



# Developing Value-added Services Facilitating the Outreach of Institutional Repositories at CAS

Zhongming Zhu, Wei Liu, Linong Lu, Xiaona Yao, Wangqiang Zhang  
Lanzhou Branch of National Science Library, CAS



## Introduction

Promoting awareness and outreach of the institutional repository (IR) is vital to its success, and is contingent much on providing attractive value-added services that meet needs not currently being met by other tools. Chinese Academy of Sciences (CAS) therefore puts top priority on focusing on building services other than merely collections while it started to promote IR services within its confines. As a result, a wide range of value-added services were implemented and aimed to respond concerns what have perceived and identified as barriers preventing researchers from engaging in using repositories. They include collection building toolset, advanced browse and search facilities, usage statistics utility, knowledge mapping module, researcher knowledge profile, and authorship claim service.

## Collection Building Services

A variety of modules and tools were implemented to uphold recruitment of content in various formats through multiple channels. As depicted in Fig. 1, these services can be further divided into following three sub groups:

- Self-archiving deposit: it supports an academic author to deposit a copy of article in IR or a department repository. The process can be performed via repository's web interface or through a SWORD client (standalone client, web based client, or a word processor add-in). As for web interface based quick submission, it is document-type aware and highly simplified and optimized for just containing two steps or pages for completing one deposit with minimum time consuming. Usually, one submission is likely to be completed within 5 minutes. Still a detailed description and deposit workflow is reserved for use at authors' will.
- Mediated deposit/import: there are needs and requirements for IR to collect and integrate related content from various existent or legacy systems or applications to speed up the rate of content recruitment. The data import utilities can bulk import data in formats range from XML to Excel to Endnote to SCI to others into IR. Still, it can capture related data from research management system, ETDs databases, or other relevant applications within information environment of the institutes on the one hand, and on the other hand, it can ingest the institutes' publications data from Web of Science via its open web service interfaces. While some publishers are beginning to provide automatic article deposit service, a SWORD based automatic remote deposit service was implemented to support such using scenario. For example, CAS has now signed an agreement with BMC to cooperate to accept papers published in BMC publications to be automatically deposited into CAS IRs.
- Harvesting metadata/objects: by incorporating an OAI-PMH service provider (with an extension of harvesting content objects) into IR platform, it can harvest metadata records and content objects from related repositories via their OAI-PMH data provider interfaces, if applicable.

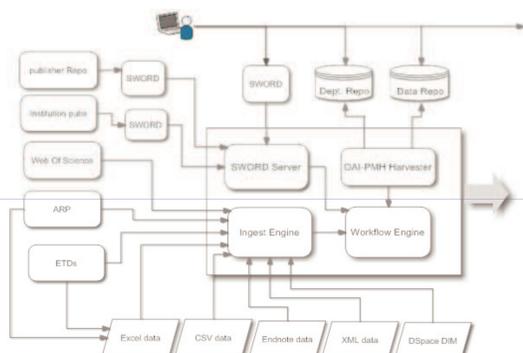


Fig. 1 Content recruitment framework for CAS OpenIR

## Advanced Browse and Search

In addition to traditional style browsing and searching services, fashionable faceted browse and search, auto-suggestion/auto-completion, and integrated neighborhood services bar were implemented to improve user's browse and search experiences.

Faceted browse and search supports refining the search result set based on facets like research community-/department, author, content type, and publication date. There are also experimental faceted navigation implementation based on automatically evaluated and assigned subject headings and DDC classifications in SKOS. Auto-suggestion and auto-completion are features popular used by search engines, they are well suited to be incorporated such a service into our IR to speed up human-computer interactions and present better user experiences.

Integrated neighborhood services bar was devised and implemented as a set of related handy services to be accompanied with each item reached during the course of browse and search. The bar will offer a set of following-on or neighborhood operations related to current item, which include bookmarking, recommending, viewing usage statistics, social bookmarking in Connotea, Cite-Ulike, Digg, etc, providing reference citation in normal style, starting an outside search with extracted information in external academic search systems such as Google Scholar, Scirus, CSDL Cross Search, etc. Thus, all services related to an item are gathered together and made available at the user's fingertips to use freely. Moreover, browse or search results can be exported to Endnote, Word, Excel/CSV, or other formatted data for reuse.

## Usage statistics service

It supports analyzing usage impact of IR in a customizable way. Each time, an analyzing process can be executed in combination of various parameters such as different content object levels (site, community, collection, item), different time interval levels (year, month, day, custom time period), different access styles (robot access, intranet access, repeated clicks), different countries or regions, etc. The result can be presented in variety of forms such as histograms, ranking lists, Excel spreadsheets, etc.

## Knowledge mapping service

The basic objective of this service is to provide researchers and research managers with tools to perform personalized knowledge asset auditing. It can support performing knowledge asset auditing in combination of a range of dimensions including content organization levels (institution, community, collection, and individual), content types, time spans, etc. The generated results are organized and visualized in various maps. Its highlights of flexible personalization are as follows:

- The overall set of audit conditions can be dynamically defined and configured. The repository manager can prescribe any statistically meaningful metadata element into audit conditions set, if needed.
- Each time of audit process can be customized according to the audit needs and requirements, based on prescribed audit conditions set.
- The audit results manifestations can be customized to display as knowledge inventory lists, histograms, line graphs, pie charts. In addition, the knowledge inventory lists can be exported as an Excel spreadsheet for later use.
- The columns of items appeared in a knowledge inventory list also can be adjusted as desired.

**Acknowledgements** This work is supported by the Knowledge Innovation Program of the Chinese Academy of Sciences and in part by the West Light Foundation of The Chinese Academy of Sciences.

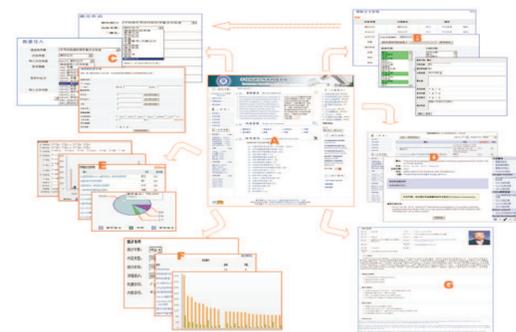
## Researcher knowledge profile

Researchers can use this service to create and maintain their research profiles based on IR. A researcher can set up and maintain his or her personal information including personal photos, blogs, educational and academic background, research interests and projects, etc. Based on these personal data, a personal research homepage and knowledge inventory list can be created automatically, and too, be exported as Excel spreadsheet for later use. An alias management module is also provided to help researchers manage their various forms of names, and combination with following authorship claim service to lay a foundation for clustering researchers works correctly and reliable.

## Authorship claim service

It works on a mechanism of combining alias control with authorship claim to link authors with their works correctly. An author or a researcher manages his or her various aliases based on researcher profile service. While an individual submission is completed and its authorship status is not clearly specified, the service will be automatically activated to check the author/creator values of the submission against alias database to find formally matched names as possible authors to send emails to them calling for claiming authorship. It is then easy and simple for authors received emails to identify whether or not an item belongs to him or her, and whether or not to follow links contained in the messages to make a claim confirmation.

## Exemplar Application Scenarios of the Value-added Services



(A) Homepage- faceted browse and search, recent sub-missions, usage ranking; (B) submission template defining and customization; (C) content recruitment: submission, bulk import, WOS data acquiring; (D) result page: faceted navigation, integrated neighborhood services bar, etc; (E) Personalized knowledge mapping; (F) usage statistics; (G) researcher profile.

Fig. 2 Exemplar pages of value-added services

## Facilitated Outreach Effects

Devising and implementing value-added services for IRs is a never-ending task to meet the ever-growing needs and requirements of researchers and other stakeholders. At present, promotion and outreach of IRs at CAS that are backed up by value-added services have seen promising early results. Now, over 90 institutes of CAS have set up or initiated their IRs. Total volume of items collected by all IRs has reached more than 28.9 million. While more than 10 of them have collected more than 5,000 items, most have accumulated more than a few thousand items. Of all items, 75 percent contain full-text content. The access usage of all IRs also shows a dramatic increase. Currently, it has more than 14 million total views and 2 million downloads in a short period of time.