Création d'un tableau de bord pour l'Archive ouverte UNIGE

Travail de Bachelor réalisé en vue de l'obtention du Bachelor HES

par :
Cédric PELLA

Conseiller au travail de Bachelor :
Patrick RUCH, Professeur HES

Genève, 15 juillet 2010

Haute École de Gestion de Genève (HEG-GE)

Filière Information documentaire

Déclaration

Ce travail de Bachelor est réalisé dans le cadre de l'examen final de la Haute école de gestion de Genève, en vue de l'obtention du titre de Bachelor en sciences HES en information documentaire. L'étudiant accepte, le cas échéant, la clause de confidentialité. L'utilisation des conclusions et recommandations formulées dans le travail de Bachelor, sans préjuger de leur valeur, n'engage ni la responsabilité de l'auteur, ni celle du conseiller au travail de Bachelor, du juré et de la HEG.

« J'atteste avoir réalisé seul le présent travail, sans avoir utilisé des sources autres que celles citées dans la bibliographie. »

Fait à Genève, le 15 juillet 2010

Cédric Pella

Remerciements

Ce travail de bachelor clôt une formation de trois ans à la HEG de Genève. Je sais gré à mes professeurs et enseignants pour les connaissances et compétences qu'ils ont transmises, à mes collègues étudiants pour les travaux de groupe et les bons moments partagés, à mes proches et amis pour leur affection et leur soutien tout au long de cette période.

Je remercie les personnes suivantes qui ont eu une importance particulière dans la réalisation de ce travail :

M. Patrick RUCH, professeur et conseiller pédagogique, pour sa liberté de ton et la justesse de ses analyses.

Mme Véronique HADENGUE-DEZAËL, directrice de l'information scientifique à l'Université de Genève, pour m'avoir accueilli dans l'équipe et confié ce mandat.

M. Jean-Blaise CLAIVAZ, coordinateur de l'Archive ouverte UNIGE, pour sa disponibilité et ses relectures attentives.

Mme Isabelle DE KAENEL, directrice de la bibliothèque du CHUV et jurée, pour son intérêt pour mon travail et son expertise du sujet.

M. Jan MELICHAR, informaticien de développement, pour sa vue d'ensemble sur le projet Archive ouverte UNIGE et son expertise technique.

M. Benoît COLLÉ, informaticien de développement, pour son implication dans ce travail et la réalisation de BaDonAo.

Sommaire

Le mandat qui fonde ce travail est la création d'un tableau de bord pour le pilotage à long terme du projet Archive ouverte UNIGE. Derrière ce nom, se cache le dépôt d'archives électroniques de l'Université de Genève.

Les objectifs généraux suivants ont été identifiés pour mener à bien le mandat :

- Connaître les enjeux et les orientations stratégiques de l'Archive ouverte UNIGE.
- Identifier des indicateurs adaptés aux services liés à des ressources numériques.
- Déterminer quels indicateurs sont les plus pertinents dans le contexte local et comment collecter les données.
- Concevoir et mettre en œuvre le tableau de bord.
- Recommander des actions correctrices pour des valeurs anormales

La première étape du travail a été réalisée par des entretiens et la consultation de sources sur le thème du libre accès. Une enquête auprès d'usagers de l'Archive ouverte UNIGE a été menée pour avoir une bonne idée des enjeux du service.

Les deux étapes suivantes se sont appuyées sur le partage d'expérience avec des responsables de dépôts d'archives électroniques et une revue de la littérature disponible sur le sujet.

La conception du tableau de bord relevait quant à elle de connaissances plus techniques.

Le tableau de bord a été mis en œuvre pendant la deuxième quinzaine de juillet ; cette étape n'est donc pas discutée dans ce mémoire.

Les recommandations d'actions correctrices n'ont pas rencontré formellement de résultat à cause d'un manque de recul sur le sujet.

Les points les plus intéressants, en terme de résultats, sont une recherche théorique sur des indicateurs adaptés aux dépôts institutionnels et le système de collecte des données.

Table des matières

Introduction **1**

1. Le mandat **2**

1.1 Le mandat **2**

1.2 Le contexte **3**

1.2.1 L'Université de Genève *3*

1.2.2 La Direction de l'information scientifique *4*

1.2.3 La Division informatique *4*

1.3 L'Archive ouverte UNIGE **5**

1.3.1 L'historique du projet *5*

1.3.2 Les ressources et contraintes *6*

1.3.3 Les usagers du service *7*

1.3.4 Les fonctionnalités de l'outil *9*

1.4 Le libre accès **10**

1.4.1 Historique *10*

1.4.2 Le libre accès en Suisse en 2010 *10*

1.4.3 Statistiques dans quelques archives électroniques institutionnelles *12*

2. Méthodologie **13**

2.1 Connaître les enjeux et les orientations stratégiques de l'Archive ouverte UNIGE **13**

2.1.1 Recherche documentaire *13*

2.1.2 Entretiens avec les responsables du projet Archive ouverte *13*

2.1.3 Entretiens avec les usagers *14*

2.1.4 Choix du tableau de bord *14*

2.2 Identifier des indicateurs adaptés aux services liés à des ressources numériques **15**

2.2.1 Recherche documentaire *15*

2.2.2 Partage d'expérience entre collègues *15*

2.3 Déterminer quels indicateurs sont les plus pertinents dans le contexte local et comment collecter les données **15**

2.4 Concevoir et mettre en œuvre le tableau de bord **15**

2.5 Recommander des actions correctrices pour des valeurs anormales **16**

2.5.1 Fixer des exigences de qualité pour les indicateurs *16*

2.5.2 Connaître les actions qui peuvent influer sur les indicateurs *16*

3. Réalisation des objectifs généraux **17**

3.1 Connaître les enjeux et les orientations stratégiques du service Archive ouverte UNIGE **17**

3.1.1 Entretiens avec les usagers *17*

3.1.2 Choix du tableau de bord *21*

3.2 Identifier des indicateurs adaptés aux services liés à des ressources numériques 22

3.2.1 Recherche documentaire 22

3.2.2 Liste d'indicateurs applicables dans le cas de ressources numériques 23

3.3 Déterminer quels indicateurs sont les plus pertinents dans le contexte local et comment collecter les données 29

3.3.1 Principes de construction des indicateurs 29

3.3.2 Liste des indicateurs retenus pour le tableau de bord 29

3.4 Concevoir et mettre en œuvre le tableau de bord 34

3.5 Recommander des actions correctrices pour des valeurs anormales 35

Conclusion 36

Critique du travail de bachelor 36

Pistes à suivre 37

Recommandations 38

Situation en Suisse 38

L'Archive ouverte UNIGE 39

Bibliographie 41

Liens vers les dépôts d'archives électroniques du tableau 1 42

Annexe 1 : Spécifications BaDonAo 44

Liste des Tableaux

Tableau 1 Quelques dépôts d'archives électroniques en Suisse et en Europe 11
Tableau 2 Cadrage du tableau de bord..... 21
Tableau 3 Sélection d'indicateurs identifiés pour les ressources numériques.. 24
Tableau 4 Liste des indicateurs retenus pour le tableau de bord..... 30

Introduction

Ce document constitue le mémoire de travail de bachelor qui clôt trois ans d'études à la Haute école de gestion de Genève pour devenir spécialiste en information documentaire. Il est disponible sous forme électronique à cette adresse :

http://www.box.net/shared/2pl2gepzyh.

Le mandat de ce travail de bachelor est de créer un tableau de bord pour le pilotage à long terme de l'Archive ouverte UNIGE. Ce nom correspond au dépôt d'archives électroniques de l'Université de Genève.

Ce travail est donc en relation avec le mouvement du libre accès (Open Access) en Suisse. Il permet de comprendre un peu mieux les enjeux actuels de ce mouvement. En particulier les stratégies mises en œuvre pour améliorer l'implication des auteurs dans le dépôt de leurs publications.

Ce mémoire présente aussi un travail de recherche sur les indicateurs qui ont du sens pour évaluer la performance d'un dépôt institutionnel. Cet aspect est encore relativement peu connu de la littérature professionnelle, il y a donc une chance pour que ce travail fasse avancer la connaissance scientifique.

Dans un troisième temps, ce mémoire délivre les spécifications pour un système semi-automatisé de collecte des mesures statistiques. Ce système a été préféré au tableau de bord initialement demandé, car il permet une plus grande flexibilité dans le temps et s'intègre dans un système d'information déjà existant.

La réalisation de ce travail s'appuie sur les connaissances et compétences acquises durant le cursus de formation et un approfondissement de certains sujets par des recherches personnelles.

1. Le mandat

Ce chapitre explicite le mandat rattaché à ce travail de bachelor. De plus, comme le contexte du service Archive ouverte UNIGE et la connaissance du mouvement du libre accès peuvent aider à comprendre les enjeux de ce travail, ils sont aussi exposés ici.

Voici donc les subdivisions de ce chapitre :

- le mandat de ce travail de bachelor
- le contexte du projet Archive ouverte UNIGE
- l'Archive ouverte UNIGE
- le libre accès

1.1 Le mandat

Le mandat de ce travail de bachelor est de créer un tableau de bord pour le pilotage à long terme de l'Archive ouverte UNIGE. Il en découle les objectifs généraux suivants :

- Connaître les enjeux et les orientations stratégiques de l'Archive ouverte UNIGE
- Identifier des indicateurs adaptés aux services liés à des ressources numériques
- Déterminer quels indicateurs sont les plus pertinents dans le contexte local et comment collecter les données
- Concevoir et mettre en œuvre le tableau de bord
- Recommander des actions correctrices pour des valeurs anormales

Le mandat doit être réalisé entre le 22 février et le 9 juillet 2010; pendant cette période, le temps est réparti sur 15 semaines à 60%, 3 semaines à temps plein et 2 semaines de congé, ce qui revient à environ 510 heures de travail pour la réalisation du mandat et la rédaction de ce mémoire. La réalisation technique du tableau de bord échoit à un informaticien de développement affecté au projet par la Division informatique.

L'enjeu de ce mandat est principalement d'identifier les indicateurs pertinents et d'automatiser au maximum la collecte des variables qui y sont reliées, ceci afin de ne pas surcharger la gestion du service. Il y a aussi un deuxième enjeu pour les responsables du projet : obtenir une appréciation qualitative sur l'outil Archive ouverte et ses fonctionnalités par un sondage sur une partie de ses usagers.

1.2 Le contexte

Pour bien cerner le mandat, il est important d'examiner attentivement dans quelle organisation s'inscrit le service. Cette analyse présente brièvement l'histoire et les missions des organismes qui ont encadré l'Archive ouverte UNIGE, elle porte sur les trois structures impliquées dans le projet :

- l'Université de Genève
- la Direction de l'information scientifique
- la Division informatique

1.2.1 L'Université de Genève

L'Académie de Genève a été fondée en 1559 par Jean Calvin et Théodore de Bèze, avec pour principal objectif de former des pasteurs et de propager la Réforme.

Elle se distancie de l'Église au fil du temps et devient l'Université de Genève en 1873. Aujourd'hui sa mission est triple :

- l'enseignement
- la recherche
- le service à la cité[1]

D'après le rapport de gestion 2009, l'Université de Genève compte 5'422 collaborateurs et 14'489 étudiants. Elle a délivré 3'272 diplômes. 10'502 personnes ont suivi ses programmes de formation continue[2]. Ces chiffres sont en relation directe avec la mission de l'Université de Genève; en effet, nombre de ses collaborateurs sont en fait des professeurs ou assistants qui dispensent un enseignement à l'attention des étudiants et/ou mènent une activité de recherche.

Le service à la cité quant à lui est assuré par deux activités :

- la formation continue qui propose 256 programmes
- les transferts de technologies

---

1 UNIVERSITÉ DE GENÈVE. *Site de l'Université de Genève* [en ligne]. http://www.unige.ch/ (consulté le 5 juillet 2010)

2 UNIVERSITÉ DE GENÈVE. *Rapport de gestion 2009* [en ligne]. http://www.unige.ch/apropos/rapport-gestion/RapportGestion2009.pdf (consulté le 8 juillet 2010)

1.2.2 La Direction de l'information scientifique

La bibliothèque est un outil de choix pour aider à remplir les missions de l'Université de Genève. Cette dernière en compte pas moins de 42, d'après un audit interne mené en 2009[3]. Pour faciliter la coopération entre ses bibliothèques, l'Université de Genève s'est dotée d'un service dédié : la Direction de l'information scientifique (DIS). Sa raison d'être est d'améliorer l'accès à l'information scientifique.

> *« Sa mission [celle de la DIS] est double : elle comprend la coordination locale genevoise liée au réseau RERO[4] et les services centralisés aux bibliothèques de l'Université. »* *(Université de Genève, 2009)*[5]

La DIS existe depuis 2007, elle reprend la structure et les missions de l'ancien Service de coordination des bibliothèques (SEBIB) en ce qui concerne RERO et pour la coopération entre les différentes bibliothèques universitaires. Quant à lui, le SEBIB avait été créé en 1982 avec pour principale mission la mise en service du catalogue en ligne géré en partenariat avec la Bibliothèque cantonale et universitaire de Lausanne.

En outre, mandat est donné à la DIS de développer des services pour les bibliothèques, ce qui ouvre la voie à de nouveaux projets communs. L'Archive ouverte UNIGE est précisément l'un de ces projets qui vise à mettre sur pied un nouveau service, en synergie avec les chercheurs et les bibliothèques de l'Université.

En mars 2009, les ressources en personnel de la DIS représentent l'équivalent de 16,4 postes à temps plein.

1.2.3 La Division informatique

La Division informatique a pour mission d'améliorer les services informatiques de l'Université de Genève.

> *« [...] La Division informatique coordonne et contribue à l'évolution du système d'information institutionnel, qui soutient l'enseignement et la recherche ainsi que la gestion administrative de l'Université. Elle met en œuvre des services de*

---

3 MCADAM, Daisy. Y a-t-il des BUGs à l'Université de Genève ? : d'où venons-nous et où allons-nous ?. *Hors-texte. AGBD*. 2009, no 90, p. 7-10.

4 Réseau des bibliothèques de Suisse occidentale. Plus d'information sur le site : http://www.rero.ch (consulté le 14 juillet 2010)

5 UNIVERSITÉ DE GENÈVE. DIRECTION DE L'INFORMATION SCIENTIFIQUE. *Site de la Direction de l'information scientifique* [en ligne]. Modifié le 16 décembre 2009. http://www.unige.ch/biblio/sinformer/dis.html (consulté le 5 juillet 2010)

qualité, basés sur des technologies de pointe et une infrastructure performante [...] » *(Université de Genève, 2009)*[6]

La Division informatique est un partenaire de l'Archive ouverte. Dans le cadre de ce projet, son rôle est d'assurer la maintenance et le développement des outils informatiques sous la conduite de la DIS.

1.3 L'Archive ouverte UNIGE

« L'Archive ouverte UNIGE est le dépôt numérique du patrimoine scientifique de l'Université de Genève. Elle a été créée afin de récolter, conserver et rendre accessibles le plus largement possible les publications des enseignants et des chercheurs de l'institution, en suivant les principes de l'Open Access [libre accès] » *(Université de Genève, 2009)*[7]

L'Archive ouverte UNIGE, ou plus simplement Archive ouverte, est au cœur de ce travail de bachelor. Sa mission est explicitée plus haut. Elle répond à la définition de dépôt d'archives numériques ou dépôt d'archives électroniques institutionnel. C'est à dire qu'elle récolte, conserve et diffuse les publications issues de la recherche scientifique qui a reçu un soutien de l'Université de Genève.

L'Archive ouverte est un service en ligne disponible sur le Web depuis le site de l'Université à l'adresse URL http://archive-ouverte.unige.ch. Pour ce qui est du nombre de documents, en juin 2010, le service a référencé 8'246 publications, dont 6'879 sont accompagnées du texte intégral.

Pour compléter cette présentation de l'Archive ouverte, quatre axes principaux se dégagent:

- l'historique du projet
- les ressources et contraintes
- les usagers du service
- les fonctionnalités de l'outil

1.3.1 L'historique du projet

Le projet Archive ouverte UNIGE a été initié pour répondre à une directive du Fonds national suisse de la recherche scientifique (FNS). Le FNS est la fondation privée la plus importante en Suisse en terme de financement de la recherche fondamentale et

6 UNIVERSITÉ DE GENÈVE. DIVISION INFORMATIQUE. *Site de la Division informatique* [en ligne]. Modifié le 10 novembre 2009. http://www.unige.ch/dinf/home/presentation.html (consulté le 7 juillet 2010)

7 UNIVERSITÉ DE GENÈVE. *Site de l'Archive ouverte UNIGE* [en ligne]. http://archive-ouverte.unige.ch/vital/access/manager/Index (consulté le 7 juillet 2010)

appliquée dans les milieux universitaires. La *Directive sur le libre accès (Open Access) aux publications scientifiques issues de projets encouragés par le FNS*[8] stipule, en substance, que les résultats des projets de recherche auxquels le FNS aura contribué financièrement devront être publiés en libre accès et déposés dans des archives électroniques institutionnelles.

Or ces archives électroniques institutionnelles n'existent pas. La DIS reçoit donc, fin 2007, la mission de mettre sur pied ce qui deviendra l'Archive ouverte UNIGE. En automne 2008, l'Archive ouverte entre en production, les premiers documents disponibles sont des thèses, auxquelles viennent rapidement s'ajouter des articles, puis d'autres types de documents. Actuellement, la collection continue à s'accroître régulièrement et le projet connaît une deuxième phase de développement jusqu'à juin 2011 afin de configurer de nouvelles fonctionnalités et d'améliorer les services proposés par l'outil[9].

1.3.2 Les ressources et contraintes

Cette section comprend une partie sur les ressources humaines affectées à l'Archive ouverte et une sur les logiciels utilisés.

1.3.2.1 Ressources humaines

La DIS délègue une personne à mi-temps au projet Archive ouverte. Ses tâches principales sont de coordonner le projet et de superviser les développements. De son côté la Division informatique affecte une personne à temps plein dont le rôle est d'assurer la maintenance et les développements techniques. La charge de travail devrait cependant être réduite dès juillet 2011, date à laquelle le projet aura terminé sa phase de développement.

Le dépôt de documents et l'attribution de métadonnées sont à la charge des chercheurs eux-mêmes, mais il existe huit correspondants dont le cahier des charges inclut une fonction d'orientation et de conseil sur le versement des documents dans l'Archive ouverte. Par ailleurs, certains professeurs laissent au secrétariat de faculté le

---

8 FONDS NATIONAL SUISSE DE LA RECHERCHE SCIENTIFIQUE. CONSEIL NATIONAL DE LA RECHERCHE. *Directive sur le libre accès (Open Access) aux publications scientifiques issues des projets encouragés par le FNS* [en ligne]. Publié le 4 juillet 2007. http://www.snf.ch/SiteCollectionDocuments/dos_OA_Weisung_f.pdf (consulté le 26 mars 2010)

9 UNIVERSITÉ DE GENÈVE. ARCHIVE OUVERTE UNIGE. *Plus d'info | Archive ouverte UNIGE* [en ligne]. http://archive-ouverte.unige.ch/vital/access/manager/UGProjet (consulté le 26 mars 2010)

soin de s'acquitter de cette tâche administrative. On peut donc estimer que les correspondants et une demi-douzaine d'autres personnes ont une partie variable de leur temps de travail (entre 5% et 20%) dévolue à des tâches diverses en relation directe avec le service Archive ouverte.

1.3.2.2 Logiciels utilisés

Les documents de l'Archive ouverte sont gérés par Fedora[10] (acronyme de Flexible Extensible Digital Object Repository Architecture), un logiciel Open Source développé pour conserver, gérer et diffuser des contenus numériques.

Comme Fedora n'a pas d'interface utilisateur propre, c'est VITAL[11], de la société VTLS, qui assure ce rôle. Les logiciels tournent en parallèle sur la version de production, accessible sur le Web, et une version de développement. La version de VITAL en production est la version 4.1, il existe une version 5.1 vers laquelle l'Archive ouverte devrait migrer d'ici à juin 2011.

La Division informatique a déjà mené plusieurs développements spécifiques pour l'Archive ouverte. Notamment un script pour gérer les droits d'accès en analysant l'adresse IP de l'ordinateur client, et un outil de correction de masse qui permet d'affecter une nouvelle valeur à une métadonnée sur un ensemble de documents.

1.3.3 Les usagers du service

Tout un chacun peut utiliser l'Archive ouverte moyennant une connexion Internet. Un des objectifs du libre accès est une diffusion aussi large que possible, cependant le texte intégral de nombreuses publications de l'Archive ouverte n'est pas disponible pour le grand public parce que leurs auteurs ont cédé les droits de diffusion à des éditeurs commerciaux.

Ainsi dans ce travail de bachelor, la population ou le public cible correspondra aux professeurs, assistants et étudiants de l'Université de Genève. En effet, on peut considérer que l'Archive ouverte s'adresse principalement à eux, usagers primaires de l'Université, qui elle-même finance le service.

Pour comprendre les usagers, il convient d'examiner leur comportement documentaire plutôt que l'appartenance à une catégorie socio-professionnelle. En effet, il est fort

---

10 FEDORA COMMONS. *Site de Fedora* [en ligne]. http://www.fedora-commons.org/ (consulté le 14 juillet 2010)

11 VTLS. *Présentation de VITAL* [en ligne]. http://www.vtls.com/products/vital (consultée le 14 juillet 2010)

probable que la majorité des utilisateurs aient eu accès à une formation académique et effectuent des recherches scientifiques dans tel ou tel domaine.

Dana McKay, dans son étude sur les dépôts d'archives électroniques institutionnels[12] présente trois catégories d'usagers en fonction de leur comportement vis-à-vis de l'outil:

- les auteurs
- les chercheurs d'information
- les gestionnaires des données

Cette section va définir ces trois catégories d'usagers.

1.3.3.1 Les auteurs

Les auteurs ou producteurs d'information sont les chercheurs scientifiques qui ont l'obligation de déposer les résultats de leurs recherches dans l'Archive ouverte. Ils sont employés par l'Université de Genève ou collaborent avec elle. Ils peuvent avoir statut de professeur, d'assistant ou de chercheur indépendant, que ce soit une femme ou un homme.

Les auteurs utilisent des formulaires de dépôt électroniques à accès réservé pour la saisie des métadonnées.

1.3.3.2 Les chercheurs d'information

Les chercheurs d'information sont le public le plus disparate et le moins connu, ce sont eux qui viennent interroger l'Archive ouverte et espèrent trouver des publications dont le sujet est en relation avec une demande d'information personnelle ou professionnelle.

Les chercheurs d'information utilisent les fonctions de recherche et de navigation du site. Ils demandent parfois à télédécharger un document.

1.3.3.3 Les gestionnaires des données

Les gestionnaires de données sont les administrateurs du logiciel et des métadonnées. Dans le cas de l'Archive ouverte, ce groupe inclut des validateurs qui ont pour fonction de vérifier la conformité des métadonnées afin de garantir la cohérence de l'ensemble.

Les gestionnaires des données utilisent des interfaces à accès réservé afin de vérifier les métadonnées et les corriger au besoin.

---

12 MCKAY, Dana. *Institutional Repositories and Their 'Other' Users : Usability Beyond Authors*. Ariadne [en ligne]. July 2007, Issue 52. http://www.ariadne.ac.uk/issue52/mckay/ (consulté le 8 juillet 2010)

1.3.4 Les fonctionnalités de l'outil

L'objet de cette section n'est pas de lister toutes les fonctionnalités de l'Archive ouverte. En effet une telle liste risque fort d'être rapidement incomplète dans un projet qui est en pleine phase de développement. Il est par contre sans doute utile de mentionner quelques spécificités du service pour en terminer la présentation.

1.3.4.1 Importation de notices

Pour alléger la saisie des métadonnées, il est possible d'importer celles-ci depuis un certain nombre de bases de données (RERO, PubMed[13], arXiv.org[14]) pour peu que le document soit déjà référencé dans l'une de ces bases et que l'auteur en aie connaissance.

1.3.4.2 Liens dynamiques vers l'Archive ouverte

L'interrogation de l'Archive ouverte est rendue possible par la technologie PHP. De ce fait les résultats d'une recherche documentaire peuvent être transmis à l'aide du lien dynamique qui contient les termes de recherche. Il est ainsi possible pour les auteurs d'établir une bibliographie sur la base des notices qu'ils ont recensées dans l'archive ouverte et de mettre un lien vers cette bibliographie.

1.3.4.3 Niveau de diffusion

La philosophie du libre accès n'est pas toujours en accord avec les politiques éditoriales. Les auteurs doivent se tenir informés des droits de diffusion sur des documents qui ont été publiés par un éditeur commercial. L'Archive ouverte prévoit trois niveaux de diffusion pour le texte intégral, lorsqu'il est rattaché à la notice :

- **privé** où seul l'auteur a accès au texte intégral dans une session authentifiée
- **restreint** qui rend possible le télédéchargement du texte intégral aux personnels et étudiants de l'Université dans une session authentifiée
- **public** dont le texte intégral est télédéchargeable depuis tout ordinateur sur Internet

Dans ce travail de bachelor, les textes dont le niveau de diffusion est public ou restreint constituent l'ensemble des documents *consultables*. Ces documents sont en effet consultables pour les personnels et étudiants de l'Université de Genève, qui forment le public cible du service.

---

13 NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. *PubMed* [en ligne]. http://www.ncbi.nlm.nih.gov/pubmed/ (consulté le 14 juillet 2010)

14 CORNELL UNIVERSITY LIBRARY. *arXiv.org* [en ligne]. http://arxiv.org/ (consulté le 14 juillet 2010)

1.4 Le libre accès

Le libre accès est un mouvement international soutenu par une partie de la communauté scientifique et de nombreuses bibliothèques académiques. Le principe du libre accès est de rendre publics les résultats de la recherche académique, avec pour argument principal que celle-ci est financée essentiellement par les contribuables.

1.4.1 Historique

Ce mouvement s'est créé en réaction à l'augmentation exponentielle des tarifs d'abonnement aux revues et journaux scientifiques que les éditeurs privés ont pratiqués depuis la fin des années 1980. De nombreuses bibliothèques ont dû alors renoncer à des abonnements à des revues scientifiques, par manque de budget.

Le libre accès s'inspire largement de la réussite d'arXiv.org qui est un serveur de preprints[15] dans le domaine de la physique des particules. Les auteurs du monde entier y publient les résultats de leurs travaux et les soumettent à leurs confrères avant la publication officielle dans une revue.

En 2002 la Déclaration de Budapest[16] préconise que les auteurs mettent à disposition leurs publications depuis un site personnel ou alors publient dans des revues en ligne et gratuites. En 2003, la Déclaration de Berlin[17] parle de publier sur des serveurs maintenus par des institutions académiques selon les principes du libre accès.

1.4.2 Le libre accès en Suisse en 2010

Le FNS et la Conférence des Recteurs des Universités Suisses sont signataires de la Déclaration de Berlin en 2006. Voici donc un état des lieux des projets de dépôts d'archives électroniques dans le milieu académique suisse.

Les Universités de Bâle, Genève, Lausanne, Saint-Gall et Zürich et les deux écoles polytechniques ont développé leur propre dépôt institutionnel. Les Universités de Fribourg, Neuchâtel et l'Université de la Suisse italienne utilisent RERO DOC comme vitrine pour leurs publications. L'Université de Lucerne présente ses publications sous forme de liste annuelle organisée par faculté et par auteur, sans accès au texte

---

15 épreuves d'auteur non encore finalisées pour la publication dans une revue

16 COLLECTIF. *Budapest Open Access Initiative*. Publiée le 14 février 2002. http://www.soros.org/openaccess/read.shtml (consultée le 12 juillet 2010)

17 COLLECTIF. *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities* [en ligne]. Publiée le 22 octobre 2003. http://oa.mpg.de/openaccess-berlin/berlindeclaration.html (consultée le 12 juillet 2010)

intégral[18]. Le site de l'Université de Berne n'a pas permis de déterminer si cette institution s'était dotée d'un service équivalent ou si elle estimait que les bibliothèques assumaient déjà cette fonction. Ce tableau présente une liste des dépôts institutionnels en Suisse, augmentée de deux projets en Europe qui sont cités plus loin :

Tableau 1

Quelques dépôts d'archives électroniques en Suisse et en Europe

Nom du projet	Institution	Pays	Logiciel
Archive ouverte UNIGE	Université de Genève	Suisse	Fedora / VITAL
Bibliothèque numérique RERO DOC	Réseau des bibliothèques de Suisse occidentale	Suisse	CDSWare
CERN Document Server	Organisation européenne pour la recherche nucléaire (CERN)	Suisse et Europe	CDSWare
Codices electronices sangallensis (CESG)	Stiftsbibliothek St. Gallen	Suisse	Non spécifié
E-collection	Eidgenössische technische Hochschule Zürich (ETHZ)	Suisse	Fedora / Fez
Edoc	Universität Basel	Suisse	EPrints
Infoscience	École polytechnique fédérale de Lausanne (EPFL)	Suisse	CDSWare
MSF Field Research	Médecins sans frontières (MSF)	Suisse	Open Repository
Research platform Alexandria	Universität St. Gallen	Suisse	Alexandria
Serval	Université de Lausanne (UNIL)	Suisse	Fedora / Fez
ZORA	Universität Zürich	Suisse	EPrints
Hyperarticles en ligne (HAL)	Centre pour la communication scientifique directe	France	HAL
Lirias	Katholieke Universiteit Leuven	Belgique	DSpace

Il est intéressant de relever que les Universités suisses ont eu quatre types de réaction face à la directive du FNS :

- ne rien faire
- dresser une liste des publications scientifiques de l'Université
- utiliser une plateforme existante
- développer un nouveau service

18 UNIVERSITÄT LUZERN. *Publikationen* [en ligne]. Modifié le 8 juillet 2010. http://www.unilu.ch/deu/publikationen_187462.html (consulté le 11 juillet 2010)

Cette disparité peut indiquer que les principes du libre accès ne sont pas défendus par l'ensemble de la communauté académique suisse. D'un autre côté, cela peut témoigner d'une prudence de certaines universités à l'égard de services résolument nouveaux et dont on a encore de la peine à chiffrer les coûts sur le long terme.

1.4.3 Statistiques dans quelques archives électroniques institutionnelles

Il est intéressant de savoir quelles sont les statistiques utilisées dans différents projets. Un échange de courriel avec les personnes suivantes a permis de se constituer une première idée :

- M. David Aymonin, directeur de la Bibliothèque de l'EPFL dont dépend Infoscience.
- Mme Jeannette Frey, directrice de la Bibliothèque cantonale et universitaire de Lausanne et cheffe de projet pour Serval.
- Une personne participant au projet HAL[19], qui est un dépôt de publications scientifiques à l'échelon national en France.

D'une façon générale, si ces projets utilisent des indicateurs, ceux-ci s'attachent à décrire la collection et à mesurer le trafic sur le site. Le jeu d'indicateurs s'élargit si l'on considère les statistiques, intégrées dans certains logiciels, à partir desquelles il est possible de faire des classements d'auteurs ou de documents les plus souvent visualisés ou les plus télédéchargés.

1.4.3.1 Statistiques relevées par l'Archive ouverte UNIGE

Avant le démarrage de ce travail de bachelor, l'équipe de projet de l'Archive ouverte avait à sa disposition des statistiques mensuelles sur le nombre total de notices et sur le nombre de celles-ci accompagnées du texte intégral collectées grâce à un outil spécifique développé à l'interne. L'interface VITAL collecte aussi des statistiques sommaires sur le trafic. Et enfin, le trafic du site était analysé par AWStats[20] qui fournit des statistiques détaillées.

Les statistiques relevées par VITAL comptent aussi le nombre de télédéchargements par document individuel, mais les responsables du projet ont décelé des erreurs. Ces statistiques-là sont donc considérées non fiables.

---

19 CENTRE POUR LA COMMUNICATION SCIENTIFIQUE DIRECTE. *Hyperarticles en ligne (HAL)* [en ligne]. http://hal.archives-ouvertes.fr/ (consulté le 14 juillet 2010)

20 AWSTATS. *Site d'AWStats* [en ligne]. http://awstats.sourceforge.net/ (consulté le 14 juillet 2010)

2. Méthodologie

L'exposé de la méthodologie reprend les différents objectifs généraux définis pour réaliser le mandat au chapitre précédent. Ces objectifs généraux sont développés en objectifs spécifiques pour permettre une vision détaillée des différentes tâches prévues pour la réalisation de ce mandat.

2.1 Connaître les enjeux et les orientations stratégiques de l'Archive ouverte UNIGE

La réalisation de cet objectif repose sur une bonne connaissance du milieu académique, de la recherche scientifique et du mouvement du libre accès. Ces notions ont un certain intérêt bibliothéconomique et font partie du cursus de spécialiste en information documentaire.

Pour cerner mieux le projet Archive ouverte, trois objectifs spécifiques se dessinent :

- Recherche documentaire
- Entretiens avec les responsables du projet Archive ouverte
- Entretiens avec les usagers

Cette partie d'analyse se complète par le choix de tableau de bord qui permettra de suivre la gestion du service Archive ouverte.

La démarche pour réaliser ces objectifs spécifiques est expliquée dans la partie qui suit.

2.1.1 Recherche documentaire

Le site de l'Université de Genève a été une mine de renseignements pour comprendre le contexte et les enjeux de l'Archive ouverte.

L'exposé du contexte se complète d'une recherche sur le thème du libre accès, et plus particulièrement sur les dépôts institutionnels en Suisse. La source principale pour identifier des dépôts d'archives électroniques est le site OpenDOAR[21].

2.1.2 Entretiens avec les responsables du projet Archive ouverte

Ces entretiens ont servi, dans un premier temps à clarifier le mandat, puis ils ont permis de prendre connaissance des spécificités de l'Archive ouverte et des

[21] UNIVERSITY OF NOTTINGHAM. *OpenDOAR* [en ligne]. http://www.opendoar.org/index.html (consulté le 9 juillet 2010)

fonctionnalités de l'outil. Les principaux interlocuteurs ont été Madame Véronique Hadengue-Dezaël, cheffe de projet, Monsieur Jean-Blaise Claivaz, coordinateur du projet, et Monsieur Jan Melichar, informaticien de développement.

2.1.3 Entretiens avec les usagers

Les usagers qui ont été interrogés sont les référents académiques dont voici une définition et la description des tâches :

> *« Professeur de la Faculté nommé par le Décanat qui a pour tâches principales :*
>
> - *la mise en œuvre de la présente directive au sein de la Faculté*
> - *la promotion de l'Archive [ouverte] auprès des enseignants des chercheurs de la Faculté*
> - *la transmission de propositions d'améliorations ou de modifications à l'équipe des correspondants Archive »* *(Faculté des lettres, 2009 : 3)*[22]

Dans la typologie des usagers de l'Archive ouverte, les référents académiques sont des auteurs. Cependant, l'un d'eux assume le rôle de validateur pour les publications de sa faculté et il rentre aussi dans la catégorie des gestionnaires de données.

2.1.4 Choix du tableau de bord

Cet objectif spécifique suit la méthodologie proposée par Marc Maisonneuve[23] pour la construction d'un tableau de bord. C'est la première étape, le cadrage, qui est réalisée à ce stade; il s'agit de définir l'objet de suivi, l'objectif du tableau de bord, quelle est le type de décisions qu'il va aider à prendre, et sa typologie selon plusieurs critères.

---

22 UNIVERSITÉ DE GENÈVE. FACULTÉ DES LETTRES. *Archive ouverte UNIGE : Directive de la Faculté des lettres* [en ligne]. Mise à jour le 21 décembre 2009. http://archive-ouverte.unige.ch/outils/Directive_Lettres.pdf (consulté le 2 juillet 2010)

23 MAISONNEUVE, Marc. Construire un tableau de bord. In : CARBONE, Pierre, dir. *Construire des indicateurs et tableaux de bord*. Paris : Tec et Doc ; Villeurbanne : ENSSIB, 2002. P. 205-218 (La boîte à outils)

2.2 Identifier des indicateurs adaptés aux services liés à des ressources numériques

Ce type d'objectif se réalise généralement par deux approches complémentaires :

- recherche documentaire
- partage d'expériences entre collègues

2.2.1 Recherche documentaire

Il a été convenu avec le mandant de rechercher les indicateurs de performance dans les projets informatiques. Cet élargissement du sujet permet en effet une meilleure expertise sur la mesure de performance. La recherche documentaire combine donc les deux axes suivants :

- la mesure de la performance
- les ressources numériques

Les sources consultées sont les catalogues de bibliothèque, les bases de données documentaires (LISA, LISTA, Dialog) et Internet.

2.2.2 Partage d'expérience entre collègues

Les responsables de plusieurs dépôts d'archives électroniques ont été contactés par courriel afin de savoir s'ils avaient mis en place un tableau de bord ou quels indicateurs ils utilisaient dans la gestion de leur service. Les résultats sont exposés au point 1.1.4.3 consacré aux statistiques dans quelques archives électroniques institutionnelles.

2.3 Déterminer quels indicateurs sont les plus pertinents dans le contexte local et comment collecter les données

Cet objectif implique une discussion avec les responsables du projet et une validation de leur part, en particulier M. Jan Melichar qui apporte une expertise technique quant à la faisabilité d'une collecte automatique des données.

2.4 Concevoir et mettre en œuvre le tableau de bord

Le tableau de bord est un nouvel outil qu'il faut intégrer dans l'environnement de l'Archive ouverte. La partie conception fait partie du mandat et se déroule en étroite collaboration avec les responsables du projet Archive ouverte. La mise en oeuvre est par contre du ressort de la Division informatique.

2.5 Recommander des actions correctrices pour des valeurs anormales

Cet objectif général se décompose en deux objectifs spécifiques :

- Fixer des exigences de qualité pour les indicateurs
- Connaître les actions qui peuvent influer sur les indicateurs

2.5.1 Fixer des exigences de qualité pour les indicateurs

Pour réaliser cet objectif, l'expérience acquise dans la mesure des indicateurs est primordiale. Celle-ci peut avoir été recueillie par d'autres services similaires. C'est l'approche de benchmarking, en face de laquelle il faut toujours s'assurer que les chiffres soient comparables. Alternativement, l'expérience peut s'être constituée par la marche de son propre service.

2.5.2 Connaître les actions qui peuvent influer sur les indicateurs

Avec une mesure régulière de l'activité de son service, la mise en œuvre d'actions ciblées et un peu d'intuition, il est possible de mettre en évidence les relations entre les actions et les indicateurs. C'est à nouveau l'expérience qui permet de réaliser cet objectif.

3. Réalisation des objectifs généraux

Ce chapitre est structuré en suivant les objectifs généraux du mandat. Les résultats des démarches entreprises sont exposés.

3.1 Connaître les enjeux et les orientations stratégiques du service Archive ouverte UNIGE

Le service Archive ouverte et le mouvement du libre accès sont déjà présentés dans le premier chapitre de ce mémoire qui est une synthèse de différentes informations disponibles sur Internet, notamment le site de l'Université de Genève, et des entretiens avec les responsables de l'Archive ouverte.

Cette section offre une synthèse des résultats des entretiens avec les usagers et expose les raisons du choix de tableau de bord.

3.1.1 Entretiens avec les usagers

Dans cette section, les usagers sont les référents académiques d'une part et les responsables du service de l'évaluation de la recherche d'autre part.

3.1.1.1 Les référents académiques

L'objectif de ces entretiens était d'évaluer qualitativement les fonctionnalités proposées par le service Archive ouverte. Cette enquête est guidée par un questionnaire semi-directif dont voici les questions :

- Déposez-vous ou faites-vous déposer vos publications dans l'Archive ouverte ?
- Quelles fonctionnalités de l'Archive ouverte vous paraissent-utiles ?
- Quelles autres fonctionnalités ou quelles améliorations sur une fonctionnalité existante devrait proposer l'Archive ouverte, selon vous ?
- Concevez-vous l'Archive ouverte UNIGE comme une source pour vos recherches ?
- Comment verriez-vous la mise en place de pôles disciplinaires à l'échelle de la Suisse ?
- Quelles mesures d'activité pourraient être collectées par l'Archive ouverte, selon vous ?
- Avez-vous des remarques à ajouter ?

Les référents académiques sont au nombre de huit, il y en a un par faculté et un pour l'École de Traduction et d'interprétation. Mme Guillemette Bolens, Mme Pierrette Bouillon, M. Pierre-Alain Carrupt, M. Dimitri Konstantas et M. Jacques de Werra se

sont prêtés à l'exercice. M. Michel Grandjean a répondu par courriel. MM. Stéphane Jouve-Couty et Gabriel Mugny n'auront pas pu faire entendre leur avis dans le temps imparti à la réalisation de ce mandat.

Toutes les personnes interrogées utilisent l'Archive ouverte comme dépôt de leurs publications. Elles utilisent généralement d'autres sources pour leurs recherches, ce qui confirme qu'elles appartiennent essentiellement à la catégorie d'usagers que sont les auteurs.

La question sur le développement de pôles disciplinaires a souvent suscité l'incompréhension dans un premier temps; avant de reconnaître un intérêt potentiel du point de vue du chercheur d'informations; cette question n'est pas discutée plus loin.

La synthèse de ces entretiens comprend des remarques générales, les fonctionnalités discutées qui sont classées selon leur apparition dans les entretiens de la plus fréquente à la plus rare et les considérations sur les mesures bibliométriques. Voici le plan de cette section :

- Remarques générales
- Exportation des résultats
- Création d'une notice bibliographique
- Recherche sur le site
- Archivage à long terme
- Considérations sur les mesures bibliométriques

3.1.1.1.1 Remarques générales

Trois personnes ont relevé l'utilité du service pour faire des liens vers le texte intégral en réponse à une demande spécifique d'autres chercheurs ou pour accéder à des thèses. L'une d'elle a relevé qu'un dépôt institutionnel permet une mise à disposition des publications de façon très rapide et une autre a souligné la possibilité d'un gain sur la fréquence de citation de ses travaux.

Un autre point qui a été soulevé est la visibilité de l'Archive ouverte. Deux des personnes interrogées parlent de l'indexation et du référencement dans des moteurs de recherche en ligne et notamment dans Google Scholar. Selon elles, le service doit améliorer sa visibilité s'il veut démontrer son utilité pour les chercheurs d'information.

Un référent académique s'étonnait de la façon dont le projet Archive ouverte avait été mené. En particulier concernant le manque de collaboration entre les institutions

académiques qui se traduit par plusieurs projets développés en parallèle, cloisonnés et sans réelle possibilité d'échanger leurs données.

Une personne interrogée a soulevé l'intérêt pour le chercheur d'information d'avoir accès à une copie de la table des matières des livres qui ne sont pas disponibles en texte intégral pour des raisons de droit. Le service pourrait même proposer un lien vers l'éditeur commercial.

3.1.1.1.2 Exportation des résultats

Trois personnes ont relevé qu'elles devaient tenir à jour leur liste de publications à la fois sur leur page personnelle et dans l'Archive ouverte. Il serait plus facile pour elles d'avoir une possibilité d'exporter leur bibliographie depuis l'Archive ouverte dans un format compatible avec EndNote ou Reference Manager ou alors de pouvoir intégrer les contenus dans leur page personnelle. Ce problème est déjà connu des responsables du projet qui souhaitent y donner une solution d'ici juin 2011.

3.1.1.1.3 Création d'une notice bibliographique

Deux référents académiques parlent d'une longue et fastidieuse saisie manuelle pour créer une notice dans l'Archive ouverte. Là aussi, le problème est déjà connu, et la valeur de certains champs peut être importée depuis PubMed, arXiv.org, RERO. Des développements ultérieurs devraient permettre d'importer des notices depuis d'autres sources.

3.1.1.1.4 Recherche sur le site

Une personne la trouve claire et opérationnelle. Un autre référent académique la trouve lente dans les temps de réponse en comparant à Infoscience. Ce dernier point s'est amélioré depuis avec l'entrée en service d'un nouveau serveur fin mai 2010.

Enfin une troisième personne souhaiterait l'existence d'un champ de recherche sur l'index des noms d'auteur en page d'accueil.

3.1.1.1.5 Archivage à long terme

Deux personnes ont soulevé la question de la suppression de notice. Pour l'une, cela la freinait dans le dépôt d'articles à paraître. Pour l'autre, le choix a été fait au niveau de la Faculté de ne pas déposer les posters présentés à des conférences, parce que la qualité de l'information scientifique est parfois très brève dans ce type de publications en raison d'hypothèses invalidées lors d'un travail de recherche plus approfondi. Il faut relever que la suppression de documents est en contradiction avec les principes du

libre accès; ainsi même si l'archivage est parfois indirectement un frein à la publication il est nécessaire dans le contexte de l'Archive ouverte.

Par ailleurs, une personne a souligné le côté pratique de pouvoir déposer une publication en texte intégral dans l'Archive ouverte et de savoir qu'elle resterait ainsi toujours disponible sans avoir à s'en soucier.

3.1.1.1.6 Considérations sur les mesures bibliométriques

Un référent académique exprime l'avis que le seul indicateur bibliométrique qui prenne son sens dans le cas d'un dépôt institutionnel serait le nombre de publications à l'actif d'un chercheur. En effet, tout indicateur basé sur la fréquence de citations risque fort d'être incomplet puisque basé sur les publications de son institution uniquement.

Une autre personne interrogée soulève la part d'arbitraire liée aux indicateurs bibliométriques. En effet, toute problématique internationale a vraisemblablement un impact plus fort qu'une problématique locale ou régionale et, par contrecoup, la publication d'une recherche dans la première catégorie sera plus citée que son équivalent dans la deuxième.

3.1.1.2 Les responsables du Service d'évaluation de l'activité scientifique

L'Université de Genève a un Service d'évaluation de l'activité scientifique. Un entretien avec ses responsables, MM. Pablo Achard et Joost Mönks, a révélé aussi un usage particulier de l'Archive ouverte qu'il semble opportun de mentionner ici.

Pour aider ce service dans sa fonction d'évaluation, il est en effet prévu la mise en chantier d'une base de données institutionnelle (BDI) qui recensera le nombre de publications issues de projets de recherche de l'Université de Genève par auteur et par type de document. La BDI aura vraisemblablement pour source unique l'Archive ouverte.

3.1.2 Choix du tableau de bord

Le mandat explicite le besoin d'un tableau de bord pour aider à la gestion du service Archive ouverte. Sur cette base, il est possible de définir le cadrage comme suit :

Tableau 2

Cadrage du tableau de bord

Élément de cadrage	Réponse trouvée pour l'Archive ouverte
Objet de suivi	L'Archive ouverte en tant que structure
Usager principal	Le coordinateur du projet
Type de décision	Identification de dérives pour appliquer des mesures correctrices
Politique ou gestion	Outil d'aide à la gestion
Approche benchmarking	Pas essentielle
Opérationnel ou de pilotage	Tableau de bord de pilotage
Dia- ou syn-chronique	Diachronique

Ce tableau de bord doit offrir un suivi régulier des axes principaux du service Archive ouverte UNIGE tels que définis d'après la mission. La collecte des données est prévue mensuellement.

3.2 Identifier des indicateurs adaptés aux services liés à des ressources numériques

Cet objectif a été réalisé par une recherche documentaire dont les éléments les plus marquants sont rapportés ici. Une liste d'indicateurs est proposée en fin de section.

3.2.1 Recherche documentaire

Les documents consultés sont classés selon leur pertinence :

- L'étude de cas de Yong Ho Kim et Hyun Hee Kim[24], une liste d'indicateurs pour évaluer des dépôts institutionnels essentiellement du point de vue des gestionnaires de données et des auteurs.
- Le site d'E-Metrics[25], un projet américain sur la mesure de performance liée aux ressources numériques. Il comprend les bonnes pratiques et des considérations sur la mesure des données statistiques.
- Le site d'Equinox[26], un projet européen sur la mesure de performance liée aux ressources numériques.
- Le rapport de Sabine Barral[27], lié à la mise en place d'indicateurs d'activité pour les ressources numériques. Ce rapport aborde des aspects techniques et propose des solutions pratiques.
- Le site de COUNTER[28], un projet international sur la mesure de performance liée aux ressources numériques. COUNTER définit des règles de conduite auxquelles de nombreux éditeurs et aggrégateurs de périodiques électroniques acceptent de se plier, ce qui garantit la qualité des données statistiques.
- la norme ISO sur les indicateurs de performance dans les bibliothèques[29], un outil de travail très intéressant pour évaluer un service et qui comporte quelques indicateurs liés aux ressources électroniques.

---

24 KIM, Yong Ho et KIM, Hyun Hee. Development and Validation of Evaluation Indicators for a Consortium of Institutional Repositories : a Case Study of dCollection. *Journal of the American Society for Information Science and Technology*. Juin 2008, vol. 59, no 8, p.1282-1294

25 ASSOCIATION OF RESEARCH LIBRARIES. *E-Metrics* [en ligne]. Modifié le 21 avril 2010. http://www.arl.org/stats/initiatives/emetrics/index.shtml (consulté le 13 juillet 2010)

26 EQUINOX. *Equinox : Library Performance Measurement and Quality Management System* [en ligne]. http://equinox.dcu.ie/index.html (consulté le 5 juillet 2010)

27 BARRAL, Sabine. *Mission : indicateurs d'usages des ressources électroniques. In : Site des bibliothèques universitaires* [en ligne]. http://www.sup.adc.education.fr/bib/Acti/Electro/mission_barral.pdf (consulté le 5 juillet 2010)

28 COUNTER. *COUNTER Codes of Practice* [en ligne]. http://www.projectcounter.org/code_practice.html (consulté le 13 juillet 2010)

29 ORGANISATION INTERNATIONALE DE NORMALISATION. *Information et documentation : indicateurs de performance des bibliothèques*. 2e éd. Genève : ISO, 2008. 91 p. Norme internationale ISO 11620 : 2008 (F)

Les études sur les périodiques électroniques et les bases de données sont très largement majoritaires dans la littérature professionnelle. Peut-être parce qu'il y a encore peu de recul par rapport aux dépôts institutionnels, et aussi parce que contrôler le budget d'acquisition est un enjeu majeur de la gestion des bibliothèques et que c'est précisément les périodiques électroniques qui sont souvent le plus gros poste de ce budget dans le milieu académique. Ainsi le choix a été fait de citer trois projets (COUNTER, E-Metrics et Equinox) qui sont des tentatives de standardisation de l'évaluation de l'usage des ressources électroniques, plutôt qu'une myriade de documents qui s'y rattachent de près ou de loin.

3.2.2 Liste d'indicateurs applicables dans le cas de ressources numériques

Certains indicateurs qui ont été mis en évidence par la revue de la littérature s'appliquent à une collection présentant une mixité de supports papier et électronique, ou encore évaluent la disponibilité de postes de travail informatisés. Ils ne sont pas pertinents dans le contexte d'un service dématérialisé et ne sont pas rapportés dans la liste qui suit pour ne pas surcharger la lecture.

Les indicateurs qui suivent ont été répartis en quatre grandes catégories :

- ressources et infrastructures
- utilisation des services
- efficience
- développement des services

Ces catégories correspondent à celles proposées dans le tableau de bord prospectif adapté en 2002 au monde des bibliothèques par Klaus Ceynowa et André Coners [30]. L'intérêt d'une telle classification est de s'assurer que les axes les plus importants du service soient au moins représentés par un indicateur. Ainsi, avec un choix d'indicateurs appropriés, il est possible d'avoir une information très synthétique sur le service qui permette de mettre en évidence tout dysfonctionnement et d'anticiper les difficultés.

30 CEYNOWA, Klaus et CONERS, André. Balanced Scorecard für wissenschaftliche Bibliotheken. *Zeitschrift für Bibliothekswesen und Bibliographie*. 2002, Sonderhefte 82. 152 p.

Tableau 3

Sélection d'indicateurs identifiés pour les ressources numériques

Référence	Nom de l'indicateur
3.2.3.1.1	Rapidité d'accès depuis la page d'accueil
3.2.3.1.2	Nombre total de notices
3.2.3.1.3	Pourcentage de notices accompagnées du texte intégral
3.2.3.1.4	Pourcentage de publications parues dans les trois ans
3.2.3.1.5	Nombre de types de document
3.2.3.1.6	Nombre de métadonnées par type de document
3.2.3.2.1	Ratio des documents télédéchargés par rapport à la population cible
3.2.3.2.2	Nombre de sessions par catégorie d'usagers
3.2.3.2.3	Pourcentage de ressources inutilisées
3.2.3.2.4	Ratio du nombre de sessions par rapport à la population cible
3.2.3.2.5	Pourcentage de la population cible utilisant les services en ligne
3.2.3.2.6	Nombre de documents télédéchargés par session
3.2.3.2.7	Classement des documents par accès à la notice développée
3.2.3.2.8	Classement des documents par télédéchargement du texte intégral
3.2.3.2.9	Ratio du nombre de participants à des cours de formation pour les ressources électroniques par rapport à la population cible
3.2.3.2.10	Satisfaction des usagers concernant les ressources électroniques
3.2.3.3.1	Coût par télédéchargement
3.2.3.4.1	Pourcentage du temps de travail dévolu aux ressources électroniques

3.2.2.1 Indicateurs portant sur la qualité des ressources et infrastructures

3.2.2.1.1 Rapidité d'accès depuis la page d'accueil

Axe mesuré : simplicité d'utilisation du site Web.

Variables associées : échelle de mesure liée au nombre de clics, échelle de mesure liée à la clarté des termes utilisés comme liens qui parcourent le chemin vers la ressource souhaitée.

Remarque : cet indicateur soulève d'emblée, sur un site aussi complexe que celui de l'Université de Genève, la question de la page d'accueil à considérer. Cet indicateur ne trouvera sa place dans un tableau de bord que si la rapidité d'accès aux ressources les

plus demandées est un objectif stratégique, cependant la démarche est intéressante en terme de suivi du projet au travers de ses différentes phases de développement.

3.2.2.1.2 Nombre total de notices

Axe mesuré : variété de la collection.

Variables associées : nombre total de notices.

Remarque : possibilité d'affiner la mesure par structure académique et/ou par type de document.

3.2.2.1.3 Pourcentage de notices accompagnées du texte intégral

Axe mesuré : réalisation de l'objectif de mise à disposition des textes intégraux en libre accès.

Variables associées : nombre de notices accompagnées du texte intégral, nombre total de notices.

Remarque : possibilité d'affiner la mesure par structure académique et/ou par type de document.

3.2.2.1.4 Pourcentage de publications parues dans les trois ans

Axe mesuré : actualité de l'information scientifique.

Variables associées : nombre de notices pour des documents publiés dans les trois ans, nombre total de notices.

Remarque : possibilité d'affiner la mesure par structure académique et/ou par type de document. Cette durée de trois ans est indicative.

3.2.2.1.5 Nombre de types de documents

Axe mesuré : diversité des publications recueillies.

Variables associées : nombre de types de documents définis.

3.2.2.1.6 Nombre de métadonnées par type de document

Axe mesuré : profondeur de la description

Variables associées : nombre de métadonnées par type de document.

Remarque : possibilité de compter les métadonnées obligatoires.

3.2.2.2 Indicateurs portant sur l'utilisation des services

Remarque : plusieurs indicateurs parlent de population cible. Dans le cas de l'Archive ouverte, la population cible est délimitée par les principaux publics de l'institution faîtière soit l'ensemble formé par les étudiants, les assistants et les professeurs.

3.2.2.2.1 Ratio des documents télédéchargés par rapport à la population cible

Axe mesuré : adéquation de la collection.

Variables associées : nombre de documents télédéchargés dans une période donnée, nombre d'individus formant la population cible.

3.2.2.2.2 Nombre de sessions par catégorie d'usagers

Axe mesuré : adéquation de la collection.

Variables associées : nombre de sessions associées à chaque catégorie d'usager.

Remarque : un tel indicateur nécessite un accès au service par session authentifiée.

3.2.2.2.3 Pourcentage de ressources inutilisées

Axe mesuré : adéquation de la collection.

Variables associées : nombre de documents qui n'ont pas été télédéchargés sur une période donnée, nombre total de documents proposé par le service.

3.2.2.2.4 Ratio du nombre de sessions par rapport à la population cible

Axe mesuré : adéquation de la collection.

Variables associées : nombre de sessions sur une période donnée, nombre d'individus formant la population cible.

3.2.2.2.5 Pourcentage de la population cible utilisant les services en ligne

Axe mesuré : pénétration du service en ligne.

Variables associées : nombre d'utilisateurs faisant partie de la population cible, nombre d'individus formant la population cible.

Remarque : cet indicateur peut se calculer par une enquête sur la population cible ou peut être relevé de façon automatique si le service est accessible à travers une session authentifiée.

3.2.2.2.6 Nombre de documents télédéchargés par session

Axe mesuré : adéquation du service en ligne.

Variables associées : nombre total documents télédéchargés pour une période donnée, nombre total de sessions pour une période donnée.

3.2.2.2.7 Classement des documents par accès à la notice développée

Axe mesuré : popularité de certains documents.

Variables associées : comptage de chaque clic pour lire la notice développée d'un document et comparaison avec les mêmes comptages pour d'autres documents.

Remarque : il est possible d'associer le titre ou de consolider par nom d'auteurs, de collecter les accès sur une période donnée ou depuis l'entrée en fonction du service, etc. pour obtenir une grande variété de classements.

3.2.2.2.8 Classement des documents par télédéchargement du texte intégral

Axe mesuré : popularité de certains documents.

Variables associées : comptage de chaque demande de télédéchargement d'un document et comparaison avec les mêmes comptages pour d'autres documents.

Remarque : il est possible d'associer le titre ou de consolider par nom d'auteurs, de collecter les demandes de télédéchargement sur une période donnée ou depuis l'entrée en fonction du service, etc. pour obtenir une grande variété de classements.

3.2.2.2.9 Ratio du nombre de participants à des cours de formation pour les ressources électroniques par rapport à la population cible

Axe mesuré : adéquation de la formation pour soutenir le développement d'une offre numérique.

Variables associées : nombre de participants à des cours de formation pour les ressources électroniques, nombre d'individus formant la population cible.

3.2.2.2.10 Satisfaction des usagers concernant les ressources électroniques

Cet indicateur se construit par une enquête. L'axe mesurés et les variables peuvent donc être très différents, selon les questions qui composent l'enquête.

3.2.2.3 Indicateurs portant sur l'efficience

3.2.2.3.1 Coût par télédéchargement

Axe mesuré : pertinence d'une offre de collection numérique.

Variables associés : coût de maintenance du service, nombre de télédéchargements sur une période donnée.

3.2.2.4 Indicateurs portant sur le développement des services

3.2.2.4.1 Pourcentage du temps de travail dévolu aux ressources électroniques

Axe mesuré : potentiel de développement des ressources électroniques.

Variables associées : nombre de postes en équivalent plein temps affectés au développement des services en ligne, nombre total de postes en équivalent plein-temps.

3.3 Déterminer quels indicateurs sont les plus pertinents dans le contexte local et comment collecter les données

Cette section présente les principes qui ont guidé la sélection des indicateurs et dresse une liste des indicateurs retenus pour le tableau de bord.

3.3.1 Principes de construction des indicateurs

L'Archive ouverte n'a pas encore connu son rythme de croisière. En particulier, il n'y a pas de plan de développement disponible sur le moyen terme, et la gestion actuelle du service implique la supervision d'une série de développements techniques.

Il en résulte que les responsables du projet n'ont pas encore d'idée très précise de ce qu'ils souhaitent mesurer. Cette incertitude diminue le caractère stable du tableau de bord, qui avait été défini de type diachronique dans l'étape de cadrage.

Une partie des indicateurs sélectionnés existe en tant que tel pour la mesure de ressources électroniques, une autre partie a été construite en adaptant les variables au contexte d'après des indicateurs existants. Tous les indicateurs ont ensuite été validés.

Dans l'optique de faciliter l'évolution du tableau de bord, la sélection d'indicateurs a pour objectif secondaire d'identifier un jeu de variables dont la valeur sera mesurée avec régularité.

Les indicateurs sélectionnés reflètent les axes de la mission en ce qui concerne la collecte (1 indicateur), la conservation (1 indicateur) et l'accès (4 indicateurs).

Pour le projet BDI en lien avec le Service d'évaluation de l'activité scientifique, il est prévu de relever le nombre de publications par auteur et par type de document à partir de l'Archive ouverte. Ces indicateurs ne rentreront pas dans le tableau de bord, en effet ils ne concernent pas la marche générale du service et ils requièrent une grande quantité de mesures qui peuvent être aisément extraites de l'Archive ouverte.

3.3.2 Liste des indicateurs retenus pour le tableau de bord

Ces indicateurs sont groupés selon les axes mis en évidence dans la section 3.2.3 :

- Ressources et infrastructures
- Utilisation des services
- Efficience
- Développement des services

En plus le tableau qui suit leur attribue un type selon qu'ils sont indicateurs d'activité, de moyens, financier ou de performance et, lorsque cela est pertinent, met en évidence la relation à l'un des axes stratégiques de la mission de l'Archive ouverte.

Tableau 4

Liste des indicateurs retenus pour le tableau de bord

Nom de l'indicateur	Type d'indicateur
Pourcentage des documents illisibles	Indicateur de performance lié à la conservation
Pourcentage d'erreurs	Indicateur de performance lié à l'accès
Ratio de texte intégral	Indicateur de performance lié à la collecte
Ratio de textes consultables	Indicateur de performance lié à l'accès
Nombre de documents versés	Indicateur d'activité
Ratio formation	Indicateur d'activité
Ratio de documents télédéchargés	Indicateur de performance lié à l'accès
Taux de télédéchargement par session	Indicateur de performance lié à l'accès
Coût par texte consultable	Indicateur financier
Coût par télédéchargement	Indicateur financier
Taux pour développement	Indicateur de moyens

Un texte sera considéré consultable lorsque son niveau de diffusion est *restreint* ou *public*. Cela garantit en effet l'accès au texte intégral aux professeurs, assistants et étudiants de l'Université de Genève qui sont aussi considéré comme le public cible en leur qualité d'utilisateurs principaux de l'Université et de ses infrastructures.

Certains des indicateurs qui suivent peuvent être déclinés par structure académique ou par type de document. Dans le cas de l'Archive ouverte, l'organisation des structures académiques a été jugée instable et n'est pas détaillé pour des indicateurs portant sur la marche générale du service. Un niveau de détail par type de document a été choisi uniquement pour le nombre de documents versés dans le mois.

3.3.2.1 Indicateurs portant sur la qualité des ressources et infrastructures

3.3.2.1.1 Pourcentage des documents illisibles

Axe mesuré : adéquation des actions de conservation.

Variables associées : nombre de documents illisibles ou corrompus sur un échantillon, nombre total de documents composant l'échantillon.

Collecte des données : comptage et saisie manuelle.

3.3.2.1.2 Pourcentage d'erreurs

Axe mesuré : fiabilité de l'accès.

Variables associées : nombre de clics résultant en erreurs 404, 503 et 504, nombre de clics total.

Collecte des données : analyse automatisée d'après les fichiers journaux (fichiers logs) du serveur Apache.

Remarque : pour avoir une meilleure idée des habitudes des chercheurs d'information il est possible de déduire les accès à des pages d'administration, dont le trafic correspond à des tâches de maintenance du site.

3.3.2.1.3 Ratio de texte intégral

Axe mesuré : sensibilisation des chercheurs à la thématique de l'accès libre.

Variables associées : nombre total de notices, nombre de celles-ci accompagnées du texte intégral.

Collecte des données : recensement automatisé depuis Fedora.

Remarque : possibilité de calculer ces nombres par type de publication et/ou par organe de l'institution.

3.3.2.1.4 Ratio de textes consultables

Axe mesuré : pertinence de l'outil pour la recherche scientifique.

Variables associées : nombre total de notices, nombres de celles-ci accompagnées du texte intégral et consultables par le public cible.

Collecte des données : recensement automatisé depuis Fedora.

Remarque : possibilité de calculer ces nombres par type de publication et/ou par organe de l'institution.

3.3.2.2 Indicateurs portant sur l'utilisation des services

3.3.2.2.1 Nombre de documents versés

Axe mesuré : adéquation de l'outil aux besoins de l'institution.

Variables associées : nombre de documents versés sur une période donnée.

Collecte des données : recensement automatisé depuis Fedora.

Affinage possible : possibilité de calculer ce nombre par type de publication et/ou par organe de l'institution.

3.3.2.2.2 Ratio formation

Axe mesuré : adéquation de l'outil avec les attentes du public cible.

Variables associées : nombre de participants à des cours de formation sur une période donnée, nombre d'individus formant la population cible sur une période donnée.

Collecte des données : saisie manuelle d'après les rapports de formation.

3.3.2.2.3 Ratio de documents télédéchargés

Axe mesuré : adéquation de la collection.

Variables associées : nombre de documents télédéchargés sur une période donnée, nombre d'individus formant la population cible.

Collecte des données : analyse automatisée d'après les fichiers journaux (fichiers logs) du serveur Apache et saisie manuelle d'après les statistiques de l'Université.

3.3.2.2.4 Nombre de documents télédéchargés par session

Axe mesuré : adéquation du service en ligne.

Variables associées : nombre total documents télédéchargés pour une période donnée, nombre total de sessions pour une période donnée.

Collecte des données : analyse automatisée d'après les fichiers journaux (fichiers logs) du serveur Apache.

3.3.2.3 Indicateurs portant sur l'efficience

3.3.2.3.1 Coût par texte consultable

Axe mesuré : pertinence du service.

Variables associées : coût de maintenance sur une période donnée, nombre de documents consultables par le public cible versés sur une période donnée.

Collecte des données : saisie manuelle d'après les rapports de gestion et recensement automatisé depuis Fedora.

Remarque : possibilité de calculer ce nombre par type de publication et/ou par organe de l'institution.

3.3.2.3.2 Coût par télédéchargement

Axe mesuré : pertinence de la collection.

Variables associées : coût de maintenance sur une période donnée, nombre de télédéchargements sur une période donnée.

Collecte des données : saisie manuelle d'après les rapports de gestion et analyse automatisée d'après les fichiers journaux (fichiers logs) du serveur Apache.

Remarque : possibilité de calculer ce nombre par type de publication et/ou par organe de l'institution.

3.3.2.4 Indicateurs portant sur le développement des services

3.3.2.4.1 Taux pour développement

Axe mesuré : potentiel d'adaptation du service.

Variables associées : nombre de postes, en équivalent plein temps, affectés au développement de l'outil (y compris tout travail de configuration sur le système), nombre total de postes affectés à la maintenance et au développement du service, en équivalent plein-temps.

Collecte des données : saisie manuelle d'après les rapports de gestion.

Remarque : seules les activités en mesure d'améliorer le niveau de service ou d'offrir de nouvelles fonctionnalités sont à prendre en compte dans la catégorie développement de l'outil.

3.4 Concevoir et mettre en œuvre le tableau de bord

Le tableau de bord n'a pas été finalisé dans le cadre de ce mandat. En effet, le manque de recul vis-à-vis des indicateurs pertinents pour un dépôt institutionnel se traduisait par une trop grande probabilité d'échec de l'outil. De plus la réalisation technique n'a pas pu être assurée dans le temps imparti à ce travail de bachelor parce que les ressources n'avaient pas été planifiées en conséquence.

Ce travail s'est donc orienté en cours de réalisation vers un de tableau de bord susceptible d'évoluer avec les besoins liés à la gestion de l'Archive ouverte lorsque ceux-ci seraient mieux définis.

En conséquence, ce travail de bachelor a pour principal résultat un système de collecte semi-automatique de données, accompagné d'une interface d'interrogation. Concrètement le système repose sur une base de données MySQL. Le nom du système est l'acronyme BaDonAo.

Pour mettre en œuvre BaDonAo, des spécifications ont été élaborées. Elles se retrouvent en annexe de ce document.

Cette solution présente l'avantage d'avoir un contrôle sur les données. De plus les indicateurs choisis pour démarrer BaDonAo présentent un éventail assez large de variables dans l'optique de pouvoir créer d'autres indicateurs en combinant ces mêmes variables.

La présentation des indicateurs dans BaDonAo ne répond pas à la définition d'un tableau de bord puisque ceux-ci sont présentés sous forme de tableaux de chiffres. Cependant, cette présentation spartiate offre l'avantage de pouvoir exporter aisément les valeurs dans un tableur, depuis lequel il est possible de créer un graphique.

Les principaux types de graphiques et leur fonction sont :

- **Graphiques en courbe** pour montrer une évolution
- **Histogrammes** pour comparer des valeurs entre elles
- **Graphiques en secteurs** pour mettre en évidence des proportions
- **Graphiques en radar** pour représenter un ensemble de variables sans interaction

3.5 Recommander des actions correctrices pour des valeurs anormales

Cet objectif n'a pas non plus été réalisé dans le cadre de ce mandat. Essentiellement parce que la littérature professionnelle est encore assez pauvre sur le sujet.

Les pistes qui peuvent donner des résultats intéressants sont l'étude de Yong Ho Kim et Hyun Hee Kim[31] et le projet DRAMBORA[32] qui propose une auto-évaluation des dépôts d'archives électroniques basée sur l'analyse de risques. En effet, passer à travers un tel processus d'évaluation peut amener à mieux connaître ses forces et faiblesses.

31 KIM, Yong Ho et KIM, Hyun Hee. Development and Validation of Evaluation Indicators for a Consortium of Institutional Repositories (op. cit.)

32 DRAMBORA. Site de DRAMBORA [en ligne] Publié le 1 février 2008. http://www.repositoryaudit.eu/ (consulté le 14 juillet 2010)

Conclusion

Cette conclusion s'articule en trois points : un regard critique sur le travail effectué, une présentation succincte de pistes qui n'ont pas été suivies et enfin des recommandations.

Critique du travail de bachelor

Ce travail se termine et le mandat est réalisé à 75%. Les objectifs de départ étaient assez ambitieux et la réalisation pratique n'a pas été conduite par la Division informatique à cause d'un manque de ressources au moment crucial.

Par chance, le Comité de pilotage du projet Archive ouverte a voté une rallonge de crédit dans sa session du 22 juin 2010. En conséquence, BaDonAo a pu être intégré à un Working package à haute priorité et est en cours de réalisation. Le système devrait être opérationnel à la fin du mois de juillet 2010.

Le mandat sera alors réalisé à 90%. Les 10% restants sont à attribuer au catalogue de mesures correctrices.

Dans le délai imparti et en regard de la charge de travail qu'a représentée l'état de l'art en début de projet, le résultat est assez satisfaisant. Cet état de l'art sur la mesure de performance dans le cadre de ressources numériques a nécessité de parcourir de nombreux articles en lien avec les périodiques électroniques. Or ce type de publications est souvent consacré à deux axes :

- la diffusion
- l'estimation de la valeur d'abonnement

Le premier point tient uniquement compte des chercheurs d'information et ne saurait couvrir toutes les dimensions d'un dépôt institutionnel alors que le deuxième est carrément hors sujet.

Ce sont souvent les liens envoyés par des responsables de services analogues à l'Archive ouverte qui ont permis de trouver des ressources plus pertinentes. Cette approche ne doit donc être en aucun cas négligée dans un tel travail de recherche. Merci à eux !

Le contexte dans lequel se sont développés les dépôts institutionnels en Suisse est assez déroutant. Le FNS émet une directive, certaines universités mordent à

l'hameçon et c'est parti. À aucun moment, il n'y a de réflexion sur la visibilité de la recherche scientifique à un échelon plus large que l'institution.

Quel était l'objectif du FNS, promouvoir le libre accès ou alors se doter d'un outil d'évaluation de l'activité scientifique ? Dans le second cas, la réticence des auteurs à collaborer à un service qu'ils ressentent comme une contrainte, fait rater la cible.

Pour améliorer la participation des auteurs, l'Archive ouverte a une double stratégie. D'une part offrir des services sur mesure pour les auteurs, par exemple la possibilité d'inclure des références et publications sur des recherches antérieures à leur entrée à l'Université de Genève ou la possibilité, encore attendue, d'exporter sa bibliographie. D'autre part la collaboration avec le Service d'évaluation de l'activité scientifique prend une dimension de sanction pour ceux qui n'alimentent pas l'Archive ouverte.

Cette double stratégie est également appliquée par Lirias[33], le dépôt institutionnel de la Katholieke Universiteit Leuven. Dans une conférence à l'Université de Genève le 9 mars 2010, Herlinde Leemans rapporte les possibilités d'exporter sa bibliographie ou encore de l'intégrer directement sur sa page personnelle au sein de l'institution. D'un autre côté, Lirias est l'unique source des bibliographies qui accompagnent les dossiers de promotion des professeurs.

Pistes à suivre

Il a été évoqué d'évaluer le pourcentage de documents référencés par rapport à la production documentaire d'une faculté sur une période donnée. Cette question a mis en évidence la difficulté à recenser toutes les publications scientifiques de l'Université. Pour sa part, le Service d'évaluation de l'activité scientifique se base sur les statistiques du Web of Science de Thomson[34]. Or en parcourant le rapport 2009, seules deux publications, tout type confondu, sont référencées pour la littérature durant toute l'année. Cela démontre le peu de fiabilité de Web of Science quant aux sciences humaines. Il serait sans doute intéressant d'avoir une étude qui comparerait le nombre de publications recensées par différents services internes ou externes à l'Université et de comparer avec l'Archive ouverte pour estimer sa couverture.

---

33 KATHOLIEKE UNIVERSITEIT LEUVEN. *Site de Lirias* [en ligne] https://lirias.kuleuven.be/ (consulté le 14 juillet 2010)

34 THOMSON REUTERS. Site du Web of Science [en ligne] http://wokinfo.com/products_tools/multidisciplinary/webofscience/ (consulté le 14 juillet 2010)

Il peut être intéressant de se poser la question concernant la catégorisation des utilisateurs de l'Archive ouverte. Sont-ils membres de la population cible? De quelle faculté font-ils partie? Un questionnaire en ligne permettrait de renseigner les responsables de l'Archive ouverte sur les chercheurs d'information, et pourquoi pas, les auteurs aussi.

Il est tout-à-fait envisageable, aussi, d'évaluer la façon qu'ont les chercheurs d'information d'utiliser les masques de recherche, par exemple en comptant le nombre de mots-clés utilisés pour chaque requête.

Suivant le contexte du dépôt institutionnel, il peut-être intéressant de relever quelle est la place dans l'institution de la personne qui saisit les métadonnées d'un enregistrement ; pour l'Archive ouverte, on estime qu'environ 95% des notices sont importées par des bibliothécaires. Cela s'explique par le fait que ceux-ci sont actuellement mis à contribution pour augmenter le corpus de l'Archive ouverte en référençant des publications de manière rétrospective.

Un indicateur intéressant serait un indicateur de pertinence, on peut le construire en décomptant les recherches formulées par les chercheurs d'information qui ont eu pour résultat un clic sur un des résultats de recherche, ce qui confirme qu'ils ont trouvé ce qu'ils cherchaient. En déterminant le pourcentage de ces recherches sur le nombre total de recherches, on obtient une valeur qui peut aisément se comparer d'une institution à l'autre.

Un autre indicateur qui a paru particulièrement intéressant est le décompte des demandes de télédéchargements qui n'ont pu aboutir. Il est possible alors d'arriver vers un auteur et de lui présenter, chiffres en main, qu'une mise à disposition plus large d'une publication peut résulter en un gain de citations pour lui. Ce type d'argument est toutefois à double tranchant, il est possible en effet que l'auteur en question croie que son œuvre en serait plus facilement plagiée.

Recommandations

Situation en Suisse

L'état des lieux sur les dépôts institutionnels en Suisse fait ressortir une diversité quant à l'implication des institutions académiques dans le mouvement du libre accès, ainsi que dans les solutions techniques adoptées. Cela s'explique par la nature même de la

Confédération dans laquelle les Universités sont sous la responsabilité directe des cantons et n'ont donc aucune obligation de collaborer.

Une telle collaboration pourrait néanmoins apporter une amélioration certaine du point de vue des usagers: notamment la création de pôles disciplinaires nationaux, c'est-à-dire de dépôts d'archives électroniques en libre accès dont la collection ne soit pas délimitée par une institution mais par un domaine de la connaissance scientifique, pourrait remplir la même fonction que les catalogues collectifs et ainsi faciliter grandement la tâche des chercheurs d'information.

C'est là une vision d'avenir assez lointain, mais qui pourrait aussi avoir d'autres conséquences positives. Par exemple :

- Renforcer le sentiment des chercheurs scientifiques d'appartenir à une communauté
- Accélérer l'évolution des mentalités de la communauté scientifique vis-à-vis du libre accès
- Améliorer la visibilité de la recherche en Suisse
- Offrir une base pour l'évaluation de la recherche scientifique en Suisse.

L'Archive ouverte UNIGE

De son côté, l'Archive ouverte UNIGE est un service soutenu par une volonté politique. Malgré cet important soutien, il doit encore démontrer son utilité à deux catégories de ses usagers.

En premier lieu les auteurs, pour lesquels déposer dans l'Archive ouverte est une contrainte et qui souhaiteraient une saisie manuelle aussi succinte que possible lors du dépôt d'un document ou encore des possibilités d'exportation pour mettre à jour leur bibliographie sur leur page personnelle.

En deuxième lieu les chercheurs d'information, qui ont d'autres habitudes de recherche documentaire et qui se retrouvent face à une répartition des publications scientifiques suisses sur plusieurs dépôts, ce qui leur complique la tâche.

Des réponses ont été esquissées par les responsables du projet, et les solutions techniques devraient être mises en place d'ici juin 2011 :

- augmentation du nombre de bases de données internationales depuis lesquelles une importation de notice sera possible
- exportation des bibliographies
- référencement par les moteurs de recherche
- compatibilité OAI-PMH

La compatibilité OAI-PMH devrait permettre le moissonnage des métadonnées de l'Archive ouverte. Concrètement, RERO ou un autre réservoir de données pourrait importer une notice dans son catalogue et proposer un lien direct vers le document dans l'Archive ouverte. Avec le référencement par les moteurs de recherche c'est l'un des meilleurs moyens d'augmenter la visibilité de l'Archive ouverte.

De plus l'Archive ouverte pourrait mettre un peu plus en avant les fonctionnalités existantes du site, en particulier les flux RSS qui peuvent représenter une source d'information de premier plan pour les chercheurs de l'Université. Cela pourrait se traduire par une démonstration systématique lors des formations ou encore par une page de démonstration sur le site.

Bibliographie

ASSOCIATION OF RESEARCH LIBRARIES. *E-Metrics* [en ligne]. Modifié le 21 avril 2010. http://www.arl.org/stats/initiatives/emetrics/index.shtml (consulté le 13 juillet 2010)

AWSTATS. *Site d'AWStats* [en ligne]. http://awstats.sourceforge.net/ (consulté le 14 juillet 2010)

BARRAL, Sabine. Mission : indicateurs d'usages des ressources électroniques. In : *Site des bibliothèques universitaires* [en ligne]. http://www.sup.adc.education.fr/bib/Acti/Electro/mission_barral.pdf (consulté le 5 juillet 2010)

CEYNOWA, Klaus et CONERS, André. Balanced Scorecard für wissenschaftliche Bibliotheken. *Zeitschrift für Bibliothekswesen und Bibliographie*, 2002, Sonderhefte 82. 152 p.

COLLECTIF. *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities* [en ligne]. Publiée le 22 octobre 2003. http://oa.mpg.de/openaccess-berlin/berlindeclaration.html (consultée le 12 juillet 2010)

COLLECTIF. *Budapest Open Access Initiative*. Publiée le 14 février 2002. http://www.soros.org/openaccess/read.shtml (consultée le 12 juillet 2010)

COUNTER. *COUNTER Codes of Practice* [en ligne]. http://www.projectcounter.org/code_practice.html (consulté le 13 juillet 2010)

DRAMBORA. Site de DRAMBORA [en ligne] Publié le 1 février 2008. http://www.repositoryaudit.eu/ (consulté le 14 juillet 2010)

EQUINOX. *Equinox : Library Performance Measurement and Quality Management System* [en ligne]. http://equinox.dcu.ie/index.html (consulté le 5 juillet 2010)

FEDORA COMMONS. *Site de Fedora* [en ligne]. http://www.fedora-commons.org/ (consulté le 14 juillet 2010)

FONDS NATIONAL SUISSE DE LA RECHERCHE SCIENTIFIQUE. CONSEIL NATIONAL DE LA RECHERCHE. *Directive sur le libre accès (Open Access) aux publications scientifiques issues des projets encouragés par le FNS* [en ligne]. Publié le 4 juillet 2007. http://www.snf.ch/SiteCollectionDocuments/dos_OA_Weisung_f.pdf (consulté le 26 mars 2010)

KIM, Yong Ho et KIM, Hyun Hee. Development and Validation of Evaluation Indicators for a Consortium of Institutional Repositories : a Case Study of dCollection. *Journal of the American Society for Information Science and Technology*, June 2008, vol. 59, Issue 8, p.1282-1294

MAISONNEUVE, Marc. Construire un tableau de bord. In : CARBONE, Pierre, dir. *Construire des indicateurs et tableaux de bord*. Paris : Tec et Doc ; Villeurbanne : ENSSIB, 2002. P. 205-218 (La boîte à outils)

MCADAM, Daisy. Y a-t-il des BUGs à l'Université de Genève ? : d'où venons-nous et où allons-nous ?. *Hors-texte. AGBD,* 2009, no 90, p. 7-10.

MCKAY, Dana. Institutional Repositories and Their 'Other' Users : Usability Beyond Authors. *Ariadne* [en ligne]. July 2007, Issue 52. http://www.ariadne.ac.uk/issue52/mckay/ (consulté le 8 juillet 2010)

ORGANISATION INTERNATIONALE DE NORMALISATION. *Information et documentation : indicateurs de performance des bibliothèques*. 2e éd. Genève : ISO, 2008. 91 p. Norme internationale ISO 11620 : 2008 (F)

RÉSEAU DES BIBLIOTHÈQUES DE SUISSE OCCIDENTALE. *Site de RERO* [en ligne] http://www.rero.ch (consulté le 14 juillet 2010)

THOMSON REUTERS. *Site du Web of Science* [en ligne] http://wokinfo.com/products_tools/multidisciplinary/webofscience/ (consulté le 14 juillet 2010)

UNIVERSITÉ DE GENÈVE. *Plus d'info | Archive ouverte UNIGE* [en ligne]. http://archive-ouverte.unige.ch/vital/access/manager/UGProjet (consulté le 26 mars 2010)

UNIVERSITÉ DE GENÈVE. *Rapport de gestion 2009* [en ligne]. http://www.unige.ch/apropos/rapport-gestion/RapportGestion2009.pdf (consulté le 8 juillet 2010)

UNIVERSITÉ DE GENÈVE. *Site de l'Université de Genève* [en ligne]. http://www.unige.ch/ consulté le 5 juillet 2010)

UNIVERSITÉ DE GENÈVE. DIRECTION DE L'INFORMATION SCIENTIFIQUE. *Site de la Direction de l'information scientifique* [en ligne]. Modifié le 16 décembre 2009. http://www.unige.ch/biblio/sinformer/dis.html (consulté le 5 juillet 2010)

UNIVERSITÉ DE GENÈVE. FACULTÉ DES LETTRES. *Archive ouverte UNIGE : Directive de la Faculté des lettres* [en ligne]. Mise à jour le 21 décembre 2009. http://archive-ouverte.unige.ch/outils/Directive_Lettres.pdf (consulté le 2 juillet 2010)

UNIVERSITY OF NOTTINGHAM. OpenDOAR [en ligne]. http://www.opendoar.org/index.html (consulté le 9 juillet 2010)

UNIVERSITÄT LUZERN. *Publikationen* [en ligne]. Modifié le 8 juillet 2010. http://www.unilu.ch/deu/publikationen_187462.html (consulté le 11 juillet 2010)

VTLS. *Présentation de VITAL* [en ligne]. http://www.vtls.com/products/vital (consultée le 14 juillet 2010)

Liens vers les dépôts d'archives électroniques du tableau 1

BIBLIOTHÈQUE DE L'ABBAYE DE SAINT-GALL. *Site de Codices electronici sangallenses* [en ligne]. http://www.cesg.unifr.ch/fr/ (consulté le 14 juillet 2010)

CENTRE POUR LA COMMUNICATION SCIENTIFIQUE DIRECTE. *Site d'Hyperarticles en ligne* (HAL) [en ligne]. http://hal.archives-ouvertes.fr/ (consulté le 14 juillet 2010)

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE. *Site d'Infoscience* [en ligne]. http://infoscience.epfl.ch/ (consulté le 14 juillet 2010)

EIDGENÖSSISCHE TECHNISCHE HOCHSCHULE ZÜRICH. *Site d'e-collection* [en ligne]. http://e-collection.ethbib.ethz.ch/ (consulté le 14 juillet 2010)

MÉDECINS SANS FRONTIÈRES. *Site de MSF Field Research* [en ligne]. http://fieldresearch.msf.org/msf/ (consulté le 14 juillet 2010)

ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE. *Site de CERN Document Server* [en ligne]. http://cdsweb.cern.ch/ (consulté le 14 juillet 2010)

RÉSEAU DES BIBLIOTHÈQUES DE SUISSE OCCIDENTALE. *Site de RERO DOC* [en ligne] http://doc.rero.ch/ (consulté le 14 juillet 2010)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN. *Site de Lirias* [en ligne]. https://lirias.kuleuven.be/ (consulté le 14 juillet 2010)

UNIVERSITÉ DE BÂLE. *Site d'edoc* [en ligne]. http://edoc.unibas.ch/ (consulté le 14 juillet 2010)

UNIVERSITÉ DE GENÈVE. *Site de l'Archive ouverte UNIGE* [en ligne]. http://archive-ouverte.unige.ch/ (consulté le 14 juillet 2010)

UNIVERSITÉ DE LAUSANNE. Site de Serval [en ligne]. http://www.unil.ch/serval (consulté le 14 juillet 2010)

UNIVERSITÉ DE SAINT-GALL. *Site de la plateforme de recherche Alexandria* [en ligne]. http://www.alexandria.unisg.ch/L-fr (consulté le 14 juillet 2010)

UNIVERSITÉ DE ZÜRICH. *Site de Zora* [en ligne]. http://www.zora.uzh.ch/ (consulté le 14 juillet 2010)

Annexe 1

Spécifications BaDonAo

Base de données de mesures statistiques

pour l'Archive ouverte UNIGE

par :

Cédric PELLA

Table des matières

1. Introduction 47
2. Contraintes sur BaDonAo 48
   2.1 Source des données 48
   2.2 Évolution future 48
3. Architecture fonctionnelle 49
   3.1 Collecte des données 49
   3.2 Base de données 50
   3.3 Interface utilisateur 50
4. Composants de BaDonAo 51
   4.1 BaDonAo_Apache 52
   4.2 BaDonAo_Fedora 52
   4.3 BaDonAo_interro 52
   4.4 BaDonAo_Rapport 53
5. Définition et typage des champs 55
   5.1 Table donnees_statistiques 55
   5.2 Table donnees_statistiques_complementaires 61
6. Conclusion 64

Liste des Tableaux

Tableau 1 Formules pour le calcul des indicateurs 53
Tableau 2 Correspondance entre les terminaisons des champs et les types de documents 55
Tableau 3 Champs de la table donnees_statistiques 60
Tableau 4 Champs de la table donnees_statistiques_complementaires 63

Liste des Figures

Figure 1 Environnement de BaDonAo 49
Figure 2 Composants de BaDonAo à développer 51
Figure 3 Capture d'écran d'un prototype de BaDonAo_interro 53
Figure 4 Capture d'écran d'un prototype de BaDonAo_Rapport 54

1. Introduction

BaDonAo est un acronyme utilisé pour « Base de données pour l'Archive ouverte ». Cette base de données a pour objectif d'aider à la gestion du service Archive ouverte UNIGE en fournissant des mesures quantifiées et comparables dans le temps. Toutefois, BaDonAo est plus qu'une simple base de données; en effet, elle comprend aussi une interface HTML, ainsi que des scripts pour l'extraction automatique de certaines variables.

Ces spécifications ont pour but de décrire l'architecture et les fonctions des différents composants de BaDonAo, afin de faciliter sa réalisation. Chaque composant est décrit individuellement avec une mise en évidence des interactions qu'il a avec d'autres composants des systèmes d'information que sont BaDonAo et l'Archive ouverte UNIGE.

2. Contraintes sur BaDonAo

BaDonAo n'est pas un projet existant pour soi mais un outil pour la gestion de l'Archive ouverte UNIGE. C'est pourquoi les variables mesurées ont été explicitement définies avec les responsables du projet Archive ouverte UNIGE, en fonction des besoins identifiés au printemps 2010.

De plus, il est souhaitable de pouvoir réutiliser et manipuler ces statistiques pour des rapports d'activité. À cette fin, BaDonAo doit permettre d'extraire les variables sous une forme organisée et aisément importable dans des logiciels bureautiques.

2.1 Source des données

Les variables stockées dans BaDonAo proviennent de plusieurs sources :

- les fichiers journaux (aussi appelés fichiers logs) du serveur Apache
- le corpus de documents gérés par Fedora
- d'autres statistiques ou comptages manuels

BaDonAo collecte automatiquement des données sur le trafic sur le serveur Apache à partir des fichiers journaux ainsi que des statistiques sur Fedora. Cependant, certaines variables doivent être importées manuellement.

2.2 Évolution future

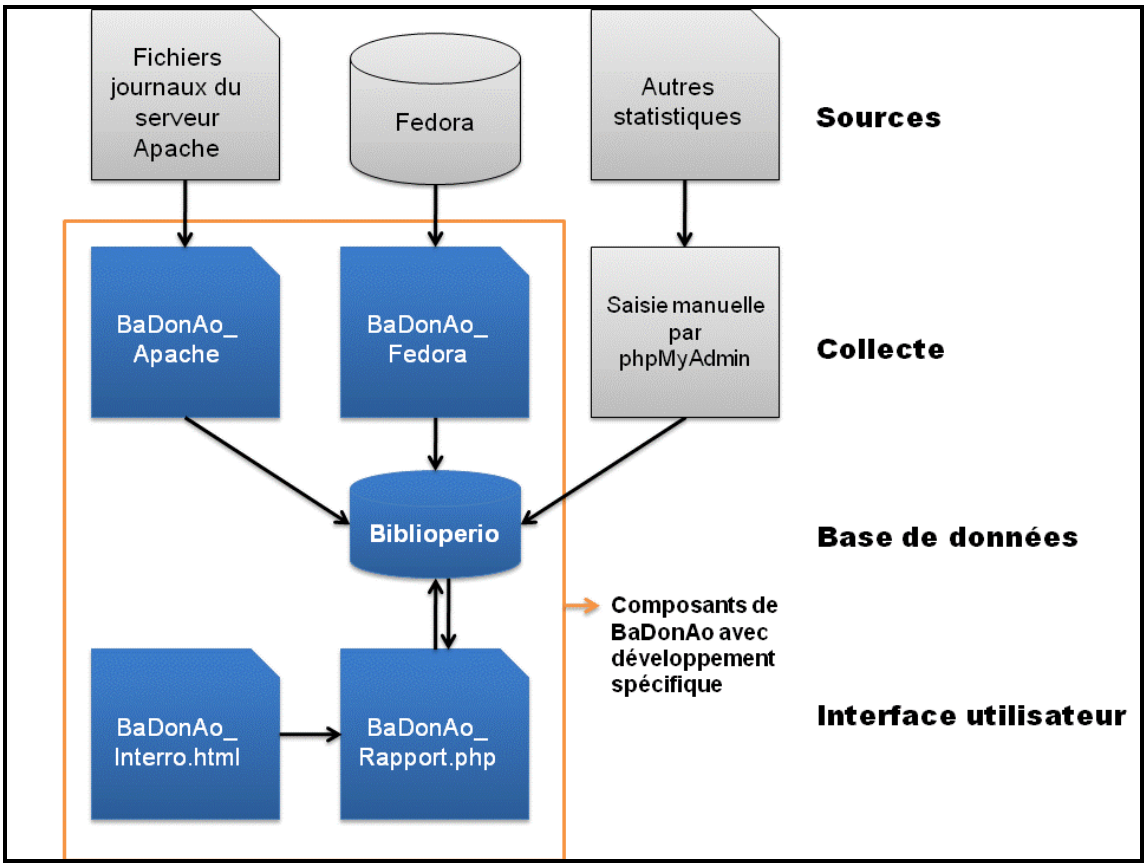
Il est possible que les mesures identifiées aujourd'hui ne suffisent pas à l'avenir. À la demande des responsables du projet Archive ouverte UNIGE, BaDonAo devra donc permettre la collecte de nouvelles données, et la création de nouveaux rapports, moyennant un développement ultérieur sur différents de ses composants.

3. Architecture fonctionnelle

BaDonAo comprend des mesures sauvegardées dans une base de données, une interface HTML pour extraire certaines données et des scripts pour collecter automatiquement certaines mesures.

Figure 1

Environnement de BaDonAo



La figure précédente met bien en évidence les 3 niveaux de BaDonAo :

- Collecte des données
- Base de données
- Interface utilisateur

3.1 Collecte des données

Deux scripts permettent l'importation automatique de certaines données. BaDonAo_Apache parcourt et extrait les informations contenues dans les fichiers journaux du serveur Apache. BaDonAo_Fedora recueille des informations statistiques sur les notices et documents dans Fedora.

La procédure de collecte est agendée pour chaque début de mois. Il est prévu d'exécuter les scripts BaDonAo_Apache, puis BaDonAo_Fedora, soit par un déclenchement manuel soit automatiquement, à l'aide d'un *cron*. Dans un deuxième temps, une saisie manuelle depuis l'interface phpMyAdmin complète la collecte.

3.2 Base de données

L'Archive ouverte UNIGE utilise déjà une base de données MySQL qui contient une liste de périodiques, du nom de *biblioperio*. Nous incorporons dans cette base les deux tables nécessaires au stockage des variables :

- donnees_statistiques
- donnees_statistiques_complementaires

Pour plus d'information, se reporter au chapitre 5 : Définition et typage des champs.

3.3 Interface utilisateur

L'interface est constituée d'une page HTML de paramétrage de rapport et d'une page de rapport, cette dernière est dynamique et repose sur la technologie PHP pour extraire et mettre en forme les données stockées dans la base selon des paramètres prédéfinis.

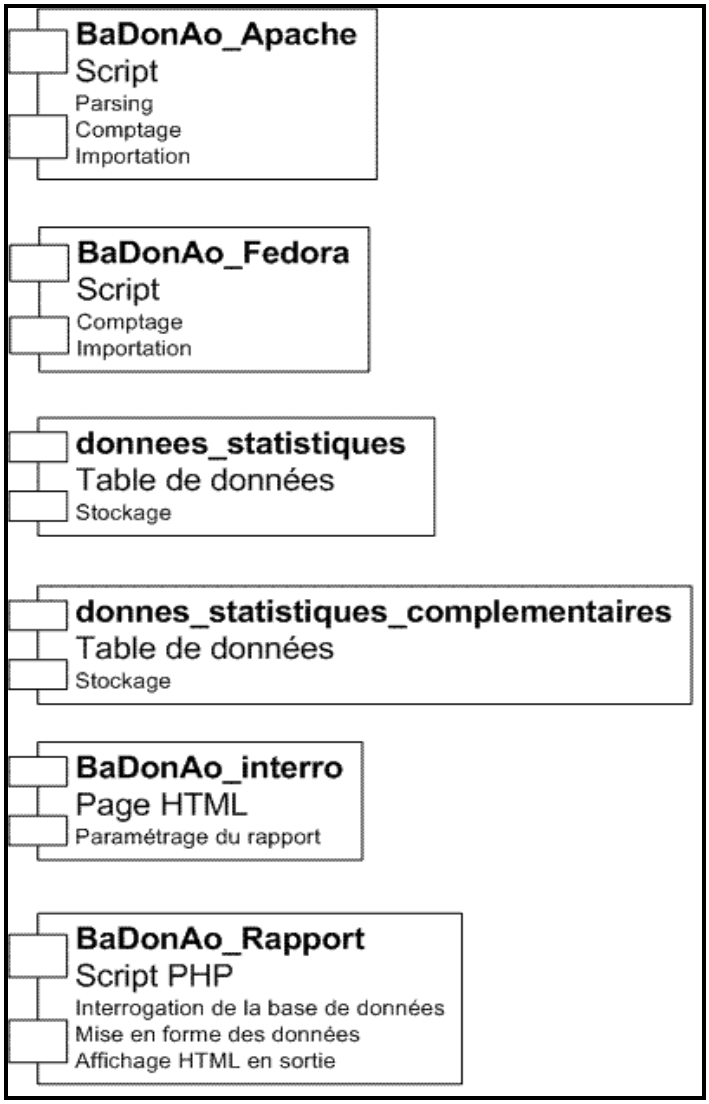
Ces pages sont accessibles dans un nouvel onglet *Reporting* au niveau de l'*Archive ouverte Manager*. Cet outil a en effet un accès réservé selon certaines IP autorisées et correspond tout-à-fait à la mission de gestion de BaDonAo. La présentation des pages HTML est soumise à la CSS de l'*Archive ouverte Manager*.

4. Composants de BaDonAo

Cette partie comporte la description des composants à développer pour BaDonAo et explicite leur fonction. La description des tables et des champs de la base de données constitue l'objet du chapitre suivant. Il était en effet préférable d'y consacrer un chapitre entier en raison de l'importance particulière des données et du calcul de celles-ci dans le projet.

Figure 2

Composants de BaDonAo à développer



4.1 BaDonAo_Apache

Ce script parcourt et analyse les fichiers journaux du serveur Apache. C'est le premier script qui est lancé pour la collecte des données. Il crée donc une nouvelle ligne de données à la date du premier du mois précédent et calcule et introduit dans la base de données la valeur des variables liées à la mesure du trafic sur le serveur :

- Erreurs_navigation
- Clics
- Sessions
- Teledechargement.

4.2 BaDonAo_Fedora

Ce script est en interaction avec Fedora ; son rôle est de recenser les notices selon trois axes :

- Publication (les notices existantes)
- Integral (celles des notices qui sont accompagnées du texte intégral)
- Consultable (celles des notices dont le texte intégral est consultable sur l'Intranet ou l'Internet)

De plus, BaDonAo_Fedora recense les notices ajoutées sur le mois et pour lesquelles les documents ont une date de publication sur l'année en cours ou l'année précédente, toujours selon ces trois axes et selon les 18 sous-types de document existant à ce jour.

BaDonAo_Fedora calcule donc la valeur de 57 variables et stocke celle-ci dans la base de données.

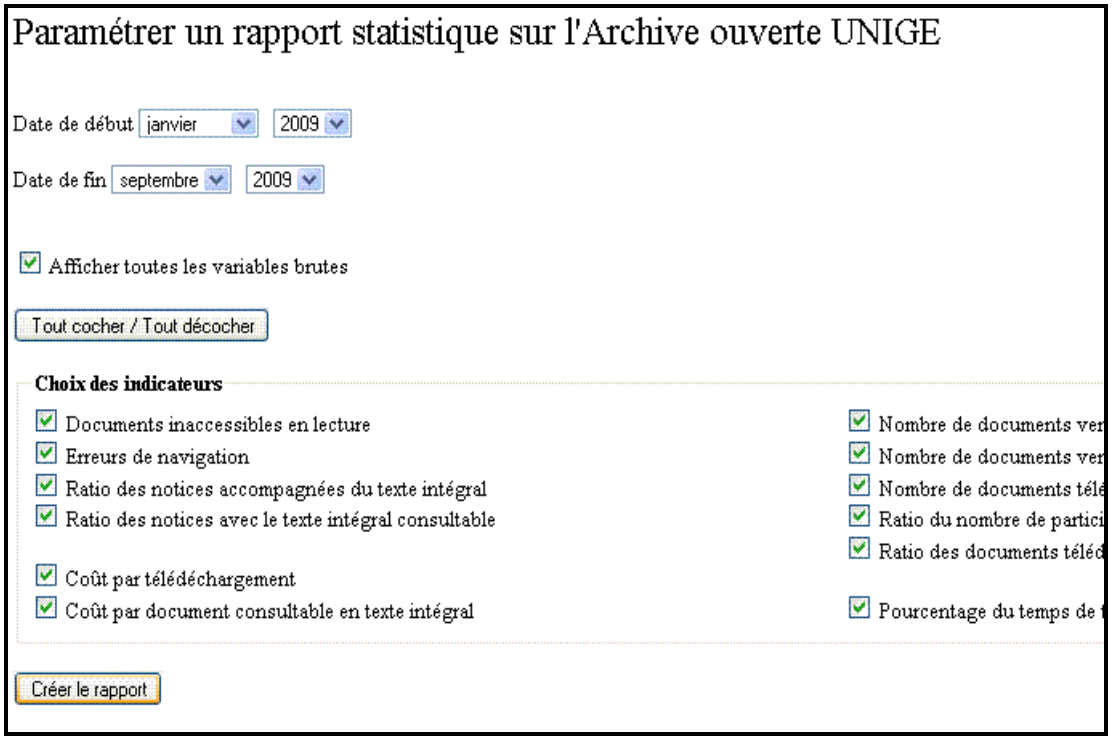
4.3 BaDonAo_interro

C'est une page HTML qui sert à paramétrer le rapport. Les options sont présentées sous forme de listes déroulantes et de cases à cocher.

Cette page permet à l'utilisateur de configurer les dates extrêmes du rapport, et le contenu du rapport, avec un une sélection d'indicateurs et une option de présentation des variables brutes.

Figure 3

Capture d'écran d'un prototype de BaDonAo_interro



La figure précédente présente les différentes options de configuration du rapport.

4.4 BaDonAo_Rapport

C'est un script PHP qui sert à extraire et mettre en forme les variables mesurées et les indicateurs. Les données sont classées par nom et par date et présentées dans des tableaux qui permettent une réutilisation dans un logiciel bureautique.

Tableau 5

Formules pour le calcul des indicateurs

Nom de l'indicateur	Calcul
Pourcentage des documents illisibles	Documents_illisibles / Taille_echantillon
Pourcentage d'erreurs	Erreurs_navigation / Clics
Ratio de texte intégral	Total_integral / Total_publication
Ratio de textes consultable	Total_consultable / Total_publication
Nombre de documents versés	Publication_*

Ratio formation	Formation / Population_cible
Ratio de documents télédéchargés	Teledechargement / Population_cible
Taux de télédéchargement par session	Teledechargement / Sessions
Coût par texte consultable	Cout_maintenance / Somme Consultable_*
Coût par télédéchargement	Cout_maintenance / Teledechargement
Taux pour développement	Heures_developpement / (Heures_maintenance + Heures_developpement)

Un certains nombre d'indicateurs sont prédéfinis. Leur mode de calcul est présenté dans le tableau précédent.

Figure 4

Capture d'écran d'un prototype de BaDonAo_Rapport

Rapport statistique sur l'Archive ouverte UNIGE

Les indicateurs

Mois	Pourcentage de doc. illisibles	Pourcentage d'erreurs	Ratio de texte intégral	Ratio de textes consultables	Ratio formation	Ratio de doc. télédéchargés	Taux de télédéchargemen par session
Jan. 09	0	4.05	75	25	0.13	2.5	7.
Fév. 09	0	5.75	75	50	0	0.06	1.
Mar. 09	1	4.63	70	32.5	0.25	5	10
Avr. 09	0	4.05	75	25	0.13	2.5	7.
Mai 09	0	5.75	75	50	0	0.06	1.
Juin 09	1	4.63	70	32.5	0.25	5	10
Juil. 09	0	4.05	75	25	0.13	2.5	7.
Août 09	0	5.75	75	50	0	0.06	1.
Sep. 09	1	4.63	70	32.5	0.25	5	10

Les variables mensuelles

Mois	Documents illisibles	Taille de l'échantillon	Erreurs navigation	Nombre de hits	Notices publiées	Notices avec texte intégral	Textes consultables	Participants formation	Population cible	Docu téléděc
Jan. 09	0	100	243	6000	40	30	10	15	12000	
Fév. 09	0	100	230	4000	20	15	10	0	12000	
Mar. 09	1	100	250	5400	40	28	13	30	12000	
Avr. 09	0	100	243	6000	40	30	10	15	12000	
Mai 09	0	100	230	4000	20	15	10	0	12000	
Juin 09	1	100	250	5400	40	28	13	30	12000	
Juil. 09	0	100	243	6000	40	30	10	15	12000	
Août 09	0	100	230	4000	20	15	10	0	12000	
Sep. 09	1	100	250	5400	40	28	13	30	12000	

Les données ponctuelles

Date	Nom	Valeur
3 Janvier 2009	Nombre de notices supprimées par dédoublonnage sur décembre 2008	3
6 Juin 2009	Nombre de notices supprimées par dédoublonnage sur mai 2009	7

Retour au paramétrage du rapport

La figure précédente illustre la présentation d'un rapport.

5. Définition et typage des champs

Les variables qui ont été identifiées durant le printemps 2010 sont mesurées mensuellement et stockées dans la table donnees_statistiques. La deuxième table, donnees_statistiques_complementaires, permet de faire évoluer la capture de données selon de nouveaux besoins, qu'ils soient ponctuels ou qu'ils s'inscrivent dans la durée.

Les champs sont présentés par table, avec leur nom dans la base de données, leur type, une définition et la source. Un tableau synthétique reprend les variables de chaque table à la suite des définitions.

5.1 Table donnees_statistiques

Remarque sur les champs Publication_*, Integral_* et Consultable_* : ces champs ont tous une définition unique, déclinée selon la valeur du champ MARC 980 $b. Le tableau qui suit met en évidence la relation entre le nom de variable et le type de document associé.

Tableau 6

Correspondance entre les terminaisons des champs et les types de documents

Terminaison du nom	Type de documents
_acte	Acte de conférence
_article_professionnel	Article professionnel
_article_scientifique	Article scientifique
_autre_article	Autre article
_chapitre_acte	Chapitre d'acte de conférence
_chapitre_livre	Chapitre d'un livre
_livre	Livre
_livre_edite	Livre édité
_maitrise	Maîtrise
_mas	Maîtrise d'études avancées
_numero_revue	Numéro de revue
_poster	Poster

_preprint	Article non publié
_presentation	Présentation à une conférence
_rapport_recherche	Rapport de recherche
_rapport_technique	Rapport technique
_these	Thèse
_these_privat_docent	Thèse de privat docent

5.1.1 Clics

Type : INTEGER.

Le nombre de hits sur le site pendant le mois, en excluant les robots.

Source : importation par BaDonAo_Apache.

5.1.2 Consultable_acte et autres champs déclinés

Type : INTEGER.

Le nombre de notices d'actes de conférence publiées dans l'Archive ouverte pendant le mois, dont la date officielle de publication n'est pas plus ancienne que l'année civile précédente et dont le texte intégral est disponible sur l'Intranet ou sur l'Internet.

Source : importation par BaDonAo_Fedora.

5.1.3 Cout_maintenance

Type : INTEGER.

Arrondi à l'entier du coût de maintenance mensualisé (à l'exclusion des développements logiciels qui la mise en place de nouvelles fonctionnalités et des migrations sur un nouveau serveur de production en parallèle).

Source : saisie manuelle d'après analyse de la gestion du service.

5.1.4 Date

Type : DATETIME.

La date au premier du mois sur lequel ont été effectuées les mesures ; ce champ sert aussi de clé primaire.

Source : créée automatiquement par l'usage du script BaDonAo_Apache.

5.1.5 Documents_illisibles

Type : INTEGER.

Le nombre de documents illisibles d'après sondage mensuel sur un échantillon aléatoire.

Source : saisie manuelle d'après comptage.

5.1.6 Erreurs_navigation

Type : INTEGER.

Le nombre d'erreurs http 404, 503 et 504 sur le trafic du site pendant le mois, en excluant les robots.

Source : importation par BaDonAo_Apache.

5.1.7 Formation

Type : INTEGER.

Le nombre de participants à des formations sur l'Archive ouverte pendant le mois.

Source : Saisie manuelle d'après comptage.

5.1.8 Heures_developpement

Type : INTEGER.

Le nombre d'heures allouées au développement du service Archive ouverte UNIGE (comprend les ajouts de nouvelles fonctionnalités, les paramétrages de serveurs de production en mirroring, etc.).

Source : saisie manuelle d'après analyse de la gestion du service.

5.1.9 Heures_maintenance

Type : INTEGER.

Le nombre d'heures allouées à la maintencance du service Archive ouverte UNIGE (comprend la gestion du service, les migrations vers des nouveaux formats, etc.).

Source : saisie manuelle d'après analyse de la gestion du service.

5.1.10 Integral_acte et autres champs déclinés

Type : INTEGER.

Le nombre de notices d'actes de conférence publiées dans l'Archive ouverte pendant le mois, dont la date officielle de publication n'est pas plus ancienne que l'année civile précédente et qui sont accompagnées du texte intégral.

Source : importation par BaDonAo_Fedora.

5.1.11 Population_cible

Type : INTEGER.

Le nombre d'étudiants, assistants et professeurs inscrits à l'Université de Genève.

Source : saisie manuelle d'après les statistiques de l'Université de Genève.

5.1.12 Publication_acte et autres champs déclinés

Type : INTEGER.

Le nombre de notices d'actes de conférence publiées dans l'Archive ouverte pendant le mois et dont la date officielle de publication n'est pas plus ancienne que l'année civile précédente.

Source : importation par BaDonAo_Fedora.

5.1.13 Sessions

Type : INTEGER.

Le nombre de sessions pendant le mois, en excluant les robots.

Source : importation par BaDonAo_Apache

5.1.14 Taille_echantillon

Type : INTEGER.

La taille de l'échantillon aléatoire sur lequel le sondage de lisibilité a été fait (voir point 5.1.4).

Source : saisie manuelle d'après comptage.

5.1.15 Teledechargement

Type : INTEGER.

Le nombre de télédéchargements demandés pendant le mois.

Source : importation par BaDonAo_Apache.

5.1.16 Total_consultable

Type : INTEGER.

Le nombre total des notices à la fin du mois précédent dont le texte intégral est consultable sur l'Intranet.

Source : importation par BaDonAo_Fedora.

5.1.17 Total_integral

Type : INTEGER.

Le nombre total des notices à la fin du mois précédent qui sont accompagnées du texte intégral.

Source : importation par BaDonAo_Fedora.

5.1.18 Total_publication

Type : INTEGER.

Le nombre total des notices à la fin du mois précédent.

Source : importation par BaDonAo_Fedora.

5.1.19 Tableau de synthèse

Le tableau qui suit reprend tous les champs prévus dans la table donnees_statistiques.

Tableau 7

Champs de la table donnees_statistiques

donnees_statistiques		
Nom	**Type**	**Source**
Clics	INTEGER	Log Apache
Consultable_acte	INTEGER	Fedora
Consultable_article_professionnel	INTEGER	Fedora
Consultable_article_scientifique	INTEGER	Fedora
Consultable_autre_article	INTEGER	Fedora
Consultable_chapitre_acte	INTEGER	Fedora
Consultable_chapitre_livre	INTEGER	Fedora
Consultable_livre	INTEGER	Fedora
Consultable_livre_edite	INTEGER	Fedora
Consultable_maitrise	INTEGER	Fedora
Consultable_mas	INTEGER	Fedora
Consultable_numero_revue	INTEGER	Fedora
Consultable_poster	INTEGER	Fedora
Consultable_preprint	INTEGER	Fedora
Consultable_presentation	INTEGER	Fedora
Consultable_rapport_recherche	INTEGER	Fedora
Consultable_rapport_technique	INTEGER	Fedora
Consultable_these	INTEGER	Fedora
Consultable_these_privat_docent	INTEGER	Fedora
Cout_maintenance	INTEGER	manuel
Date	DATETIME	automatique
Documents_illisibles	SMALLINT	manuel
Erreurs_navigation	INTEGER	log Apache
Formation	INTEGER	manuel
Heures_developpement	INTEGER	manuel
Heures_maintenance	INTEGER	manuel
Integral_acte	INTEGER	Fedora
Integral_article_professionnel	INTEGER	Fedora
Integral_article_scientifique	INTEGER	Fedora
Integral_autre_article	INTEGER	Fedora
Integral_chapitre_acte	INTEGER	Fedora
Integral_chapitre_livre	INTEGER	Fedora
Integral_livre	INTEGER	Fedora
Integral_livre_edite	INTEGER	Fedora
Integral_maitrise	INTEGER	Fedora
Integral_mas	INTEGER	Fedora
Integral_numero_revue	INTEGER	Fedora
Integral_poster	INTEGER	Fedora
Integral_preprint	INTEGER	Fedora
Integral_presentation	INTEGER	Fedora
Integral_rapport_recherche	INTEGER	Fedora
Integral_rapport_technique	INTEGER	Fedora

Nom	Type	Source
Integral_these	INTEGER	Fedora
Integral_these_privat_docent	INTEGER	Fedora
Population_cible	INTEGER	manuel
Publication_acte	INTEGER	Fedora
Publication_article_professionnel	INTEGER	Fedora
Publication_article_scientifique	INTEGER	Fedora
Publication_autre_article	INTEGER	Fedora
Publication_chapitre_acte	INTEGER	Fedora
Publication_chapitre_livre	INTEGER	Fedora
Publication_livre	INTEGER	Fedora
Publication_livre_edite	INTEGER	Fedora
Publication_maitrise	INTEGER	Fedora
Publication_mas	INTEGER	Fedora
Publication_numero_revue	INTEGER	Fedora
Publication_poster	INTEGER	Fedora
Publication_preprint	INTEGER	Fedora
Publication_presentation	INTEGER	Fedora
Publication_rapport_recherche	INTEGER	Fedora
Publication_rapport_technique	INTEGER	Fedora
Publication_these	INTEGER	Fedora
Publication_these_privat_docent	INTEGER	Fedora
Sessions	INTEGER	log Apache
Taille_echantillon	INTEGER	manuel
Teledechargement	INTEGER	log Apache
Total_consultable	INTEGER	Fedora
Total_integral	INTEGER	Fedora
Total_publication	INTEGER	Fedora

5.2 Table donnees_statistiques_complementaires

Cette table est prévue pour la collecte de données ponctuelles ou de données qui n'ont pas été incluses dans la table donnes_statistiques mais qui ont leur importance et nécessitent un suivi mensuel. Chaque mesure de variable prend une ligne complète.

5.2.1 Date_dsc

Type : DATETIME.

Date de fin de la mesure. Pour les mesures mensuelles, dater au premier du mois sur lequel les mesures sont effectuées.

Source : saisie manuelle.

5.2.2 Etendue_temporelle

Type : VARCHAR (30).

Ce champ comporte les dates extrêmes du temps de mesure d'une variable. Coder selon le format suivant : YYYY/MM – YYYY/MM pour des intervalles de dates, laisser le champ vide pour des variables mensuelles (dont la valeur du champ mesure_mensuelle est TRUE).

Source : saisie manuelle.

5.2.3 Increment

Type : INTEGER.

Valeur numérique incrémentale qui sert de clé primaire.

Source : automatique.

5.2.4 Mesure_mensuelle

Type : BOOLEAN.

Affecter TRUE pour des variables relevées mensuellement et FALSE pour les autres.

Source : saisie manuelle.

5.2.5 Nom

Type : VARCHAR (255).

Nom de la variable mesurée, aussi significatif que possible.

Source : saisie manuelle.

5.2.6 Valeur

Type : REAL.

Valeur de la variable mesurée.

Source : saisie manuelle.

5.2.7 Tableau de synthèse

Le tableau qui suit reprend tous les champs prévus dans la table donnees_statistiques_complementaires.

Tableau 8

Champs de la table donnees_statistiques_complementaires

donnees_statistiques_complementaires			
Nom	**Type**	**Source**	**Remarque**
Date_dsc	DATETIME	manuel	Date de la mesure de la variable
Etendue_temporelle	VARCHAR	manuel	Laisser vide pour les mesures mensuelles
Increment	INTEGER	automatique	Clé primaire
Mesure_mensuelle	BOOLEAN	manuel	TRUE pour des variables relevées mensuellement
Nom	VARCHAR	manuel	Nom d'une variable
Valeur	REAL	manuel	Valeur d'une variable

6. Conclusion

Ces spécifications permettent d'appréhender BaDonAo, comme un service rattaché à l'Archive ouverte UNIGE pour aider à la gestion de celle-ci. Plusieurs dimensions sont explicitées afin de faciliter le développement et l'intégration des différents composants de BaDonAo.

Ces spécifications ont aussi pour but de garantir que les composants de BaDonAo auront le comportement que les responsables du projet Archive ouverte UNIGE attendent d'eux.