

De la numérisation à la diffusion en ligne de thèses et mémoires africains : l'expérience réussie d'une gestion de projet de bibliothèque numérique par le CAMES

Zakari LIRE

Conseil africain et malgache pour l'enseignement supérieur CAMES, Burkina Faso, 01 BP 134 Ouagadougou, zlire@hotmail.com

Abstract

Within the framework of the French Ministry of Foreign and European Affairs SIST (scientific and technical information systems) project, whose aim is disseminating and promoting African research, CAMES, the African and Malagasy Council for Higher Education, has successfully completed a project for the digitization and dissemination of theses and dissertations. The 2008/2009 project was launched during an African regional training workshop organized by CIRAD with support of CTA, FAO and IRD for the digitization, creation, management and dissemination of a collection of electronic documents. Based on the action plan designed during the workshop, the specially created CAMES project team coordinated and implemented all activities concerned with collecting, scanning, processing and managing electronic documents, and making the digital library available via the CAMES website.

The implementation of the project was supported and facilitated by technical workshops on digitization and digital library management organized in the framework of the SIST project. An awareness raising campaign targeting the academic community and scientific bodies of CAMES was organized to promote African research results and resulted in electronic submission and online dissemination of theses. The CAMES digital library was launched in 2009 during a session of the Inter Consultative Committee. A DVD version of the digital library was presented to all rectors, presidents, chancellors and directors of national research centers of the 18 member countries of CAMES. The sustainability of the CAMES digital library will be ensured with the implementation of the CAMES Center for Information and Documentation.

Résumé

Dans le cadre du projet SIST (système d'information scientifique et technique) du Ministère français des affaires étrangères et européennes (MAEE) qui vise à désenclaver et promouvoir la recherche africaine, le Conseil africain et malgache pour l'enseignement supérieur (CAMES) a mené avec succès un projet de numérisation et de diffusion des thèses et mémoires dont il est le dépositaire. Le projet conduit entre 2008 et 2009 a été lancé à l'occasion d'un atelier régional africain de formation à la numérisation, à la création, à la gestion et à la diffusion d'une collection de documents électroniques organisé par le CIRAD avec l'appui du CTA, de la FAO et de l'IRD. A partir du plan d'action élaboré lors de l'atelier, l'équipe projet du CAMES constituée à cet effet a coordonné et mené l'ensemble des opérations de collecte, de numérisation et de traitement des documents électroniques, ainsi que la gestion et la mise en ligne d'une bibliothèque numérique sur le site Internet du CAMES.

La mise en œuvre du projet a été accompagnée et facilitée par des ateliers techniques de numérisation et de gestion de bibliothèque électronique organisés dans le cadre de projet SIST. Une campagne de sensibilisation de la communauté universitaire et des instances scientifiques du CAMES à la valorisation des résultats de recherche africains a été menée et a abouti à la mise en place d'un dépôt électronique et la diffusion en ligne des thèses. La bibliothèque électronique du CAMES a été lancée en 2009 lors de la session des Comités Consultatifs Interafricains (CCI). Une version DVD de la bibliothèque numérique a été présentée à l'ensemble des recteurs, présidents, chanceliers d'université et des directeurs de centres nationaux de recherche des 18 pays membres du CAMES. La pérennité de la bibliothèque numérique du CAMES sera assurée avec la mise en œuvre du Centre d'information et de documentation du CAMES.

Introduction

Les technologies de l'information et de la communication ont durablement affecté tous les domaines de la vie économique et sociale. Le secteur de la gestion de l'information et de la documentation n'est pas en reste. Le concept de bibliothèque numérique est la preuve de cette mutation induite par les nouvelles technologies. Il est possible aujourd'hui d'avoir accès à des collections documentaires numérisées à partir de son ordinateur via Internet. Tirant profit de ces technologies, le Conseil africain et malgache pour l'enseignement supérieur, a initié et conduit, de septembre 2008 à décembre 2009, un projet de numérisation, avec l'appui du Ministère français des affaires étrangères à travers l'initiative SIST (Système d'information scientifique et technique). Le présent article fait l'économie de l'expérience du CAMES en matière de numérisation et de diffusion des thèses dont il est le dépositaire.

1. Présentation du projet de numérisation

Le contexte du projet

Créé en 1968, le Conseil africain et malgache pour l'enseignement supérieur est une institution régionale qui œuvre pour la promotion de la qualité de l'enseignement supérieur dans 18 pays en Afrique et à Madagascar à travers des programmes tels :

- la reconnaissance et l'équivalence des diplômes ;
- les comités consultatifs interafricains (CCI) ;
- les concours d'agrégation de médecine, de sciences juridiques, économiques et de gestion ;
- l'ordre international des palmes académiques du CAMES (OIPA/CAMES).

En plus de ces programmes statutaires et conformément à l'article 3 de la convention portant statuts du CAMES, une des missions essentielles de l'institution est la collecte, le traitement et la diffusion de l'information : documents universitaire ou de recherche (thèses), statistiques, annuaires, palmarès, informations sur les systèmes d'enseignement supérieur, les institutions universitaires, les programmes d'études, les diplômes etc.

C'est pour prendre en charge cette mission que le Conseil des ministres, en sa session d'avril 2001 tenue à Ouagadougou, par la résolution n°6, a décidé de créer le Centre d'information et de documentation du CAMES (CID-CAMES).

Pour la mise en œuvre des activités de ce centre, le CAMES a bénéficié d'un appui de la part de certains partenaires extérieurs dont le Ministère français des affaires étrangères à travers le projet SIST (Système d'information scientifique et technique). Le SIST vise en effet à désenclaver la recherche africaine, à promouvoir une dynamique de l'expertise et à mettre la science africaine au service du développement durable. Il se décline en trois volets : la mise en place d'un système d'information dans chaque pays partenaire du projet, la création de réseaux d'expertise sur des thèmes prioritaires, formations et transferts d'expertise. Il s'adresse en premier lieu aux chercheurs, aux laboratoires, aux enseignants et aux universités et favorise l'accès aux informations produites en Afrique et sur les autres continents, la diffusion de leurs recherches, la constitution d'équipes scientifiques sur des thématiques communes. Un plan d'actions a été proposé par le CAMES lors de l'atelier de formation à la numérisation, à la création, à la gestion et à la diffusion d'une collection de documents électroniques organisé par le CIRAD avec l'appui du CTA, de la FAO et de l'IRD à Madagascar en octobre 2007. Faisant suite à ce plan d'actions le CAMES a obtenu un financement de 18 000 € pour la numérisation et la mise en ligne d'une partie de son fonds documentaire de thèses et de mémoires. Ces thèses et mémoires, il faut le rappeler, ont été

déposés au CAMES par les enseignants chercheurs des universités et des centres de recherche de l'espace CAMES en vue de leur promotion dans des différents grades de l'enseignement supérieur et de la recherche (maître-assistant, maître de conférences, professeur titulaire pour les enseignants ; chargé de recherche, maître de recherche, directeur de recherche pour les chercheurs).

La mise en œuvre concrète du projet a commencé avec à un appui technique de l'Institut de recherche pour le développement (IRD).

Les objectifs du projet :

Le projet du CAMES avait pour objectifs :

- de diffuser dans des conditions optimales de communication et de consultation les thèses du CAMES ;
- de valoriser le fonds documentaire des thèses et des mémoires ;
- d'aider et susciter la recherche au niveau de la communauté scientifique et universitaire africaine et francophone.

Les résultats attendus étaient :

- la mise en ligne du catalogue des thèses ;
- la création d'une bibliothèque numérique à l'aide de Greenstone ;
- la mise en ligne de 500 thèses numérisées pour la phase I et de 1500 thèses pour la phase II, soit un total de 2000 thèses.

Les besoins

Pour la mise en œuvre du projet, les besoins exprimés étaient relatifs :

- aux ressources humaines : deux (2) documentalistes à plein temps, 1 ingénieur informaticien à temps partiel ;
- à l'équipement : deux (2) scanners de production, un (1) scanner A 3, deux (2) disques externes de stockage et un serveur de diffusion ;
- à la formation : prise en main des scanners de production et maîtrise des logiciels Greenstone et Adobe professionnel par les acteurs de la mise en œuvre du projet.

Les publics cibles

Les publics cibles du projet sont entre autres :

- la communauté universitaire et de recherche africaine francophone d'une manière générale.
- les enseignants chercheurs africains et/ou francophones;
- les doctorants.

Le fonds documentaire

Le fonds documentaire traité est estimé à 4000 thèses dont deux mille sont enregistrées dans une base de données Winisis. Il est essentiellement en format papier. Dans le cadre du projet, il a été envisagé de récupérer 500 thèses déjà numérisées en format image PDF pendant la phase pilote du projet de numérisation financée par l'Agence Universitaire de la Francophonie en 1997, mais ce corpus de documents, compte tenue de sa mauvaise qualité a été finalement mis de côté.

Le fonds qui est essentiellement en langue française, s'accroît annuellement de 500 nouvelles thèses.

2. Le circuit de numérisation et de production de documents électroniques

Le circuit de numérisation et de production des documents électroniques est le suivant :

- sélection des thèses dans le fonds documentaire ;
- massicotage à l'atelier de reprographie de l'Université de Ouagadougou,
- numérisation ;
- océrisation et optimisation par lot ;
- vérification et optimisation pour l'affichage web ;
- construction de la collection à partir de base de données Winisis ;
- publication sur le Web à partir du serveur du Campus numérique francophone de Ouagadougou.

Les opérations de numérisation sont faites par une équipe de deux documentalistes sous la supervision du chef de projet. La construction de la collection est faite avec l'appui d'un informaticien. A l'exception du massicotage qui se fait à l'Université de Ouagadougou, toutes les opérations sont faites en interne au CAMES.

Pour la mise en ligne, le CAMES dispose d'un espace sur le serveur du Campus numérique Francophone de Ouagadougou.

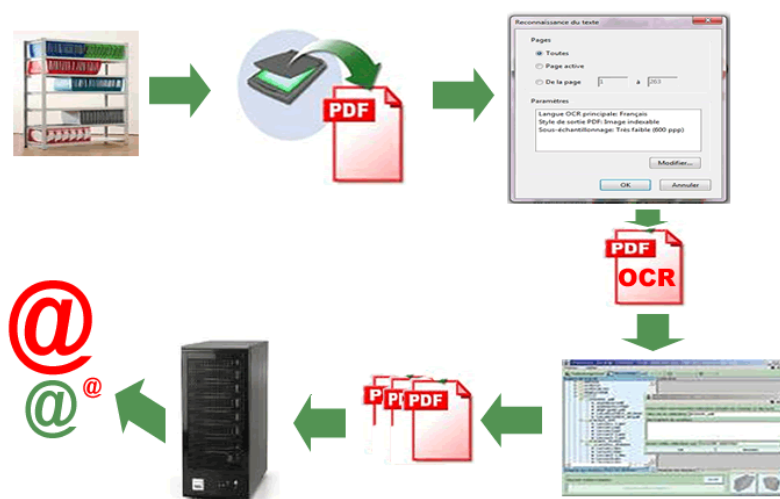


Fig. 1. Circuit de numérisation et de publication des documents

3. Les choix opérationnels et techniques

Les thèses sont numérisées en mode image à 300 dpi pour les documents en noir et blanc et à 150 dpi pour les documents en couleur et/ou niveau de gris. Le format de stockage et de diffusion est le PDF. Les équipements utilisés à cet effet sont les suivants :

- deux ordinateurs de bureau ;
- deux scanners de production de marque Fujitsu et un scanner A3;
- un ordinateur dédié à l'OCR ;
- une licence d'Adobe Acrobat 8 professionnel, deux licences de la version standard livrée avec les scanners à l'acquisition.

Une fois numérisés, les documents subissent des traitements complémentaires tels l'océrisation et optimisation des fichiers pour un affichage rapide sur le web, l'intégration des vignettes. Tous ces traitements se font par lots et permettent ainsi de gagner en temps. Les métadonnées sont extraites de la base de données Winisis et associées aux fichiers PDF. La classification des références se fait suivant les catégories du plan de classement PASCAL dont les principales sont :

- Sciences exactes
- Sciences appliquées
- Sciences de la terre, des océans, de l'espace ;
- Sciences de la vie
- Sciences de la santé
- Sciences humaines
- Sciences sociales
- Etudes.

Les métadonnées suivantes sont utilisées pour la description des documents : auteur, directeur de thèse, titre de la thèse, date de soutenance, type de thèse, spécialité, université de soutenance, pagination

4. Modalités d'accès et de diffusion du fonds numérisé

Le fonds est en accès libre à partir du site web du CAMES à l'adresse suivante : <http://greenstone.refer.bf/cgi-bin/library?>. Il est possible de rechercher l'information par l'auteur, le titre de la thèse, les mots clés, la catégorie et/ou la sous-catégorie, le texte intégral. Les fonctionnalités proposées sont la recherche, l'affichage des résultats et la consultation.

La mise en accès et/ou la diffusion a été réalisée à l'aide de Greenstone qui est un logiciel libre dédiée à création et la gestion de collections numériques destinées à être publiées en ligne ou sur CD-ROM.

La bibliothèque numérique du CAMES est référencée sur le site exemples de Greenstone : <http://www.greenstone.org/examples>



Fig. 2. Page d'accueil du site web du CAMES

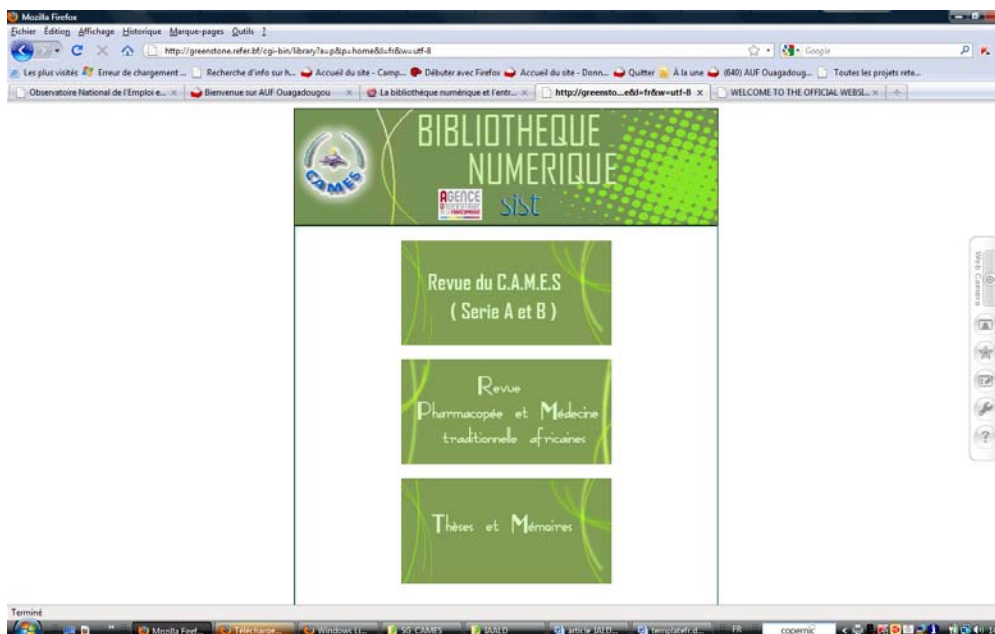


Fig. 3. Page d'accueil de la bibliothèque numérique du CAMES



Fig. 4. Formulaire de recherche de la bibliothèque numérique

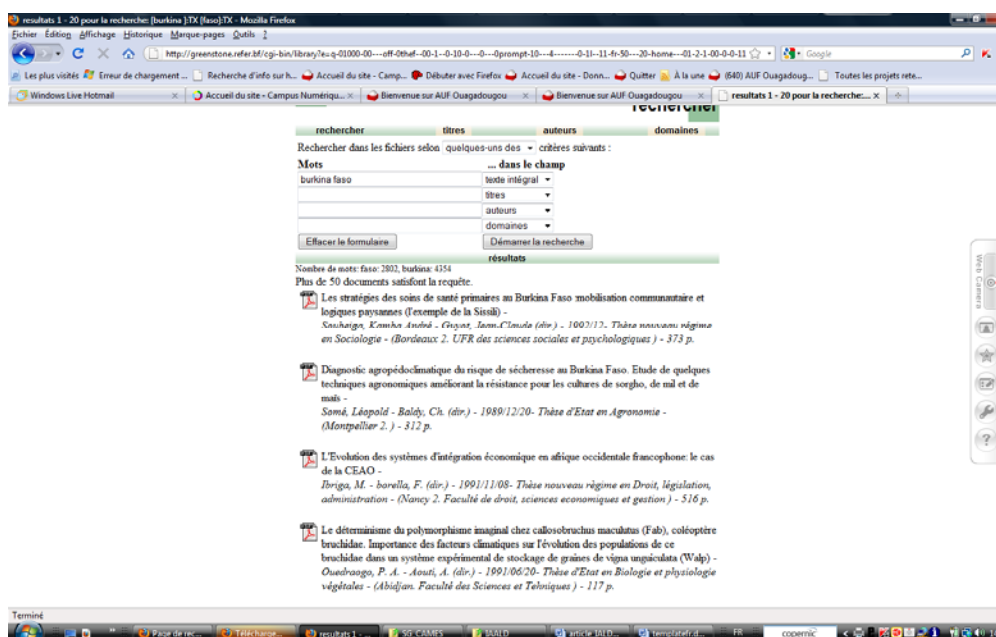


Fig. 5. Affichage des résultats de la recherche

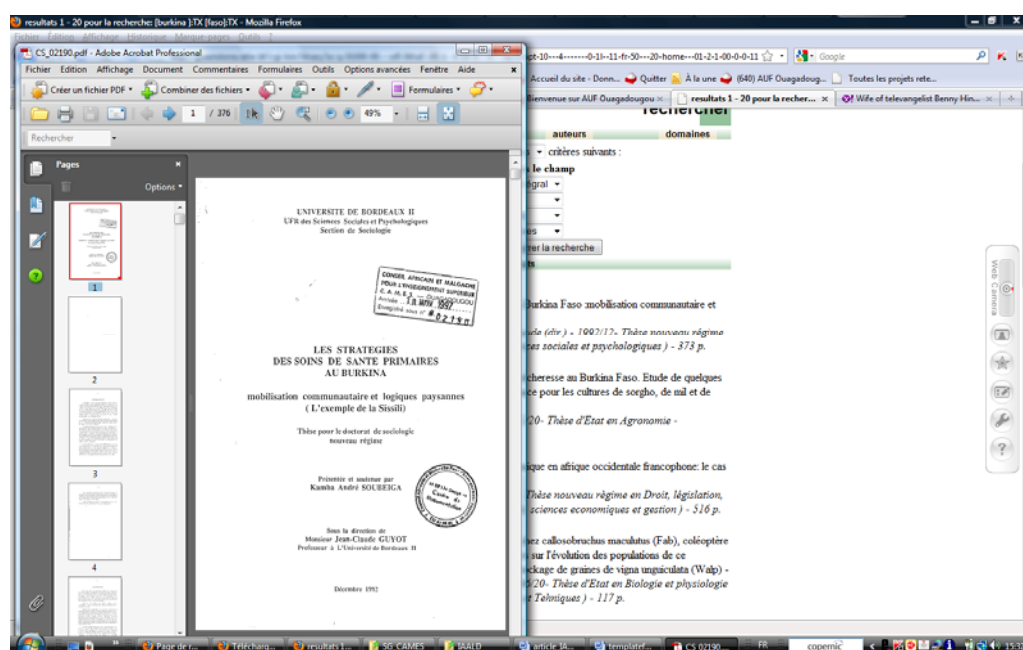


Fig.5. affichage du document (thèse)

5. La gestion des droits d'auteur

Nul n'est censé ignorer la loi, a-t-on coutume de dire. Cela est aussi vrai pour la gestion du droit d'auteur. Quand nous parlons de droits d'auteur, il s'agit pour l'essentiel des droits patrimoniaux liés à la reproduction des documents le droit moral étant perpétuel et imprescriptible. La gestion du droit d'auteur est avant tout une question de sensibilisation sur les avantages qu'offre la diffusion électronique. En effet, une large diffusion de la thèse est une valorisation de celle-ci et de son auteur. Sans disposer de données statistiques, l'expérience sur la diffusion des thèses en format papier au sein des structures documentaires montre que les thèses sont en général très peu consultées pas par défaut de

qualité mais parce qu'elles ne sont simplement pas connues de la communauté des utilisateurs potentiels (enseignants, chercheurs, doctorants). La diffusion en ligne apparaît nettement ici comme un moyen de remédier à cette situation et donc de faire connaître ces travaux contenant des données scientifiques et/ou empiriques sur nos pays qu'il est parfois impossible de trouver ailleurs.

Fort de ce constat, le CAMES a procédé à la sensibilisation de la communauté universitaire de son espace et a pu ainsi obtenir de ses instances scientifiques tout le soutien nécessaire à la mise en œuvre du projet de numérisation, ceci dans le strict respect du droit d'auteur. Cela s'est traduit par une décision du Comité Consultatif Général instituant le dépôt numérique des thèses et une sorte de « moratoire » sur le droit d'auteur permettant ainsi de diffuser en ligne les thèses n'ayant pas fait l'objet de dépôt électronique au CAMES. Toutefois, les titulaires des droits ont la possibilité de mettre fin à la diffusion de leurs œuvres s'ils le souhaitent. Les thèses sont diffusées en texte intégral depuis 2008, et nous n'avons, à ce jour, enregistré aucune demande de retrait de la diffusion en ligne. Bien au contraire, la plupart des enseignants chercheurs souhaitent que leurs thèses soient diffusées le plus rapidement possible.

Depuis le début de l'année 2009, le dépôt électronique est effectivement mis en œuvre. Les auteurs remplissent un formulaire de dépôt et une autorisation formelle de diffusion en ligne.

6. Evaluation des ressources du projet de numérisation

Le coût total du projet est de 18 000 € dont 8 000 € correspondant à la première phase et 10 000 € à la dernière. Ce montant a permis de couvrir les dépenses relative à l'équipement d'un atelier de numérisation et surtout d'assurer le paiement des salaires de deux documentalistes et de l'informaticien jusqu'à la fin du projet en décembre 2009. Le coût total des équipements s'élève à 2 500 € (2 scanners de production, un scanner A3, deux disques durs externes de stockage). Le plus gros poste budgétaire reste à n'en pas douter, le volet ressources humaines. 15 500 € y ont été consacrés en termes de vacations et de contrats à durée déterminée. La durée totale du projet est de 18 mois dont :

- 3 mois pour la mise en route du projet avec une phase de traitement documentaire de quelques thèses ;
- 2 semaines de formation ;
- 14 mois pour la numérisation et la mise en ligne des documents.

Qualifications des intervenants

Le projet a été conduit par une équipe de deux documentalistes respectivement de niveau licence et de niveau BAC (avec une bonne expérience de plusieurs années), un ingénieur informaticien de niveau BAC+3 à temps partiel. Toute l'équipe a été piloté par un chef de projet documentaire de niveau Master en information documentation et applications informatiques.

7. Formation et communication sur le projet

Actions de formation mises en œuvre

Le projet a été mis en œuvre grâce en partie à des actions de formation avec l'appui technique de l'IRD. En dehors de l'atelier de Madagascar, le projet a bénéficié de deux autres ateliers. Le premier est intervenu au lancement du projet en septembre 2008 et il était dédié à l'utilisation du matériel de production (les scanners), aux techniques de numérisation et à l'utilisation du logiciel Adobe Acrobat professionnel pour les traitements du corpus numérique. Le second atelier qui a eu lieu en décembre de

la même année, a porté essentiellement sur la construction des collections numériques à l'aide de Greenstone.

Actions de communication du projet

Deux actions majeures ont été entreprises en termes de communication autour du projet de numérisation. La première est relative à la sensibilisation des instances scientifiques du CAMES, notamment le Comité Consultatif Général (CCG) sur le bien fondé de la diffusion électronique des thèses et la nécessité du dépôt électronique. Cette action a abouti en juillet 2008 à l'adoption du formulaire de dépôt électronique et l'autorisation de procéder à la diffusion en ligne.

La deuxième action phare de communication a été le lancement de la bibliothèque électronique lors de la 31^{ème} session des Comités Consultatifs Interafricains (CCI) à Yaoundé en juillet 2009. A cette occasion une version DVD de la bibliothèque numérique a été présentée au Comité Consultatif Général (CCG). Le CCG est formé de l'ensemble des recteurs, présidents, chanceliers d'université et des directeurs de centres nationaux de recherche des 18 pays membres du CAMES.

8. Bilan global du projet de numérisation

Résultats positifs

La fin du projet est intervenue en décembre 2009. A en juger par les résultats, le bilan global est largement positif. Au total :

- 1 700 thèses ont été déjà numérisées et océrisées ;
- une bibliothèque numérique comprenant trois collections (thèses et mémoires, pharmacopée et médecine traditionnelle africaines, revues CAMES) a été créée à l'aide de l'outil Greenstone (<http://greenstone.refer.bf>).

Le projet aura également permis de capitaliser les acquis suivants :

- la mise en place d'un atelier de numérisation ;
- la formation du personnel du projet à la numérisation, à la création et mise en ligne de collections numériques grâce à l'outil Greenstone ;
- le développement d'une expertise dans le domaine des bibliothèques numériques.

Difficultés rencontrées : humaines, techniques, financières

Bien évidemment comme toute entreprise de ce genre, le projet a connu quelques difficultés :

- le début tardif des opérations de numérisation compte tenu la mise en œuvre tardive du premier atelier de numérisation (scanners + formation) ;
- un équipement insuffisant en ordinateurs de bureau : deux ordinateurs servaient en même temps pour la numérisation et l'océrisation des documents (facteur contre-productif), du moins dans la première phase du projet.
- le problème lié à la précarité du contrat étant donné que l'équipe du projet en dehors du chef de assurait plutôt des vacances, ce qui n'est pas certainement pas une source de motivation.

Le souci majeur sur le plan technique, c'est la construction des collections à partir d'une base Winisis. Cette opération est souvent laborieuse et nécessite des aménagements de la base existante. L'autre souci est lié à l'usage même de Greenstone : chaque ajout de documents à la collection donne lieu à une reconstruction. Enfin, le manque d'autonomie du CAMES au niveau de l'hébergement est également un problème notable.

9. Perspectives d'évolution du projet

Les objectifs fixés par le projet initial ont été largement atteints, en dépit des difficultés, qui ont sans doute retardé le travail. L'équipe de projet a travaillé avec dévouement et sans relâche sous la supervision directe du chef de projet. Au terme du projet, le CAMES dispose d'une collection numérique d'une partie de ses thèses, d'une expertise dans le domaine de la numérisation et de collections numériques en ligne. Il serait donc souhaitable que les autorités du CAMES prennent les mesures adéquates pour la pérennisation de cette opération de numérisation par le renforcement des capacités informatiques du CAMES (Serveur de diffusion) et la stabilisation de certains postes clés (documentaliste + informaticien). La poursuite de l'opération de numérisation permettra à terme de diffuser en ligne toutes les thèses des enseignants chercheurs de l'espace CAMES et de donner une meilleure visibilité à la recherche scientifique africaine. L'inauguration très prochaine du Centre d'information et de documentation devra permettre la pérennisation des actions de ce projet.

Conclusion

A l'analyse, le projet de numérisation et de diffusion des thèses au CAMES a été mené avec succès. La collection diffusée en ligne comprend 1700 thèses et mémoires contribuant ainsi à donner une meilleure visibilité à une part non négligeable des résultats de la recherche scientifique africaine. Ceci s'inscrit dans la dynamique de mise en œuvre du Centre d'information et de documentation dont l'ouverture très prochaine permettra de prendre en charge l'une des missions statutaires du CAMES qui est la collecte et la diffusion de l'information scientifique et technique.

Bibliographie

- BURESI, C., CEDELLE-JOUBERT, L., 2002. Conduire un projet de numérisation, Paris, France : Tec et Doc, 326 p.
- SCOPSI, C., SOUAL, L., 2007. Mener un projet open source en bibliothèque, documentation et archives, Paris, France : Cercle de la librairie, 155 p.
- JACQUESSON, A., RIVIER, A., 2005. Bibliothèques et documents numériques : concepts, composantes, techniques et enjeux, Paris, France : Cercle de la librairie, 573 p.