



L'évaluation de l'utilisation des documents du patrimoine culturel : le projet Allemand de statistiques à libre accès

Matthias Schulze
University of Stuttgart
Stuttgart, Germany

Traduit par :
M'BAHIA N'Guessan
Bibliothécaire, responsable adjoint de la biblio ex Flash,
Université de Cocody

Meeting: 92. Statistics and Evaluation, Information Technology and Preservation and Conservation

WORLD LIBRARY AND INFORMATION CONGRESS: 75TH IFLA GENERAL CONFERENCE AND ASSEMBLY
23-27 August 2009, Milan, Italy

<http://www.ifla.org/annual-conference/ifla75/index.htm>

Résumé

Le projet Allemand des "Statistiques à Libre Accès"(OA-S) est un projet d'infrastructure commun, pour une accumulation normalisée de l'hétérogénéité des données web enregistrées, avec un accent particulier sur les dépôts institutionnels. Les partenaires du projet sont l'université Georg-August de Göttingen (l'Etat de Göttingen et la bibliothèque universitaire), l'université Humbolt de Berlin (service des medias et de l'informatique), l'université de Saarland (l'université de Saarland et la bibliothèque gouvernemental) et l'université de Stuttgart (bibliothèque universitaire). Les actions entreprises sont liées à des coopérations nationales et internationales parmi lesquelles la DINI (Deutsche Initiative für Netwerkinformation/Initiative Allemande pour le Réseau d'Information), la Représentation d'infrastructure de dépôt Numérique (DRIVER) et le Comité des Systèmes d'Information Joint (JISC). Dans les OA-S les dépôts fournisseurs d'informations seront les principaux participant favorisant l'accès aux données (Berlin, Göttingen, Saarbrücken et Stuttgart), et dans les prochaines phases d'expansions tous les dépôt certifiés par la DINI. Avec le temps les connections à d'autres dépôts et aussi à de différents contenus numérisés sont attendues comme avec celle des documents du patrimoine culturel.

Introduction

"Un document d'accès enregistré peut révéler la pertinence du contenu comparé à d'autres documents. Le dépôt d'accès enregistré peut être une base qualitative/quantitative et technologique pour l'évaluation d'un document en provenance d'un dépôt. L'évaluation d'un

document d'accès numérique par rapport à la durée de vie d'un document, peut donner une indication de la constance de sa pertinence. Afin de permettre les comparaisons des différentes données statistiques des services de documentation et de publications, il est essentiel que des normes auxquelles il faudrait adhérer existent. Cependant aujourd'hui, dans les dépôts institutionnels il n'existe pas de normes propres aux statistiques d'accès. En effet différer les protocoles et les technologies d'accès créent des problèmes d'acquisitions et d'évaluation des données. Ils n'existe aucune base pour une évaluation comparative des documents dans les différents dépôts et au sein des services documentation et de publications¹"

Cette citation tirée du chapitre 4.7 de "logs and statistics" du certificat de la DINI dans lequel le groupe de publication électronique de l'Initiative Allemande pour le Réseau d'Information (DINI) impose une explication et des exemples concernant le certificat-DINI².

Cette position peut être vue comme la base sur laquelle le projet décrit ici à été initié, le point de départ du projet des statistiques à libre accès. Les phrases suivantes des textes de la DINI décrivent cela avec plus de précision : "la DINI observe les développements internationaux, ex : le projet du COUNTER (les Recensements des Ressources Electroniques Connectés en Réseau) et le projet des Statistiques de Dépôt Interoperable (IRS) et fera de nouvelles recommandations si nécessaire. D'autres approches comme l'évaluation de la séparation des liens enregistrés seront aussi suivi."³

Comme le projet, cet article parle généralement des dépôts à libre accès et des documents qui y sont stockés. Ainsi dit, les principales mises aux points sont les naissances de matériaux numérique tel que les mémoires, les revues électroniques et les articles etc. aujourd'hui la probabilité d'attribuer l'étiquette de "patrimoine numérique" à ces matériaux en partie est difficile à dire. Les aspects et les dimensions qui se présentent comme appropriés pour ce projet sont de préférences axés sur les matériaux à libre accès et naturellement sur les publications. Mais dans le chapitre de cet article, le centre d'intérêt sera développé et explicitement centré sur le sujet des documents du patrimoine culturel, et d'autre option (le choix) de collecte de données (comme le financement des projets de Numérisation, les coopérations, les questions politiques).

L'évaluation de l'impact des recherches a une longue tradition. Spécialement le Facteur d'Impact des Revues (JIF ou IF) crée par Eugene Garfield est une évaluation bien connue et respecté des citations des revues littératures scientifiques dans certaine disciplines.⁴ Mais le IF a aussi été souvent critiqué.⁵

¹ DINI- Certificate. Document and publication services 2007, Working Group "Electronic Publishing", Version 2.0, September 2006, p.32-<<http://edoc.hu-berlin.de/series/dini-schriften/2006-3-en/PDF/3-en.pdf>>

² Ici en général, sont expliquées les "normes minimums" et les "recommandations" pour les dépôts qui sont définis dans la section des critères (for logs and statistics *ibid.*, p. 16).

³ *Ibid.*, p. 32-33

⁴ Compare Richard Monastersky: the number that's Devouring Science the Chronicle of Higher Education, 2005-<<http://chronicle.com/free/v52/i08/08a01201.htm>>.

⁵ "Malgré ses mérites, le Facteur d'Impact a des défauts qui s'accroisse progressivement au fur et à mesure que les procédés de publication changent. (C'est un problème auquel toutes les méthodes d'évaluations sont confrontées, ex : le teste QI) aussi à ces caractéristiques suivantes : composition arbitraire d'échantillon, période d'observation arbitraire, le niveau de rigidité des revues, la préférence de publication en anglais, distinguer les contenus qui ne proviennent pas des revues, ne gênent pas vraiment les qualité de l'IF dans certains domaines comme celui de l'établissement des sciences naturelles. Mais aussitôt qu'une communauté abandonne le chemin pour n'importe quel autre, elle risque son rang et donc son financement causé par les valeurs déclinantes de l'IF. Et peut importe si cela est provoqué par : des publications ou des archivages via votre page d'accueil personnel (informatique), les longs cycles de réception (mathématiques) ; le mode alternatif de publication ou les barrières de langues(linguistiques). Le Facteur d'impact ne peut compenser les violations de son modèles préalablement acquis." Tirée d'un document non diffusé (pour établir un statut) par Björn Mittelsdorf : Metrics proposal for Institutional Repositories as DINI Recommendation, 2009 (ce document n'est pas à la DINI mais il est adressé à la DINI).

Dans le domaine du libre accès, de nouvelles métriques sont exigées : "le changement rapide de nature concernant les communications des érudits, nous rend conscience du fait que les métriques traditionnelles ne pouvaient pas être exclusivement suffisante pour décrire l'impact des recherches"⁶

Certains développement internationaux , ex : le projet de normalisation COUNTER (Recensements des Ressources Electronique Connecté en Réseau) ou les approches comme l'évaluation des enregistrements des séparateurs de liens développés à la Bibliothèque de Recherche du Laboratoire Los Alamos, devraient être mentionnées ici. L'architecture de base pour entreprendre les enregistrements des dépôts et des séparateurs de liens fut développée par Bollen/Van de Sompel.⁷ Ces infrastructures de développements étaient une partie de la base sur laquelle avaient débuté les statistiques à libre accès. Le projet a été initié par le groupe de publication électronique de la DINI - comme on l'a vue dans l'introduction, à propos des centres d'intérêt des travaux de la DINI- qui à incité les partenaires du projet à y prendre part.

L'effort de production et de la publication d'un texte est en vain, si ce texte n'est pas lu avec une attention particulière. Ici, il n'y a pas de différence entre publications scientifique et celle effectuée dans un monde de fiction. La facilité d'accès expérimentée sur les publications à libre accès montre que, le fait de ne manquer d'aucun besoin d'authentification, d'identification personnelle et d'être suivi financièrement simplifie beaucoup les choses et permet de satisfaire hautement la communauté scientifique. De semblables hypothèses et celle mentionné ci-dessus peuvent être vérifié par des analyses empiriques.

Les demandes peuvent être évaluée facilement étant donné que les serveurs emmagasinent la plus part des pièces d'informations utiles pour la gestion interne à toute utilisation. (Toute tentative d'approche qui essaye de surpasser ce niveau, doit faire face au protocole http. A ce niveau cela rend impossible d'apprendre quelques chose définitivement sur les variables dépendantes, pareil pour la réussite de transfert de fichier ou le temps qu'un utilisateur passe sur un document avant de progressé).

Les publications scientifiques couvrent une large variété de publicités, des hébergements, des modèles de logiciel bureautique, d'utilisation de ces modèles, des étapes de la publication et des techniques et logiques de présentations. Aussi il est important de connaître les possibles espaces de publication, et quels sont les représentants qui veulent être inclure dans l'échantillonnage. Pour toutes ces bonnes volontés de participer, seulement deux aspects sont appropriés :

1. quelles données doivent être recueillies?
2. comment les données peuvent être transférées aux fournisseurs de statistiques ?

Dans ce contexte la technique de base des statistiques à libre accès et l'évaluation de l'impact des recherches, n'est pas le principal centre d'intérêt de cet article. A cet effet le texte suivant ne traitera pas intensivement de ces sujets. Le meilleur compte rendu concernant ce sujet est celui de Frank Schloze : Measuring Research Impact in an Open Access Environnement, 2007, p. 6-8-<http://elib.in-stuttgart.de/opus/volltexte/2007/3234/pdf/liber_2007.pdf>.

⁶ Ibid. p.2. Compare also p.2-3

⁷ Johan Bollen/Herbert Van de Sompel: architecture for aggregation and analysis usage data. In: Proceedings of the 6th ACM/IEEE-CS joint Conference on Digital libraries, 2006, pp 298-307-<<http://doi.acm.org/10.1145/1141753.1141821>>.

Description du projet : statistiques à libre accès/Open Access-Statistics

Le projet de statistiques à libre accès (OA-S/Open Access-Statistics) est un projet commun de recherche s'adressant aux questions posées ci-dessus. Débutant en juillet 2008, une infrastructure sera construite pour normaliser l'accumulation de l'hétérogénéité des données des sites web enregistrées avec un accent particulier sur les dépôts institutionnels. Ainsi les coopérations étroites avec d'autres projets de libre accès variés ajoutés aux services de valeur, permettront leur mise à disponibilité au service des usagers.⁸



FIGURE 1 : le logo des statistiques à libre accès⁹

Les débuts de l'existence de l'OA-S a été financés par la "Fondation Allemande de Recherche"(Deutsche Forschungsgemeinschaft-DFG). Plus tard, les partenaires du projet l'alimenteront comme un service permanent. L'idée initial était-comme mentionnée ci-dessus-choisi par la DINI (Deutsche Initiative für Netzwerkinformation/Initiative Allemande pour le Réseau d'Information) plus précisément par son groupe de Publication Electronique.¹⁰ Les partenaires du projet OA-S sont : l'université Georg-August de Göttingen (l'Etat de Göttingen et la bibliothèque universitaire), l'université Humbolt de Berlin (service des medias et de l'informatique), l'université de Saarland (l'université de Saarland et la bibliothèque gouvernemental) et l'université de Stuttgart (bibliothèque universitaire). Les actions entreprises sont liées à des coopérations nationales et internationales parmi lesquelles la Représentation d'infrastructure de dépôt Numérique (DRIVER)¹¹, la Ligue des Bibliothèques Européennes de Recherche (LIBER) et le Comité des Systèmes d'Information Joints (JISC)¹².

⁸ Pour plus d'information à propos des statistiques à Libre Accès voir :<<http://www.dini.de/projekte/oa-statistik/#c1203>>. Voir aussi Ulrich Herb: OA-Statistik: Standardisierte Nutzungsanalysen als alternative Impact-Messungen wissen: chaftlicher Publikationen. Workshop zu Standardisierung and Vernetzung von Open-Access-Repositoryen für die Wissenschaft, Göttingen, 19.02.2009<<http://www.dini.de/fileadmin/workshops/oa-netzwerk-februar2009/herb.pdf>>

⁹ Source: <http://www.dini.de/projekte/oa-statistik/>.

¹⁰ Pour plus d'information sur le DINI<<http://www.dini.de/english/>>. "la DINI s'est engagée à développer les infrastructures d'information et aussi l'amélioration des services de communication et d'information et aussi l'amélioration des services de communications et d'information des institutions de l'enseignement supérieur et dans le monde des apprenants, régional et suprarégional" consultez le lien :<<http://www.dini.de/english/ago/e-pubo/>>. Pour plus d'information sur le groupe de publication électronique de DINI.

¹¹ "Le projet DRIVER 'Représentation d'Infrastructure d'Entrepôt Numérique pour la recherché Européenne', est conduit par le consortium Ec (de financement) qui construit un cadre organisationnel et technologique des couches de données pour un panorama Européen permettant l'usage avancé des contenus des ressources dans la recherche et au sein de l'enseignement supérieur. DRIVER développe un service d'infrastructure et la fondation des données. Tous les deux sont distinguées pour orchestrer les ressources et les services existant du décor (site) des dépôts." Tirée de "DRIVER guidelines 2.0. guidelines for content providers-Exposing textual resources with OAI-PMH" November 2008, p.9<http://www.driver-support.eu/document/DRIVER_Guidelines_v2_Final_2008-11-13.pdf>

¹² JISC Usage Statistics Workshop Berlin, July 7-8, 2008, <<http://www.dini.de/veranstaltungen/workshops/jisc-workshop-2008/>>. Voir aussi JISC Usage Statistics Review. Final Report. September 24, 2008<<http://ie-repository.jisc.ac.uk/250/>>.

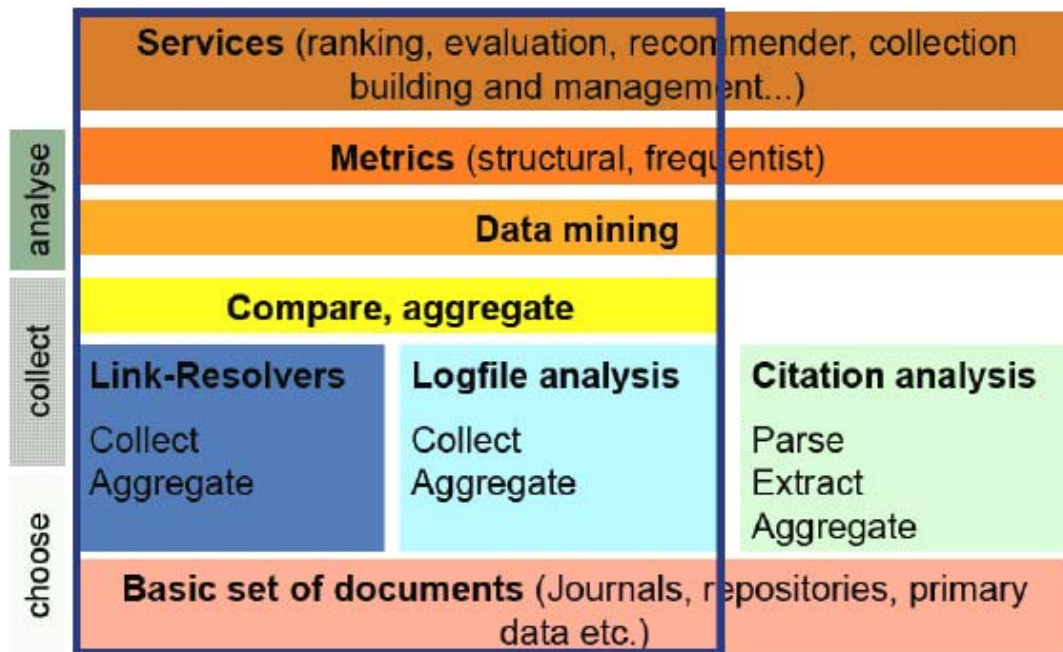


FIGURE 2 : l'Evaluation de l'Impact des Publication : les Eléments (shema)¹³

Selon le point de vue des fournisseurs des services/statistiques centraux, la diversité des fournisseurs constitue une source pour les données d'accès. Ces fournisseurs de données du projet statistiques à libres accès OA-S, seront les dépôts participatifs (Berlin, Göttingen, Saarbrücken et Stuttgart) et dans les prochaines étapes d'expansion, les dépôts certifiés-DINI¹⁴. A la longues les connections à d' autre dépôt sont attendues. L'infrastructure est préparé à recevoir les dépôt fournisseurs nationaux et internationaux, qui y prenent part et à bénéficier de la transformation et à l'assemblage des services fournis par le centre des services fournisseurs. Les dépôts ne font qu'utiliser seulement ces différents logiciels (Dspace, OPUS, e-doc) comme solutions, mais il y a aussi des différences qualitatives entre l'information des documents rassemblés sur un serveur hôte et l'étendue de l'information sur un serveur séparateur de lien (SFX, Ovid) ou contrôleur de licence (HAN).

¹³ Tirée de Frank Scholze: Measuring Research Impact in an Open Access Environment, 2007, p.6<http://elib.uni-stuttgart.de/opus/volltexte/2007/3234/pdf/liber_2007.pdf>.

¹⁴ Serveur certifié-DINI :<http://www.dini.de/no_cache/service/dini-zertifikat/zertifizierte-server/>.

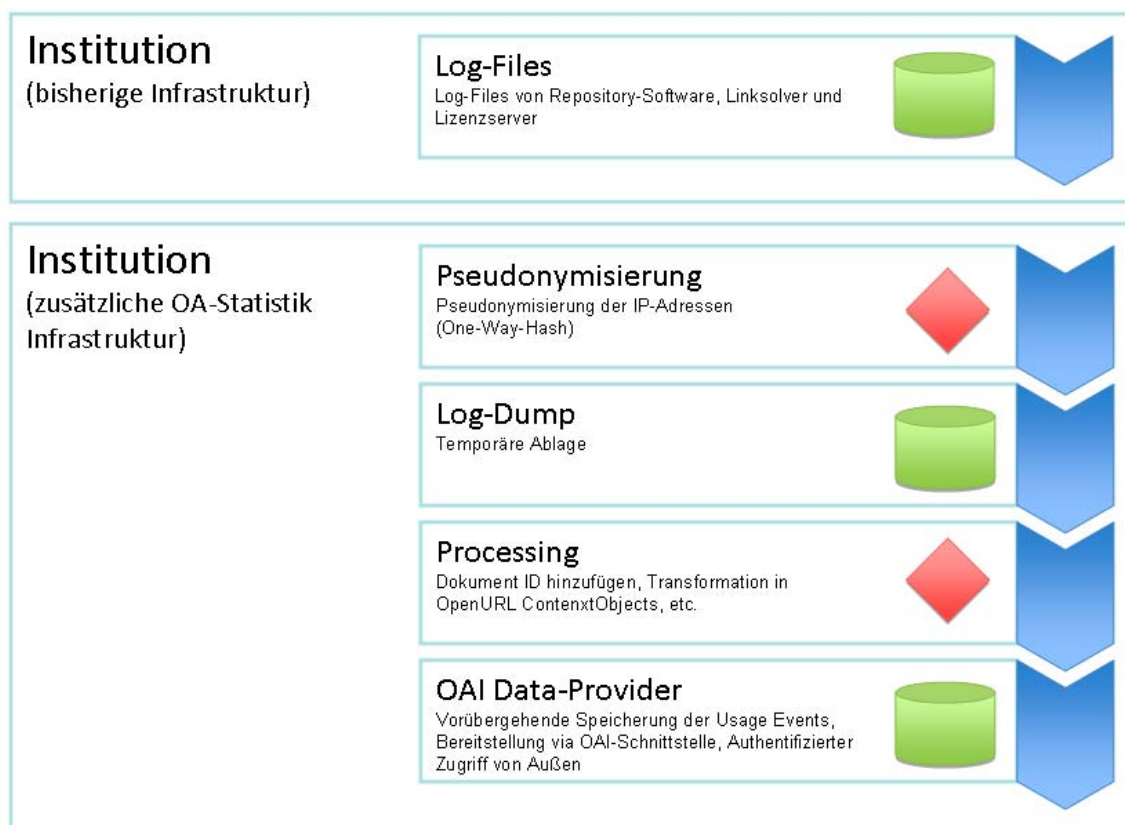


FIGURE 3 : Préparation de l'échange de donnée

Les collections dérivées des données générées localement par les fournisseurs de statistiques seront stockées localement dans un serveur central, se référant ainsi à un service de fournisseur. Les dépôts locaux réussiront à créer des services de valeurs ajoutées de document en intégrant les statistiques dans les portes d'entrées des documents. Un autre exemple actuellement en vogue, serait un système de recommandation basé sur l'analyse du flux des cliques. Initialement les dépôts seront probablement centrés sur la portion des données. Surtout sur celles qui décrivent leurs objets numériques, comme la distinction d'utilisation fiables de fréquence entre les visiteurs locaux et internationaux.

Le certificat-DINI¹⁵ sera développé par les sujets proposés et par l'appui des collections de données. La présentation des statistiques d'information et l'intégration des dépôts appropriés dans la mise en réseau des infrastructures OA - N, permettront de propager les résultats de ce projet. Les directives, la documentation technique et l'introduction de logiciel seront profitables pour les dépôts intéressés. Ainsi ils peuvent avec les efforts et en se concentrant réellement sur leurs prestations¹⁶-satisfaire l'utilisateur dans sa quête du savoir.

Le projet OA-Statistics/statistiques à Libre accès vise deux objectifs. Le premier objectif est la recherche de l'établissement d'infrastructure de sécurité-connecté au réseau des Dépôts à Libre Accès (OA-N)-pour la collection et le traitement de l'utilisation de donnée. Les usages des ressources d'informations électroniques sont collectés à plusieurs niveaux. En plus, sont

¹⁵ Pour plus d'information sur le certificat-DINI voir : <<http://www.dini.de/english/dini-certificate/>>.

¹⁶ Le projet des Statistiques à Libre Accès a reçu plusieurs lettres de soutien, notamment celle de la Division d'Accès de la LIBER

aussi considérés, les différents partenaires projet (Dspace, OPUS et e - doc) et les prétendus liens (SFX, Linkresolver si nécessaire) et la licence du serveur (HAN).

Le second objectif du projet OA-S est de développer et d'établir une norme incontournable afin de mesurer le trafic et les statistiques pour la recherche des publications et des dépôts institutionnels. Aussi les normes et les services déjà établis comme le COUNTER, LogEc ou IFABC sont évaluées et pris en considération.¹⁷

Les données utilisées sont collectées par les partenaires des projets adhérents et par la suite traitée. Quelques temps après la préparation de l'utilisation des données, ils seront transférés sur un format unique et normalisé (Open url Context Objects). Ces données normalisées sont fournies sur une interface OAI (figure 4)

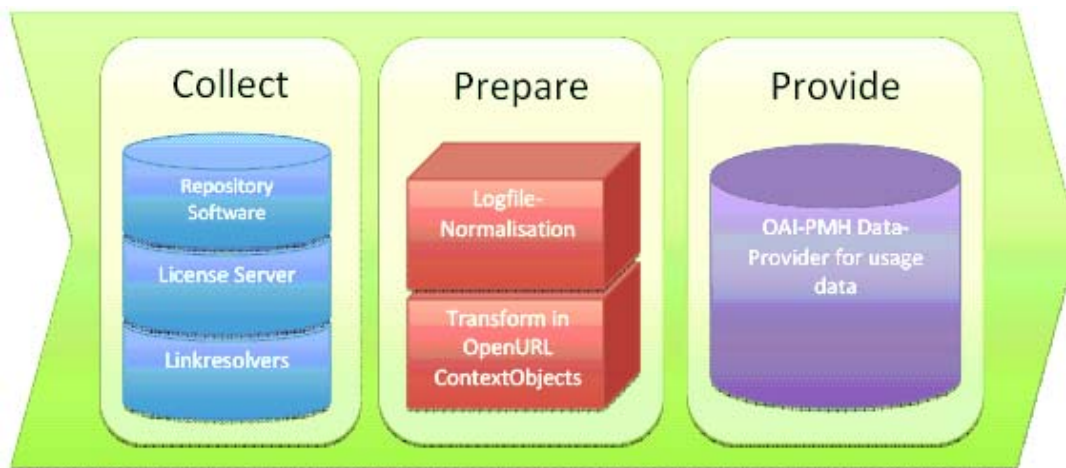


FIGURE 4 : plan de travail des fournisseurs de donnée

Un service de fournisseur a accès aux données et est disponible via l'interface OAI et les collectes à partir des partenaires du projet. Le service fournisseur peut générer une variété de statistiques et de métrique à parti des données, tant que leur utilisation est disponible en un format unique. En plus, une production de l'utilisation de données sera exécutée. Les données regroupées seront mises à la disponibilité du projet "OA-Netzwerk". En complément, il est impossible que les dépôts mettent à profit leur ressource de statistique en les intégrant dans leurs installations locales. (Figure 5)

¹⁷ Concernant "l'utilisation des échanges de statistiques" voir l'annexe de "Driver Guidelines 2.0. Guidelines for Content providers – Exposing textual resources with OAI-PMH", November 2008, p. 132-135<http://www.driver-support.eu/documents/DRIVER_Guidelines_v2_Finla_2008-11-13.pdf>.

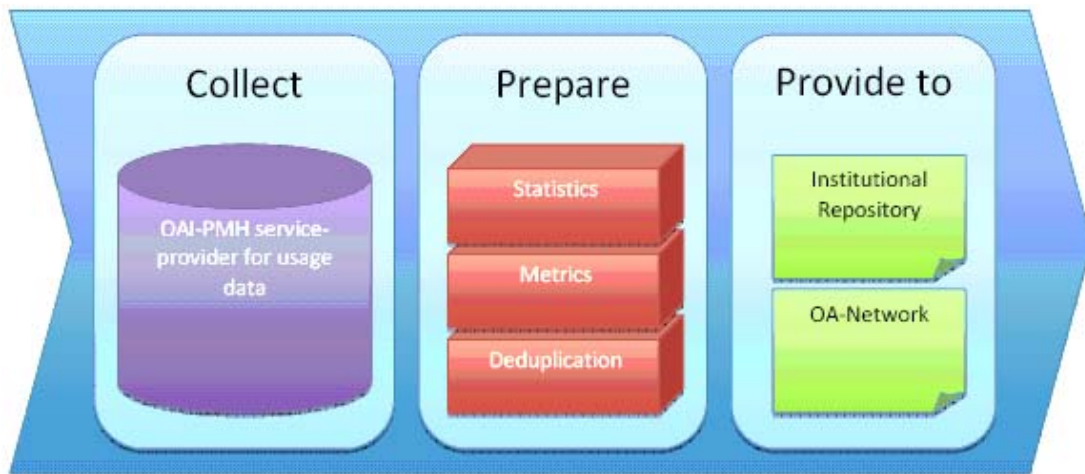


FIGURE 5 : plan de travail au service de fournisseur

Les objectifs du projet comprennent le design (le plan de construction) des divers services de valeurs ajoutées, lesquels peuvent être présentés aux utilisateurs par exemple, et au niveau des dépôts institutionnel. Dans les deux cas l'utilisateur des services obtient une valeur additionnelle. Tels services de valeurs ajoutées enrichissent les actuels offres en fournissant de nouvelles possibilités pour la sélection et l'affichage des documents électroniques. En outre la donnée additionnelle à chaque document est disponible. Par exemple : comment il a été fréquemment consulté au niveau local mais aussi au niveau national. Cela s'avérera être un service attractif pour les scientifiques, les universités, la recherche et aussi pour les institutions de financement. Il indique la fiabilité et l'uniformité de la façon dont les publications scientifiques sont utilisées et l'impact que le libre accès a sur l'utilisation et la clarté des publications.

Afin de percevoir la plus large diffusion possible des résultats du projet, des recommandations seront ajoutées au certificat-DINI qui auront à promouvoir la collection normalisée et les dispositions de l'utilisation de données des statistiques comme les connections au Réseaux. D'un autre côté les directives, les documents d'assistance et les logiciels seront mis à la disponibilité des gérants de dépôt, ainsi ils utiliseront simplement les données normalisées et fourniront les statistiques descriptives. Cela leur permettra d'offrir à un prix raisonnable-les services de valeurs ajoutées (Figure 6-8).¹⁸

¹⁸

Les figures 6-8 contenant des exemples de réalisation (sélectionnées des services des valeurs ajoutées) sont tirées de <<http://www.dini.de/projekte/oa-statistik/#c1190>>.



Figure 6 : service de valeur ajoutée "Statistiques/Métriques" pour les dépôts (esquisse).

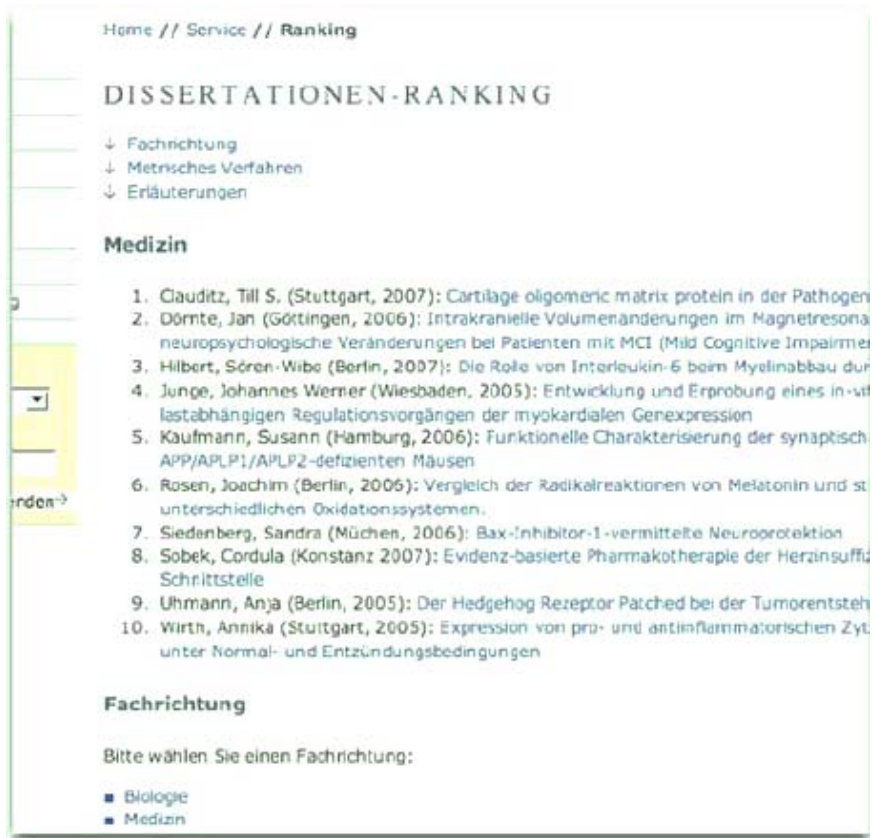


Figure 7 : service de valeur ajoutée "classement de mémoire" pour dépôt (esquisse)

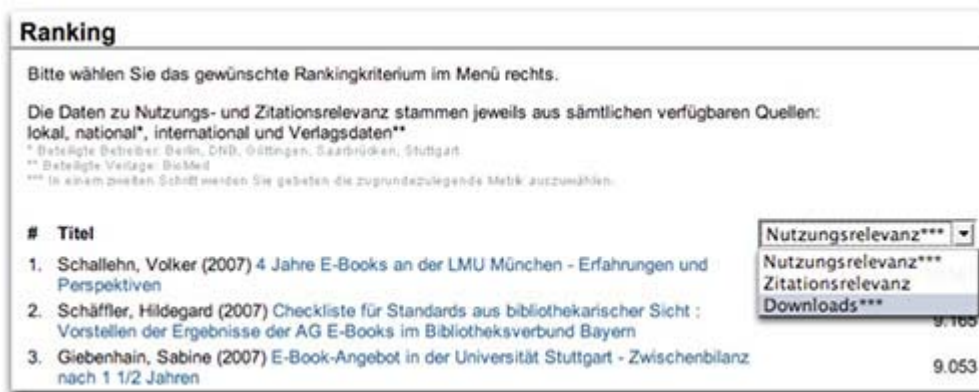


Figure 8 : service de valeur ajoutée "pertinence de l'utilisation" dans les dépôts (esquisse)

Mise en réseau : les projets Allemand de libre accès et le contexte international

La clarté et l'utilisation absolue des réseaux de publication numérique sont les piliers de sa croissance. Toutefois les dépôts en réseau sont ainsi plus attractifs aux chercheurs en ce qui concerne la résolution de leur fournir des œuvres. L'activité en réseau existe au niveau disciplinaire et interdisciplinaire (le projet présenté ici fait référence aux activités interdisciplinaires) aussi bien sur le plan national et international. En particulier, deux initiatives de valeurs sont mentionnées : au niveau Allemand, le projet mentionné ci-dessus "Réseau de Dépôts à Libre Accès (OA-Network)", au niveau Européen le projet de la "Représentation d'infrastructure de dépôt Numérique (DRIVER) ".

L'initiation du projet national "OA-Network"¹⁹ par la DINI et financé par le DFG, porte sur la mise en réseau virtuel de tous les documents certifiés-DINI et les services de publications. A cet effet le OA-Network à un portail contenant divers services d'utilisateurs finaux (spécialiste en navigation, fonction exportatrice) et sa salle de donnée est constamment disponible pour utilisation à nouveau. Aussi, elle collabore avec d'autre projet pour répandre l'influence des services, par exemple concernant l'utilisations des statistiques et des citations. De plus, l'OA-Network soutient les opérateurs de dépôt durant le processus de certification DINI et contribut à l'accroissement du nombre des dépôts certifiés-DINI. Actuellement vingt neuf dépôts Allemand possèdent le certificat DINI, cinq autres sont dans le processus d'application.²⁰ Le certificat est reconnu au plan international et est disponible en version anglaise, allemande et espagnole. Les "lignes directives DARE" ainsi que le certificat DINI servaient comme base pour le développement des grandes lignes du DRIVER, afin de satisfaire les fournisseurs²¹. Ainsi, tous les dépôts certifié-DINI sont simplement appropriés à DRIVER et après adoption sont intégrés dans les deux réseaux.

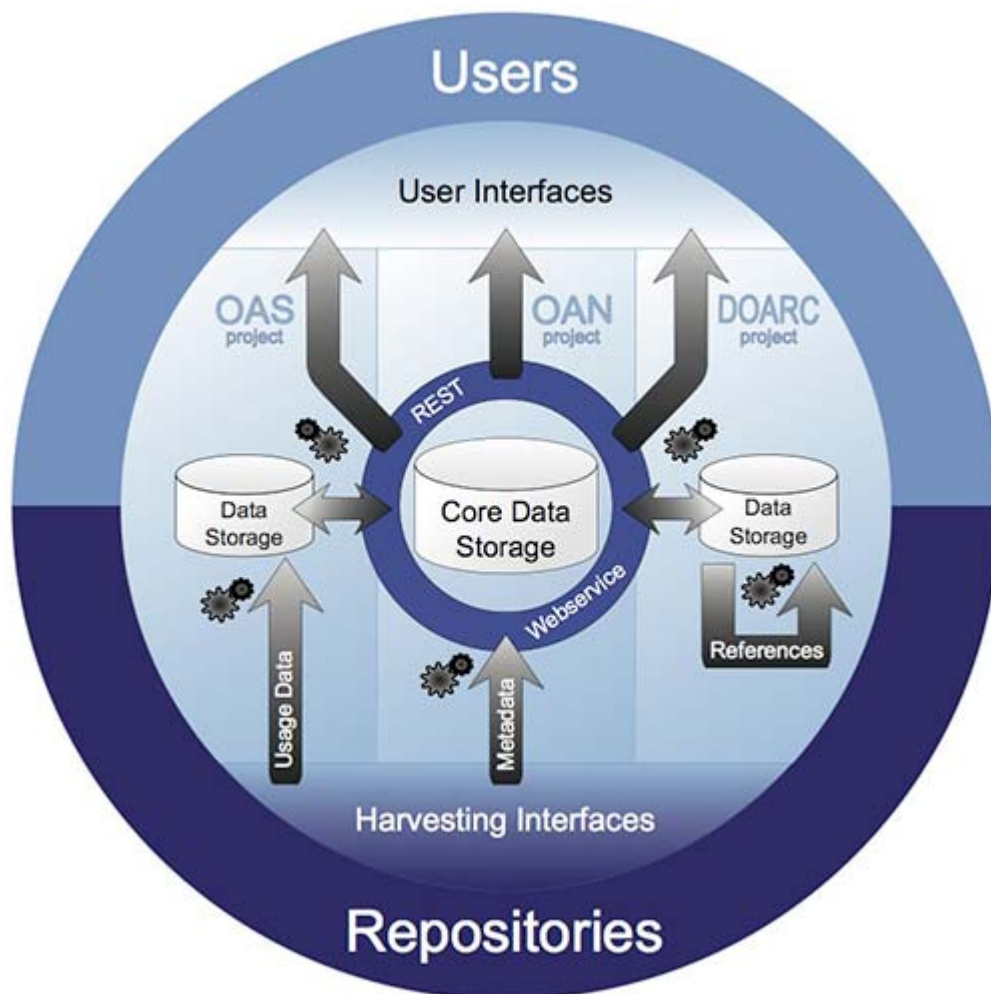
Comparé à un fonds imprimé, les dépôts numériques offrent une plus grande diversité d'opportunités pour développer les nouveaux services de valeur ajoutée et pour les approvisionner. Cela s'applique en particulier à l'assemblage de donnée webométrique comme l'utilisation de statistique ou de citation. En complément du projet de DFG et des "statistiques à libre accès" (OA-S), et le récent projet lancé appelé "Service de Consultation

¹⁹ OA-Netzwerk : <<http://www.dini.de/projekte/oa-netzwerk/>>.

²⁰ Pour consulter une liste des serveurs certifié-DINI voir : <http://www.dini.de/no_cache/service/dini-zertifikat/zertifizierte-server/>.

²¹ "Driver guidelines 2.0 Guidelines for Content providers – Exposing textual resources with OAI-PMH" November 2008 : <http://www.driver-support.eu/documents/DRIVER_Guidelines_v2_Final_2008-11-13.pdf>.

Répartition de Citation à Libre Accès (DOARC)²² valent aussi d'être mentionné. Dans ce service les exigences techniques devrait être mise à disponibilité pour une cohérence dans l'assemblage et l'analyse des citations dans les offres du libre accès. Motivé par le fait que- comme mentionné auparavant- le paysage scientifique dans le passé se servait de l'effet des modèles, qui souvent ne couvre pas les offres du libre accès et en plus étaient défaillant, les deux projets devraient développer des approches alternatives pour mesurer l'impact.²³ Les statistiques à libre accès peuvent donc être vu comme une partie de l'initiative de DINI pour construire un réseau de dépôts certifiés à travers l'Allemagne(figure 9).²⁴



Figures 9 : Structure et inter-relation des projets OA-statistics, OA-Network et DOARC²⁵

Le projet de l'UE "DRIVER" mentionné ci-dessus, dont le but est de construire un réseau de dépôt Européen, traite dans sa deuxième phase actuel de l'origine de bancs d'essai, laquelle sera répandu afin de financer les infrastructures de la technologie. En plus DRIVER projette

²² DOARC : <<http://doarc.projects.isn-oldenburg.de/>>

²³ Un autre projet semblable est celui de "CARPET" – Communauté Académique pour la Révision de la Technologie d'Édition et de Publication, ce projet vise à soutenir le bon rendement de l'usage des services et des outils électronique pour la publication scientifique. Pour plus d'information sur le CARPET : <
<http://www.dini.de/projekte/eng/carpet>>

²⁴ Pour ce sujet et les thèmes généraux de l'Open Acces en Allemand voir : Birgit Schmidt/Karin Iig-Hartbecke : Open Acces in Deutschland – Erweiterte Perspektiven für die Wissenschaft. In : GMS Medizin – Bibliothek – Information Volume 9, Issue 1, 2009 (forthcoming), <<http://egms.de/de/journals/mbi/index.shtml>>.

²⁵ Source : <<http://www.dini.de/projekte/oa-statistik/>>

de construire une structure organisationnelle durable comme la "Confédération des Dépôts Européen" et d'inclure explicitement les organisations régionales et nationales (DINI, SHERPA, etc.), les techniques d'initiatives, les promoteurs de logiciels et d'autre participants (ex : SPARC Europe, LIBER). Par ailleurs, DRIVER enquête dans sa phase actuelle sur les récentes techniques de développement et teste le contenu des publications (afin de les améliorer) aussi elle enquête sur la durée de vie des archives dans les dépôts qui emmagasinent ce genre de publications complexe.²⁶

Perspective : les statistiques à libre accès II et l'intégration des documents du patrimoine numérique.

Dès l'entame le projet de statistiques à libre accès était destiné à une application ultérieure. La première application se concentre principalement sur la conception et le model d'une infrastructure. Dans une application ultérieure, l'un des buts principaux devrait être d'intégrer d'autres dépôts institutionnels dans l'infrastructure qui était développé dans la première phase. En effet, plus de plan existe pour la seconde phase²⁷ du projet (la première phase prendra fin en février 2010). Comparé à l'actuel projet qui est principalement un projet de recherche, l'intention des projets OA-S II est de développer et d'établir une infrastructure. Plus loin le travail sera d'examiner les données collectées et regroupées concernant les caractéristiques qui devraient être développées et offertes sur la base de ces données, ex : les caractéristiques de recommandations basées sur les affinités des relations ou sur la similarité. En plus d'autre thèmes devraient avoir de l'intérêt et peuvent faire partis de nouvelles applications : de nouveau modèle métrique devrait être développé et évalué scientifiquement : mise en place de site web avec les informations concernant l'importance du projets (les lignes directives, FAQ, contact, offre d'assistance, etc.), la coopération avec d'autres acteurs, et aussi les acteurs commerciaux (comme les projets PEER²⁸ par exemple)

²⁶ Voir "Technology Watch Reports", <<http://www.driver-repository.eu/>>, sous la direction de "Documentations & Downloads".

²⁷ Ce chapitre parle seulement des thèmes d'initiative, les points mentionné ici ne sont finalement pas discuté ou ne font pas partie de l'application

²⁸ PEER – Publication des Recherches Ecologique Européenne :<Publishing and the Ecology of European Research>. Le projet Ec est l'innovateur de la collaboration entre la publicité, les dépôts et les chercheurs : "PEER, soutenu par le programme eContenplus de Ec, enquêtera sur les effets à grande échelle et (la mise en dépôt) systématique des auteurs. La revue finale manuscrite de PEER (aussi appelé Green Open Access/ Libre Accès à l'Ecologie ou stage- two research output/la section-deux de la productivité des recherches) sur l'accès au lecteur, la clarté de l'auteur, la viabilité des revues et aussi bien sur la large recherche écologiques européennes ;"Ibid.



Figure 10 : Goobi. L'utilisation des statistiques sur les dépôts numérisés²⁹

Une partie importante de la seconde phase du projet serait l'intégration de matériel de numérisation non numérique. Dans un contexte d'utilisation, les statistiques des documents du patrimoine numérisé seront intégrées automatiquement. L'Etat de Göttingen et la bibliothèque universitaire partenaire du projet OA-S est impliqué au cœur du développement de "Goobi"³⁰ : un logiciel qui inclut différents modules soutenant les processus de numérisation au sein des bibliothèques. Deux tâches essentielles constituent le centre d'intérêt de Goobi : soutenir la qualité supérieure de la numérisation et de subvenir aux besoins des chercheurs à propos de la présentation des objets numérisés. Un accent particulier est mis sur la quête et le traitement des métadonnées : Goobi utilise de manière cohérente le modèle de donnée METS³¹ qui est un "logiciel à code source ouvert établi pour la bibliothèque numérique" (voir fig.11)

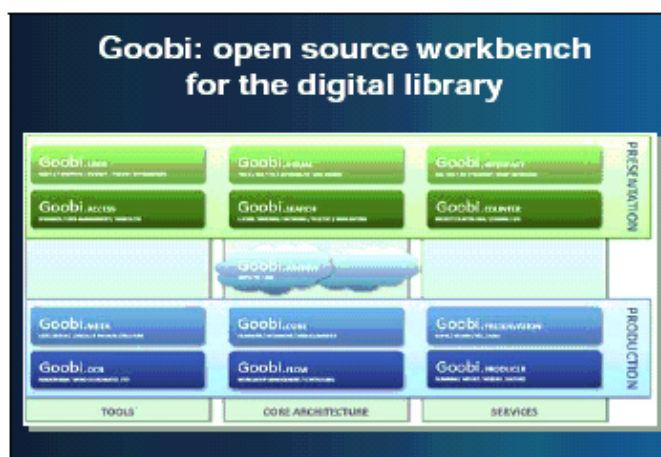


Figure 11 : Visualisation de différentes couches établies de Goobi

²⁹ Les diapositives de Goobi (Fig.10-12) sont confectionnées avec l'autorisation de Ralf Stockmann, SUB Göttingen.

³⁰ Goobi : les modules des bibliothèques numériques. <<http://www.goobi.org/>>.

³¹ METS – Codage de Métadonnée & Transmission de Norme : <<http://www.loc.gov/standards/mets/>>.

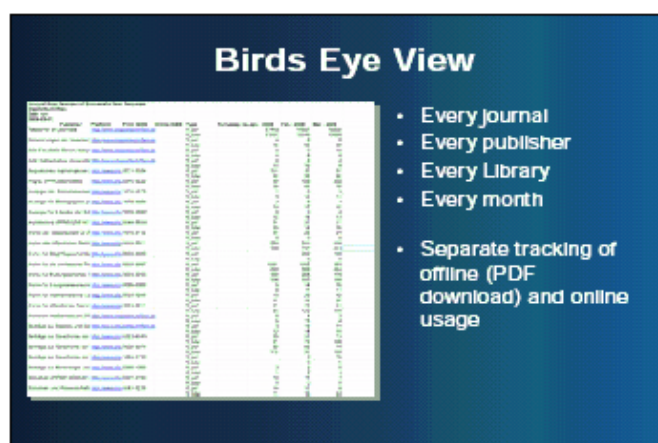


Figure 12 : "Birds Eye View" dans le cadre l'utilisation des statistiques créées par Goobi

L'un des modules de Goobi dans la section "présentation" est "Goobi.COUNTER" (fig.11). Goobi a une vaste expérience avec Counter (un exemple de la présentation de l'utilisation des statistiques est démontré sur la figure 12), aussi une collaboration entre "Goobi" et le projet OA-S laquelle dans le but de l'utilisation des statistiques analyse aussi COUNTER³² serait très promettant dans le domaine. Dans le Centre de Numérisation Rétrospectif de Göttingen (GTZ) plus de cinq millions de pages ont été numérisées jusqu'aujourd'hui.³³

Conclusion

Dans un article publié récemment, il est clairement mis en exergue que la numérisation peut aider à lire les textes anciens : "dans la version du 21^e siècle à l'époque des découvertes, les équipes scientifiques de l'informatique, les conservateur et les érudits se sont déployées à travers le monde dans une course à la numérisation des trésors littéraire en voie de disparition. Dans ce processus, de nouvelles trouvailles inattendues se sont opérées, incluant les versions jamais vues du Gospel Chrétien, des fragments de poète Grec et des commentaires sur Aristote. Le développement de la technologie a permit aux chercheurs de parcourir des anciens textes qui était jadis illisible—noircir par le feux ou par une érosion chimique, peint par-dessus ou tout simplement trop fragile à dérouler. Aujourd'hui les savants étudissent les travaux avec : les rayons fluorescents, le multi spectral numérique utilisé par la NASA pour photographier et aussi les scanographe CAT utilisé par les techniciens des examens médicaux."³⁴ Partant de cette déclaration, le nombre de documents numérisés s'accroîtra plus tard. Ainsi donc l'utilisation de ces contenus numérisés apparaîtrons aussi comme un fait important dans le futur. Par exemple, pour la gestion de collection, les propositions du projet concernant la numérisation et le financement de responsabilité civile.

³² Site web COUNTER : <<http://www.projectcounter.org/index.html>>.

³³ Pour plus d'information voir : Center for Retropsective Digitization, Göttingen (GDZ) <<http://gtz.sub.uni-goettingen.de/index.php?id=2&L=1>>.

³⁴ Alexandra Alter: The next Age of Discovery. In. The Wall Street Journal, May 8th 2009, <<http://online.wsj.com/article/SB124173896716198603.html>>.

Sans doute que les infrastructure, les services de métrique et de valeurs ajoutées développés dans le projet de statistiques à libre accès peuvent contribuer à l'utilisation des données statistiques qui sont utiles pour les deux objectifs : d'un côté pour le besoin d'utilisation des statistiques, des chercheurs ou des utilisateurs des matériaux³⁵ (à libre accès), et d'un autre côté, ex : pour les gérants de collection ou les bailleurs de fonds qui sont intéressés à la numérisation et à l'utilisation de statistiques provenant de différentes perspectives.

Presentation/contact

Dr. Matthias Schulze
University Stuttgart – University Library
Holzgartenstraße 16
70174 Stuttgart
Tel: +49 (0) 711 – 685 82509
Fax: +49 (0) 711 – 685 83502
E-Mail: mathias.schulze@ub.uni-stuttgart.de

Biographie:

Mathias Schulze a étudié l'histoire moderne et médiévale, la science politique et la communication à Göttingen, en Allemagne où il a obtenu une Maîtrise et un Ph.D en histoire. De plus, il a une Maîtrise en Bibliothéconomie et en science de l'information de l'université de Humbolt de Berlin. De 2004 à 2007, il a participé à plusieurs projets de bibliothèque numérique à la bibliothèque universitaire et dans l'Etat de Göttingen. Il a aussi travaillé dans les différentes bibliothèques universitaires des universités de Helmut Schmidt et des forces armées de Hamburg comme bibliothécaire spécialisé en sciences sociales et sciences humaines. Il a aussi donné des conférences sur l'alphabétisation. Depuis 2009, il est en fonction à la bibliothèque universitaire de Stuttgart en tant que bibliothécaire spécialisé en sciences sociales, en plus il est coordinateur de plusieurs projets de bibliothèque et de libre accès. Il est intéressé par tous les aspects de la bibliothèque numérique, spécialement sur le sujet de la préservation à long terme, du libre accès et de publication électronique.

³⁵

Et aussi pour l'augmentation des métriques existantes comme le Facteur d'Impact des Revues (JIF)