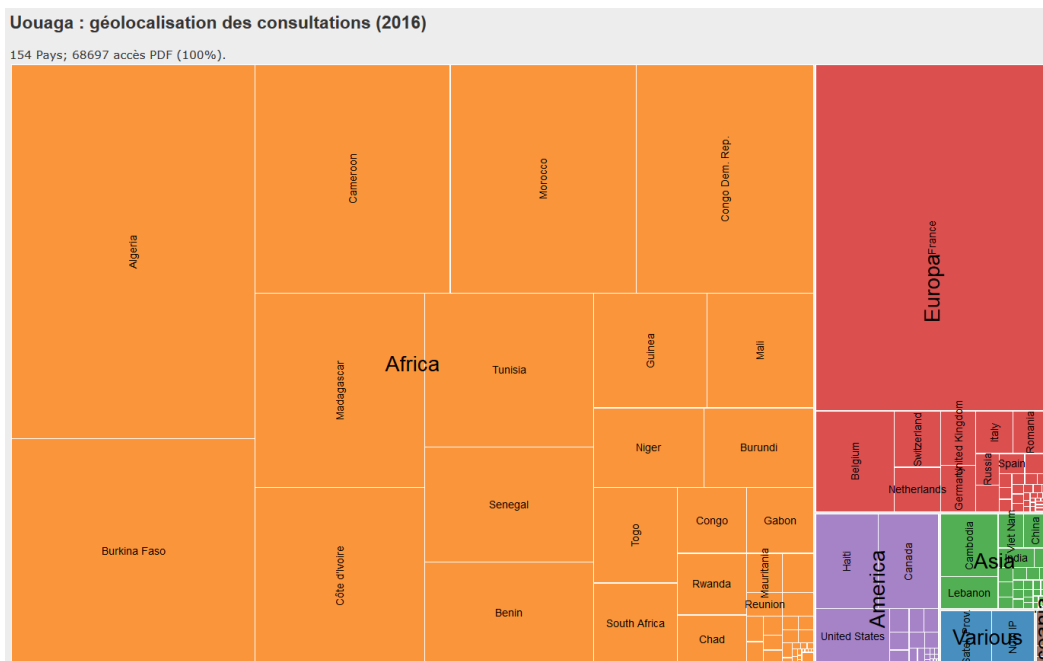


FIGURE 6 – Géolocalisation des consultations de la collection uouaga

5 Questions posées par les utilisateurs et text mining

A partir des fichiers de log il est possible de récupérer les requêtes que les utilisateurs posent avant d'atteindre les documents présents sur le serveur. Les utilisateurs posent leurs questions soit sur les pages des moteurs de recherche soit sur les pages de recherche du serveur BEEP. Pour la collection upb nous avons pu récupérer 10.922 questions et pour la collection uouaga nous avons pu récupérer 12.224 questions.

Pour analyser le texte des questions, nous avons développé un script en php qui intègre l'outil d'annotations de textes TreeTagger⁹ (SCHMID 1997). Après l'annotation des textes, nous ne conservons, pour chaque question, que les verbes à l'infinitif, les noms et adjectifs avec une forme au singulier masculin.

Après les traitements réalisés avec TreeTagger, il est possible de calculer les fréquences des mots présents dans les questions. Pour la collection upb le mot le plus fréquent est « mémoire » et pour la collection uouaga le mot le plus fréquent est « thèse » ce qui correspond à la

9. Cet outil est en libre accès à l'adresse : www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/

typologie des documents. L'utilisateur définit dans ses questions le type de document qu'il recherche (une thèse, un mémoire) en y associant l'expression de ses besoins (des mots caractérisant le sujet de sa recherche).

Pour mieux caractériser les questions des utilisateurs nous avons développé un script en php qui calcule les co-occurrences des mots dans chaque question pour l'ensemble des questions se rapportant à une collection. Les graphes des co-occurrences générés avec Gephi (BASTIAN, HEYMANN et JACOMY 2009) permettent de représenter les résultats de ces analyses. La taille des nœuds est proportionnelle à la fréquence d'occurrences. Les nœuds sont regroupés dans des clusters calculés en fonction des classes de modularité et la spatialisation est obtenue avec l'algorithme « force atlas » (JACOMY et al. 2014).

La figure 7 montre le graphe des co-occurrences de mots pour les questions se rapportant à la collection upb¹⁰.

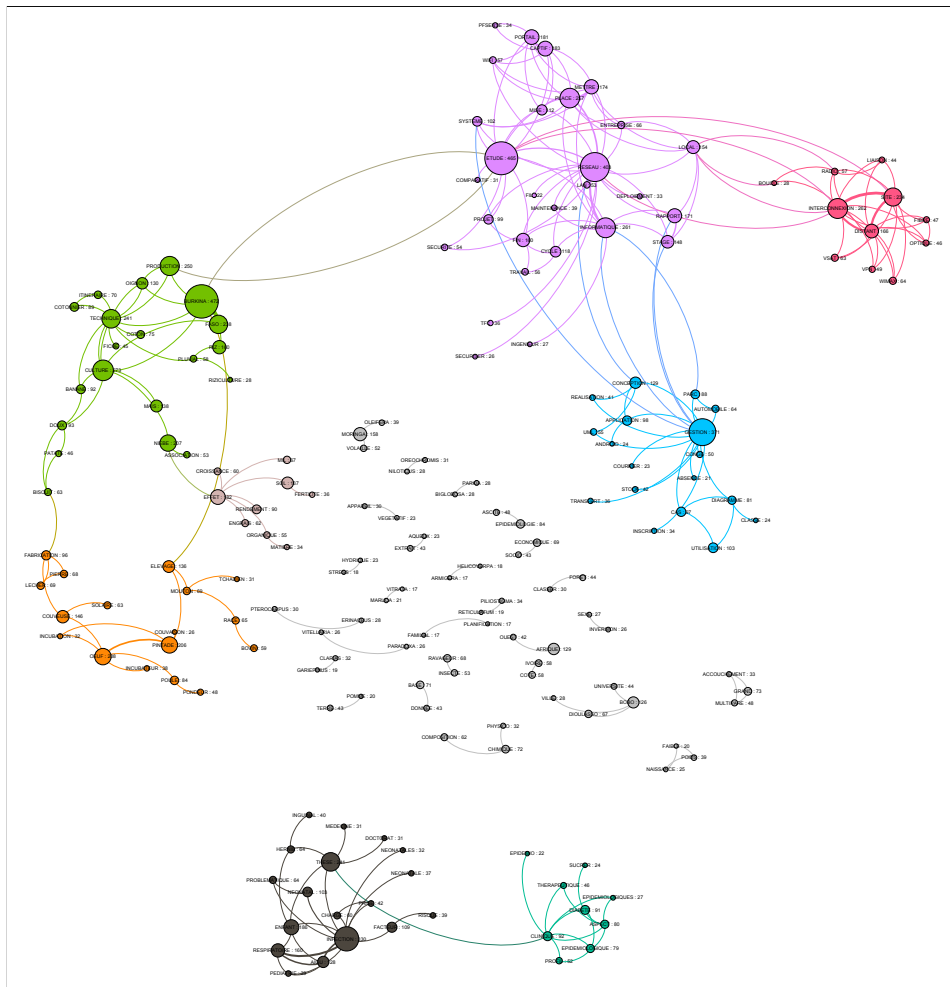


FIGURE 7 – Co-occurrences de mots pour les questions se rapportant à la collection upb

10. La version dynamique du graphe est en libre accès à : www.beep.ird.fr/ecas7/gexf-js/index.html#upb.gexf. Pour l'outil gexf-js voir : github.com/raphv/gexf-js

Lors du calcul du graphe les mots « mémoire » et « pdf » ont été filtrés. Les clusters violet, magenta et cyan, en haut à droite, représentent des questions relatives aux documents de l'Ecole supérieure d'informatique. Le cluster cyan illustre les questions concernant la gestion informatique, le cluster magenta illustre les questions concernant l'interconnexion à distance des réseaux informatiques et le cluster violet illustre les questions concernant la sécurité et la maintenance des réseaux informatiques. Les clusters vert et orange à gauche représentent des questions relatives aux documents de l'Institut de développement rural. Le cluster vert illustre les questions concernant la culture du maïs, du niébé, de l'oignon au Burkina Faso et le cluster orange illustre les questions concernant l'élevage des pintades. Les deux clusters du bas du graphe représentent des questions relatives aux documents de médecine, pédiatrie et concernant le diabète.

La figure 8 montre le graphe des co-occurrences de mots pour les questions se rapportant à la collection uouaga¹¹.

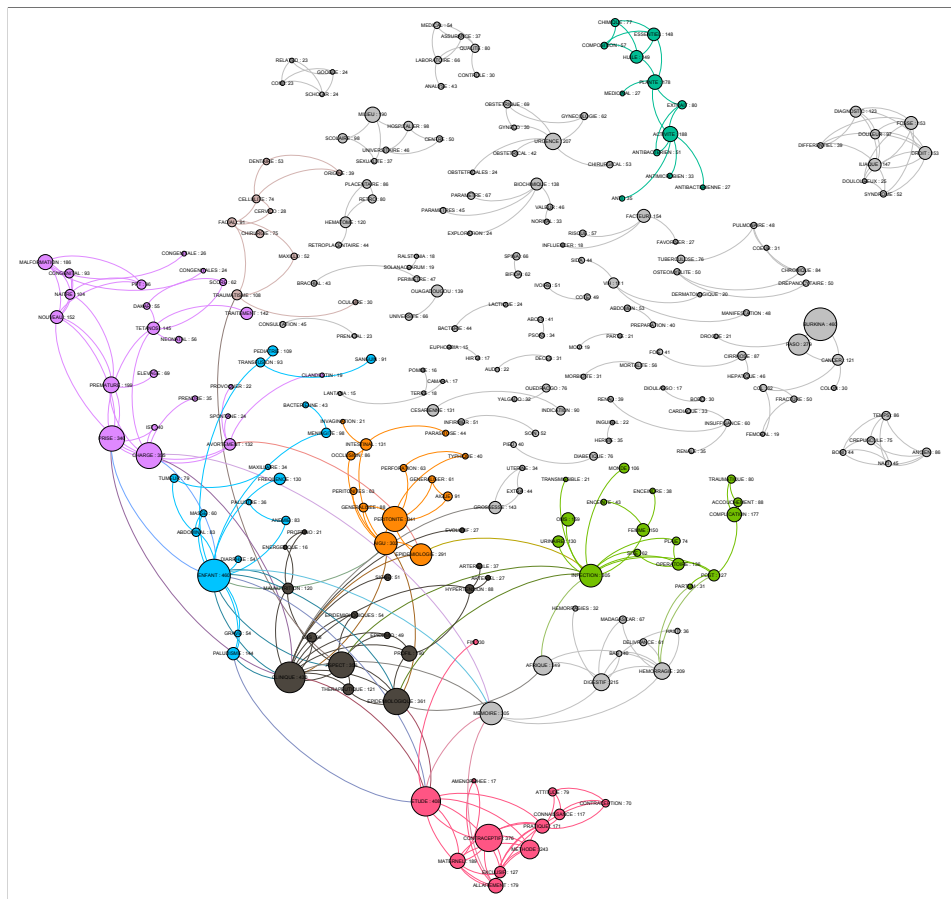


FIGURE 8 – Co-occurrences de mots pour les questions se rapportant à la collection uouaga

Lors du calcul du graphe, les mots « thèse » et « pdf » ont été filtrés. La plupart des clusters concernent des questions relatives à la médecine et ses différentes thématiques et problématiques. Le cluster vert, en haut du graphe, se détache de ces thématiques et concerne

11. La version dynamique du graphe est en libre accès à : www.bep.ird.fr/ecas7/gexf-js/index.html#uouaga.gexf

les plantes médicinales et les huiles essentielles. Les clusters les plus significatifs concernent la prise en charge médicale (cluster violet), la pédiatrie au sens large (cluster cyan), les péritonites (cluster orange), la médecine clinique (cluster noir), la contraception (cluster magenta).

6 Indicateurs d'impact des collections

En associant les tables des données de consultation et les tables des métadonnées des documents des collections avec une base de données relationnelles sous MySQL il est possible de générer dynamiquement, avec un script PHP que nous avons développé, des tableaux de bord proposant plusieurs indicateurs d'impact.

Nous avons retenu différents indicateurs : le nombre de fichiers de la collection consultés, le nombre total de consultations des documents de la collection, le nombre moyen de consultations par fichier de la collection, le facteur D (nombre de documents consultés au moins D fois), la somme des consultations des Top D, les consultations Top D sur le total des consultations des documents de la collection. De plus, le tableau de bord présente les documents les plus consultés (jusqu'au rang D) avec leur nombre de consultations et le premier pays consultant en indiquant son indice d'appropriation (nombre de consultations du premier pays consultant / nombre total des consultations). Les tableaux de bord présentent également la géolocalisation des consultations et un graphique basé sur la loi de Pareto¹² pour identifier le pourcentage de documents nécessaire à atteindre 80 % des consultations.

Le tableau 2 résume les principaux indicateurs de deux collections¹³. Pour les deux collections pratiquement tous les documents ont été consultés au moins une fois au cours de l'année 2016 (920/922 et 492/493). Le nombre moyen des consultations par document et le facteur D sont plus importants pour la collection uouaga : ceci pourrait s'expliquer par le fait qu'il s'agit de thèses avec une thématique assez cohérente (essentiellement des thèses en Médecine). L'indice de Pareto est, pour les deux collections, supérieur à 20 % ce qui indique une distribution des consultations plus large que la normale. Pour la collection upb ce sont les documents de l'École supérieure d'informatique qui « captent » le plus grand nombre de consultations ce qui détermine aussi un indice de Pareto moins important que celui de la collection uouaga.

Pour avoir une vision comparative des indicateurs des deux collections des Universités du Burkina Faso, nous avons pris en considération les indicateurs des documents de l'IRD qui sont en libre accès (environ 61.000 en 2016). L'ensemble de ces documents peut être reparti en fonction d'un code thématique qui est associé à chaque document, ce qui permet de construire des « collections virtuelles » identifiées par grands domaines. Le tableau complet des ces indicateurs est disponible en libre accès¹⁴. Un indicateur qui peut permettre de faire des comparaisons entre les différentes collections nous paraît être le nombre moyen de consultations par fichier. On peut ainsi constater que les moyennes de consultations des

12. Voir à ce propos : fr.wikipedia.org/wiki/Principe_de_Pareto

13. Les deux tableaux de bord complets sont en libre accès à : www.beep.ird.fr/ecas7/upb.htm et www.beep.ird.fr/ecas7/uouaga.htm

14. Le tableau des indicateurs de consultations des documents en libre accès de l'IRD est accessible à : www.beep.ird.fr/ecas7/chap-2016.htm. A la fin du tableau se trouve la synthèse des indicateurs, en fonction des thématiques retenues.

Indice	upb	uouaga
Nombre total de fichiers de la collection consultés	920	492
Nombre de consultations de la collection	75674	68697
Nombre moyen de consultations par fichier	82.3	139.6
Facteur D (nombre de documents consultés au moins "n" fois)	125	142
Somme des consultations des Top D	50344	48780
Consultations Top D sur le total des consultations	66.5 %	71.0 %
Pareto	25.1 %	39.0 %

TABLE 2 – Principaux indicateurs des deux collections

meilleures collections des documents de l'IRD sont comparables à celle de l'upb. Toutefois, la moyenne des consultations des documents de la collection uouaga est plus importante par rapport aux meilleures moyennes de consultations des collections de l'IRD. Le facteur D des collections de l'IRD est comparable à celui des collections upb et uouaga en considérant des collections ayant un nombre de documents similaire.

7 Conclusions

Plusieurs institutions de pays d'Afrique mettent leurs productions scientifiques en libre accès sur internet. Il s'agit le plus souvent d'institutions présentes dans les pays d'Afrique anglophone (Afrique du Sud, Kenya, Nigeria, ... (MUTIIS et KITCHEN 2016; BARO, GODFREY et EZE 2014; MOLTENO 2016)). Ces dynamiques concernent également certains pays d'Afrique francophone, comme le Burkina Faso et le Sénégal (KANE 2014).

Le constat que nous pouvons faire par rapport à la mise en libre accès des mémoires et des thèses des Universités du Burkina Faso est que les utilisateurs d'internet consultent de façon régulière et significative les ressources mises à disposition. Ces utilisateurs sont en partie ceux qui utilisent les interfaces de consultation développées au niveau des sites (5 à 10 % des consultations dues essentiellement aux étudiants de ces institutions) et, pour la plupart, des utilisateurs des moteurs de recherche tels que Google (environ 90 % des consultations).

S'agissant de documents écrits en français, les localisations géographiques des utilisateurs se situent dans les pays francophones, essentiellement les pays d'Afrique (89 %).

L'analyse des données concernant les consultations montre que certaines thématiques (les réseaux informatiques, l'interconnexion des sites, la gestion informatique, les huiles essentielles) intéressent de façon significative les internautes.

La comparaison des indicateurs et des volumes de consultations avec des collections d'autres institutions scientifiques comme celles de l'IRD montre des résultats tout à fait semblables, voire des moyennes de consultation plus importantes pour les documents produits par les Universités du Burkina Faso. Nous espérons que ces constats pourront encourager les institutions des pays francophones à poursuivre ces stratégies de valorisation de leurs propres productions scientifiques.

Références

- BARO, E. E., GODFREY, V. Z. et EZE, M. E. (2014). Electronic Theses and Dissertations Initiatives : A Survey of University Libraries in Nigeria. In : *International Information & Library Review* 46.1-2, p. 41–50. DOI : [10.1080/10572317.2014.936265](https://doi.org/10.1080/10572317.2014.936265).
- BASTIAN, M., HEYMANN, S. et JACOMY, M. (2009). Gephi : an open source software for exploring and manipulating networks. In : *Proceedings of the Third International ICWSM Conference*, p. 361–362. URL : <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/download/154/1009>.
- JACOMY, M., VENTURINI, T., HEYMANN, S. et BASTIAN, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. In : *PloS one* 9.6, e98679. DOI : [10.1371/journal.pone.0098679](https://doi.org/10.1371/journal.pone.0098679).
- KANE, K. (2014). *Réalités sur l'accès à l'information scientifique numérique dans les bibliothèques des universités du Sénégal : l'exemple de l'université de Dakar*. CILA 2014, 27-28 novembre 2014, Tunis. URL : icoa2014.sciencesconf.org/37928/document.
- KAUR, N. et AGGARWAL, H. (2017). A novel semantically-time-referrer based approach of web usage mining for improved sessionization in pre-processing of web log. In : *International Journal of Advanced Computer Science and Applications* 8.1, p. 158–168.
- MERCEUR, F. (2015). *Statistiques d'usage des documents accessibles en Libre Accès via Archimer. Période 2005-2014*. URL : archimer.ifremer.fr/doc/00251/36224/37709.pdf.
- MOLTENO, R. (2016). Digital repositories – making Africa's intelligentsia visible ? In : *Bulletin of the National Library of South Africa* 70.2, p. 167–182. URL : <https://www.internationalafricaninstitute.org/downloads/Molteno-Making%20Africa's%20Intelligensia%20Visible.pdf>.
- MUTIIS, A. de et KITCHEN, S. (2016). African Digital Research Repositories : Survey Report. In : *Africa Bibliography* 2015, p. vii–xxv. DOI : [10.1017/S0266673116000027](https://doi.org/10.1017/S0266673116000027).
- ROSSI, P. L. (2011). Electronic libraries in partnership : BEEP for Africa. In : *African Research & Documentation* 115, p. 69–75. URL : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers11-11/010052810.pdf.
- (2013). *Mise en place d'ateliers de numérisation et création de bibliothèques électroniques en Afrique francophone : transfert de compétences et partenariats*. Ifrane (Maroc), ICADLA 3, 27-31 Mai 2013. URL : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers13-06/010059325.pdf.
- (2017). *Numérisation et bibliothèques électroniques : bilan de 15 années de formations réalisées par l'IRD au bénéfice des professionnels de pays de l'Afrique francophone*. EDAD, Colloque du cinquantenaire, 2017/11/14-16, Dakar. URL : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers17-12/010071833.pdf.
- SCHMID, H. (1997). Probabilistic part-of-speech tagging using decision trees. In : *New methods in language processing*, p. 154–164. URL : <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/tree-tagger1.pdf>.
- SCHÖPFEL, J. et PROST, H. (2010). *Les statistiques d'utilisation d'archives ouvertes-Etat de l'art*. URL : https://hal.archives-ouvertes.fr/sic_00480538/document.

Liens internet pour les tableaux de bord et les graphes de co-occurrences de mots

Tableau de bord de la collection upb (2016) : www.beep.ird.fr/ecas7/upb.htm

Tableau de bord de la collection uouaga (2016) : www.beep.ird.fr/ecas7/uouaga.htm

Tableau de bord des collections de l'IRD (2016) : www.beep.ird.fr/ecas7/chap-2016.htm

Réseau des co-occurrences de mots pour les questions se rapportant à la collection upb (2016) : www.beep.ird.fr/ecas7/gexf-js/index.html#upb.gexf

Réseau des co-occurrences de mots pour les questions se rapportant à la collection uouaga (2016) : www.beep.ird.fr/ecas7/gexf-js/index.html#uouaga.gexf