



Date : 11/05/2006

**“dCollection”; The National Network of University-Based Repository in Korea.**

**Park, Hong Seok**  
KERIS<sup>1</sup>  
[honspark@keris.or.kr](mailto:honspark@keris.or.kr)

**Jang, Keum Yeon**  
KERIS  
[jkylib@keris.or.kr](mailto:jkylib@keris.or.kr)

**Meeting:** 157 Asia and Oceania

**Simultaneous Interpretation:** -

WORLD LIBRARY AND INFORMATION CONGRESS: 72ND IFLA GENERAL CONFERENCE AND COUNCIL  
20-24 August 2006, Seoul, Korea  
<http://www.ifla.org/IV/ifla72/index.htm>

**Traduction**

**Rosa María Gómez de Regil**  
[rosa-maria.gomez@insa-lyon.fr](mailto:rosa-maria.gomez@insa-lyon.fr)  
Doc'INSA – INSA de Lyon  
France

## Résumé

*Le KERIS (Korea Education & Research Information Service)<sup>1</sup> est en train de créer un système national appelé « dCollection ». Le système est en cours de développement depuis 2003. Il se compose d'un fournisseur de données (DP – data provider) et d'un fournisseur de service (SP – service provider). Le DP permet de collecter les métadonnées et les documents numériques comme les thèses et les articles scientifiques des communautés scientifiques au sein de l'université. Le SP moissonne périodiquement les métadonnées du fournisseur de service et les intègre dans un service national. Les principales fonctions du DP dans l'université sont la collecte de métadonnées et de documents électroniques, la gestion de la qualité des données collectées, rend les métadonnées disponibles au moissonnage par le SP et la gestion des droits d'auteur. Les principales fonctions du SP de dCollection au niveau du KERIS sont de moissonner les métadonnées des DP, gérer les fichiers log, le planning et le flux des fournisseurs de données, et d'intégrer les données fournies par les DP. Quarante universités utilisent ce système. Cette année, nous allons développer un système d'hébergement sur Internet pour dCollection afin de toucher les petites universités. Le système dCollection sera connecté avec d'importants dépôts d'archives ouvertes d'autres pays. Dans un futur proche, dCollection va dépasser les frontières nationales du système d'information et de recherche, et devenir un centre international pour les flux d'informations de recherche.*

<sup>1</sup> Le KERIS (Korea Education & Research Information Service) est une institution publique sous la tutelle du Ministère de l'Éducation et du Développement des ressources humaines. Il a été créé afin d'informatiser les bibliothèques universitaires et les établissements scolaires du primaire et du secondaire.

## 1. Introduction

Depuis quelque temps, on s'intéresse de près au libre accès à l'information, notamment aux informations de recherche. Cette prise de conscience du libre accès à l'information est liée aux récentes évolutions du marché de l'information scientifique : la prédominance de quelques éditeurs de bases de données, qui augmentent de plus en plus les prix des revues scientifiques, la limitation des budgets des bibliothèques, la reconnaissance du libre accès à l'information par les auteurs, etc. Le libre accès à l'information scientifique facilite le partage des résultats de la recherche, et augmente la visibilité des productions des auteurs. Le libre accès à l'information est une façon de faire valoir les droits des auteurs, qui ne sont généralement pas rétribués pour leurs publications.

L'auto-archivage des papiers de recherche d'une institution en pré- ou post-publication est l'un des aspects du libre accès à l'information. En 2004, le comité parlementaire de science et technologie du Royaume Uni a publié un rapport recommandant l'ouverture d'archives institutionnelles pour offrir un accès ouvert et en ligne des informations provenant des organisations publiques<sup>2</sup>. En Corée, le KERIS (Korea Education & Research Information Service) a ouvert un système d'archives institutionnelles, appelé « dCollection » pour « digital Collections », aux bibliothèques universitaires, et a mis en place un système central pour intégrer les systèmes distribués des universités au partage des ressources. L'objectif de notre article est de présenter dCollection et les résultats obtenus sur ses deux années de fonctionnement.

## 2. Bref aperçu de l'archive institutionnelle

Il existe deux types d'archives : les archives disciplinaires, comme ArXiv.org pour la physique et les disciplines connexes ou le NCSTRL pour l'informatique et les disciplines connexes, et les archives institutionnelles comme DSpace du MIT ou ePrints de l'université de Southampton. Toutes ces archives servent à conserver les documents numériques et à assurer le libre accès à l'information par dessus les barrières de la loi, l'économie et la technique.

Raym Crow<sup>3</sup> définit l'archive institutionnelle universitaire comme « les collections numériques qui collectent et conservent la production intellectuelle d'une université ou d'un ensemble d'universités », et en signale deux aspects stratégiques :

1) la réforme de la communication scientifique qui « étend l'accès à la recherche, réaffirme le contrôle des bourses d'études par les universités, augmente la compétitivité et réduit le monopole des revues, soulage économiquement et augmente la visibilité des établissements et de leurs bibliothèques » ;

2) le développement de la compétitivité des universités avec « le potentiel de servir d'indicateur de qualité d'une université et de démontrer la pertinence scientifique, sociétale et économique de ses activités de recherche, en augmentant la visibilité de l'institution, son statut et sa valeur publique ».

La définition et les points de vue stratégiques mentionnés ci-dessus sont pris en compte dans le système dCollection du KERIS. Les objectifs de dCollection ne diffèrent pas de ceux de DSpace : fortifier la compétitivité des bibliothèques universitaires pour la gestion et la fourniture d'information, mettre en valeur la connaissance du rôle des bibliothèques à l'intérieur et à l'extérieur de l'université et augmenter la valeur des universités qui participent à dCollection.

---

<sup>2</sup> Albanese, A. 2004 UK Report Calls for Publicly Available STM Research. *Library Journal*, 129(13): 16-17

<sup>3</sup> Crow, R. 2002. The Case for Institutional Repositories: A SPARC Position Paper. Washington, DC: SPARC. <[http://www.arl.org/sparc/IR/IR\\_Final\\_Release\\_102.pdf](http://www.arl.org/sparc/IR/IR_Final_Release_102.pdf)>

### 3. Principales fonctions du système dCollection de KERIS

Le projet dCollection a été lancé en 2003. La première version du système a été développée dans 20 bibliothèques universitaires : quatre ont participé au projet pilote en 2003 et les 16 autres ont rejoint le projet l'année suivante. Jusqu'en 2005, 3.5 millions de dollars ont été investis pour développer le système, le diffuser, faire migrer les données et les corriger, et mettre en place des programmes de formation et de promotion auprès de 40 universités.

dCollection a quatre particularités : le fournisseur de données (DP) de la bibliothèque universitaire collecte et gère les métadonnées et les documents numériques ; le fournisseur de service (SP) de KERIS collecte et intègre les métadonnées provenant des systèmes DP ; le protocole OAI-PMH (Open Archive Initiative Protocol for Metadata Harvesting) assure la communication entre le système du DP et celui du SP ; et la gestion des droits d'auteur est adaptée à l'accès ouvert à l'information et dCollection prémunit les informations d'une mauvaise utilisation.

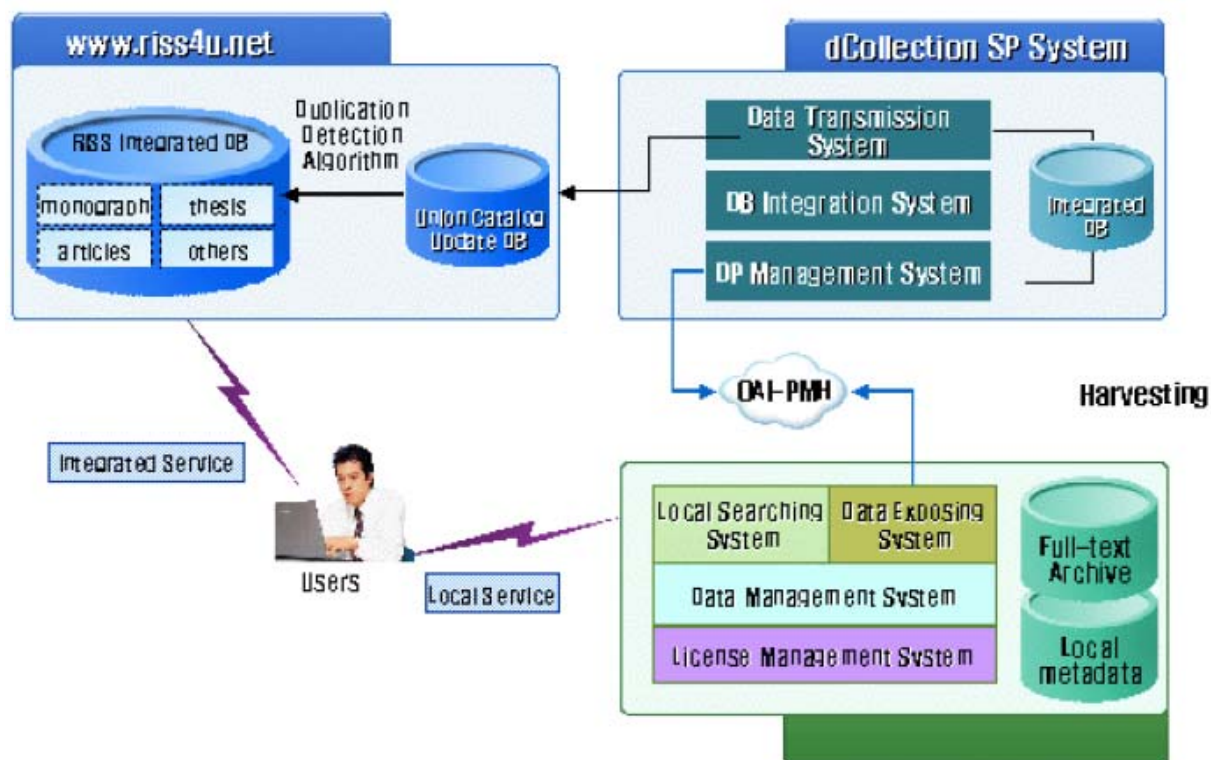
Le flux de fonctionnement général de dCollection est linéaire. Une bibliothèque universitaire, sur un système dCollection installé ou sur son propre système, collecte les métadonnées et les documents numériques auprès de l'auteur. Le bibliothécaire vérifie et corrige les métadonnées et les documents collectés. Si nécessaire, le bibliothécaire demande un nouveau dépôt des documents. Le document numérique est automatiquement converti au format PDF. Il est accessible par tous les services de l'université. Le SP du KERIS moissonne les métadonnées des DP des universités et les intègre à son propre système. Toutes les métadonnées collectées sont transférées au RISS, le système national intégré d'information pour la recherche fourni par le KERIS.

### 4. Architecture du système dCollection

dCollection, nom donné par le KERIS pour « digital collection », suit le modèle SP/DP de l'OAI<sup>4</sup>. L'architecture du système est présentée dans la Figure 1. Les bibliothèques qui ont participé au projet collectent, gèrent, conservent et partagent des contenus numériques. Le protocole OAI-PMH (OAI Protocol for Metadata Harvesting) utilise le langage HTTP pour les applications Web et le XML et SOAP (Simple Object Access Protocol) pour l'échange de données.

---

<sup>4</sup> « L'Open Archive Initiative développe et promeut des normes d'interopérabilité qui ont pour objectif de faciliter la diffusion efficace des contenus. L'Open Archive Initiative est née de la volonté d'améliorer l'accès aux archives électroniques pour augmenter la disponibilité de la communication scientifique. »  
<http://www.openarchives.org/organization/index.html>



[Figure 1.] Architecture du système dCollection

Dans ce système, les documents numériques de l'université sont accessibles au grand public et aux usagers de l'université. Le chemin classique du service est :

- 1) enregistrement de la provenance scientifique (département, laboratoire, etc.) et du type de document (thèse, article, etc.) dans le système dCollection local,
- 2) dépôt des documents et des métadonnées,
- 3) accord de l'auteur pour le libre accès au document,
- 4) vérification et correction d'erreurs par le bibliothécaire,
- 5) demande à l'auteur (ou à la communauté) de corriger ou re-déposer les métadonnées ou le document,
- 6) accord du bibliothécaire pour la mise à disposition dans le service local,
- 7) moissonnage des métadonnées locales par le SP,
- 8) vérification de la qualité et amélioration par le KERIS,
- 9) intégration avec les données anciennes,
- 10) le RISS (Research Information Service System) du KERIS installe le système DP de dCollection dans la bibliothèque universitaire. La fonction principale du DP est de collecter les métadonnées et les documents numériques, de gérer la qualité des métadonnées collectées et des documents, de rendre les métadonnées disponibles au moissonnage par le SP, et de gérer les droits d'auteur.

Le système est composé de quatre modules principaux :

- Système de gestion des données : collecte et re-collecte des métadonnées et des documents des auteurs, vérification et correction d'erreurs dans les métadonnées et les documents, information sur les erreurs et demande d'un nouveau dépôt à l'auteur
- Système de mise à disposition des données : réponse aux requêtes du DP et mise des métadonnées collectées à disposition du moissonnage par le SP
- Système de recherche locale : système de recherche dans les données locales pour le bibliothécaire et l'utilisateur de l'institution

- Système de gestion des licences : information de l'utilisateur sur les conditions et options d'utilisation du document ; protection du document conformément aux lois relatives au droit d'auteur

La principale fonction du SP de KERIS est de moissonner et d'intégrer les métadonnées des DP. Le système a trois modules principaux :

- Système de gestion des DP : gestion des logs, des plannings et du flux du fournisseur de données et du moissonnage des données
- Système d'intégration de la base de données : enlever les doublons et intégrer les données des fournisseurs de données ; gérer les métadonnées intégrées de dCollection
- Système de transmission des données : corriger les erreurs dans les métadonnées intégrées ; et transférer les données au système du service.

## 5. À propos des droits d'auteur

Les droits d'auteur forment une question essentielle dans le cadre d'un service d'information. Ils sont gérés de trois façons. Tout d'abord, un accord de l'auteur est nécessaire pour la diffusion du document en libre accès. Un formulaire d'accord a été établi par le KERIS et partagé par les bibliothèques membres du projet avec des modifications mineures. Le second est l'établissement des conditions d'utilisation conformément aux licences Creative Commons<sup>5</sup>. Les conditions d'utilisation sont « Paternité » « Absence d'oeuvre dérivée », « Pas d'usage commercial », et « Partage à l'identique ». La dernière relève de la protection du document. La bibliothèque peut limiter l'utilisation du document à l'affichage, à l'impression ou à l'enregistrement, afin de le protéger des mauvais usages.

## 6. État actuel du service dCollection

40 universités participent au projet dCollection. Elles fournissent environ 74 000 documents au total, des thèses de master ou de doctorat pour la plupart (93% approximativement) comme le montre le Tableau 1. De nombreuses universités exigent de leurs étudiants diplômés un dépôt de leur thèse pour leur conservation et leur diffusion. Les étudiants diplômés acceptent ou non la diffusion de leur thèse en libre accès. La plupart des étudiants acceptent le libre accès. Pour les articles de revues, la bibliothèque contacte directement l'individu, le laboratoire ou le professeur. Très peu d'universités règlementent le dépôt des papiers de recherche des professeurs.

Dans la plupart des universités, le système dCollection est essentiellement utilisé pour la conservation et au service des thèses. Les thèses mises à part, il existe plusieurs types de documents nécessitant une collecte et un partage entre universités. Afin de faciliter et étendre l'utilisation de dCollection, le KERIS monte un programme-type destiné aux bibliothèques universitaires afin de proposer un modèle de bonnes pratiques pour dCollection. En 2005, l'Université Nationale de Pusan et l'Université de Sungkyunkwan ont été désignées comme universités modèles. Les produits phares développés par ces universités dans le cadre du programme sont la préparation d'une réglementation universitaire pour le dépôt numérique des thèses et des papiers de recherche, un contrat avec les laboratoires de l'université pour participer à dCollection et la rédaction d'un manuel pour assurer le dépôt correct des documents et gérer les informations de la recherche en coopération avec le centre de Recherche & Développement (R&D) de l'université. Ces produits ont été diffusés à l'université entière afin de partager l'expérience des bibliothèques universitaires modèles.

---

<sup>5</sup> CCL(creative commons license, <http://creativecommons.org>)

Afin d'aider les bibliothèques universitaires et de raccourcir le temps de dépôt dans un service national, le KERIS a centralisé sa gestion qualité en 2005. Cela a permis de réduire la charge de travail des bibliothécaires universitaires dans le cadre de dCollection. Le dépôt des thèses se concentre essentiellement sur trois mois dans l'année. La plupart des bibliothèques consacrent moins d'un professionnel au système dCollection. L'activité principale de la gestion centralisée de la qualité est la correction des métadonnées et des documents numériques que l'auteur dépose par l'auteur.

## **7. Perspectives**

Ce projet a commencé en 2003 et il est actuellement en phase de développement. Pour poursuivre sa croissance, le système a besoin de s'étendre au niveau national et international. Tout d'abord, d'ici un an, un système web d'hébergement va être développé pour les petites universités. Une bibliothèque universitaire sera libérée du fardeau que représentent les opérations de maintenance du système. dCollection sera distribué par l'hébergeur web dans toute l'université et il sera le meilleur moyen de connecter toutes les bibliothèques numériques universitaires de Corée.

En deuxième lieu, un système de paiement des droits est en cours de développement pour les documents dont la diffusion en libre accès n'a pas été autorisée par leurs auteurs. Selon le droit d'auteur coréen, un document publié peut être transféré entre bibliothèques sous un format numérique cinq ans après sa publication, en contrepartie du paiement des droits d'auteur correspondants. Avec le système de paiement des droits, le partage des ressources numériques entre les universités va s'étendre.

En troisième lieu, dCollection sera connecté aux principales archives ouvertes d'autres pays : aux archives des universités à grande échelle, puis aux journaux électroniques et enfin à d'autres archives. Un projet à moyen terme visera le réseau international d'archives ouvertes des informations de recherche.

Enfin, toutes les universités participant à ce projet seront évaluées. Pour être considérée comme réussie, une archive universitaire doit répondre à différents critères. Par exemple, la mise à jour et la diversité des collections numériques sont importantes pour la satisfaction de l'utilisateur. Le système doit également fournir une interface facile à utiliser par les auteurs qui déposent des métadonnées et des documents. Si une université a une réglementation pour dCollection, le système pourra facilement être répandu dans la communauté universitaire. Les critères d'évaluation sont des facteurs de succès importants pour dCollection. Toutes les universités participantes seront évaluées selon ces critères afin de pointer les forces et faiblesses de chacune d'entre elles, et de fournir des réponses personnalisées à chaque université. Les critères d'évaluation sont listés dans le Tableau 2.

## **8. Conclusion**

Ce système, que nous avons appelé « dCollection », est le produit de la coopération entre le KERIS et 40 universités pendant les trois dernières années. Le facteur de succès dominant est la participation active des utilisateurs (bibliothèques universitaires) dans le processus de développement du système. Leur avis a été pris en compte dans le système. La participation active des bibliothèques universitaires est à souligner. À partir de l'année prochaine, le système va se répandre rapidement grâce au système d'hébergement web. dCollection va devenir un système national d'information de recherche et un centre d'import des flux internationaux d'informations de recherche.

	Thesis	Journal Article	Total
Daegu University	3,913	2,555	6,468
Pusan National University	4,660	509	5,169
Kyungpook National University	2,399	-	2,399
Sungkyunkwan University	2,353	2	2,355
Ewha womans University	2,378	401	2,779
Chungnam National University	2,259	-	2,259
Gyeongsang National University	1,128	760	1,888
Hongik University	953	343	1,296
University of Seoul	463	452	915
Keimyung University	878	-	878
others	47,659	161	47,820
<b>Total</b>	<b>69,043</b>	<b>5,183</b>	<b>74,226</b>

<Tableau 1> Nombre de contenus dans le service (31 Mars 2006)

Evaluation Area	Evaluation Criteria
Digital Collection and Metadata	Diversity of digital collection, Up-to-dateness of contents, Proportion of documents submitted to dCollection, # of access points (# of metadata elements), Quality of metadata
System/Network	Integration of service with OPAC, Linkage with related administration system, Quality of homepage, Additional services for community, # of linkage with other open archives
Policy and Management	University committee for dCollection, Flexibility of document submission, dCollection advertisement in the university, Portion of budget for dCollection, Portion of dCollection staff member, Provision of copyright information, Formalized dCollection manual, Incentive for self-archiving, Diversity of digital collection development, A library organization for dCollection, Formalized preservation policy, Awareness of librarians/head of library/student/professor on dCollection.
Author/ User	Satisfaction of the user, Satisfaction of submitting author, Assistance for submitting author, # of use of dCollection documents, Proportion of successful download of digital documents.

<Tableau 2> Critères d'évaluation dCollection

## **Park, Hong Seok**

Email: [honspark@keris.or.kr](mailto:honspark@keris.or.kr)  
Téléphone: 82-2-2118-1331



Diplômé du Département de Chimie, Dankook University (1985)  
Master au Département des études en sciences des bibliothèques et de l'information, Emporia State University (1989)  
Doctorat de l'école des études en sciences des bibliothèques, de l'information et de la communication, Rutgers University (1995)  
Chercheur au Korea Research Information Center (1998-1999) Chercheur senior au Korea Education & Research Information Service (1999 - )

## **Jang, Keum Yeon**

Email: [jkylib@keris.or.kr](mailto:jkylib@keris.or.kr)  
Téléphone: 82-2-2118-1231



Diplômée du Département des en sciences des archives, des bibliothèques et de l'information, Pusan National University, (1995)  
Master et doctorat au Département des études en sciences des bibliothèques et de l'information, Sungkyunkwan University (2000 -)  
Assistant chercheur au Korea Education & Research Information Service (1999 - )