EKT

ΕΘΝΙΚΟ ΚΕΝΤΡΟ
ΤΕΚΜΗΡΙΩΣΗΣ

NATIONAL
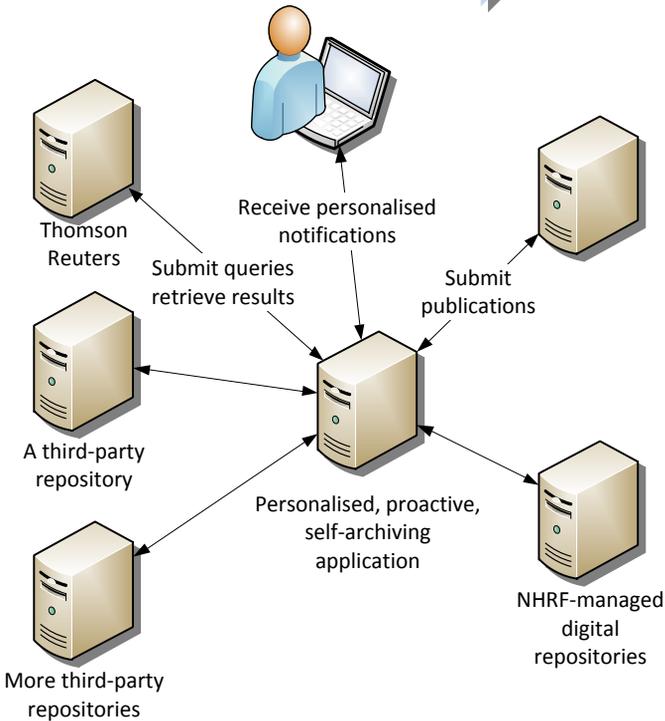DOCUMENTATION
CENTRE

Proactive personalized self-archiving

Nikolaos Konstantinou, Nikos Houssos *National Documentation Centre / NHRF*

I want to verify that my papers are correctly indexed

I want to keep track of my publication record

I want to ensure maximum reach-out of my ideas

Thomson Reuters

Receive personalised notifications

Submit queries retrieve results

Submit publications

A third-party repository

Personalised, proactive, self-archiving application

NHRF-managed digital repositories

More third-party repositories

The problem

Nowadays, institutional repositories are widespread. However, researchers do not invest time to archive their publications to repositories. This leads to the phenomenon of "empty archives" – institutional repositories containing only a modest fraction of the research output of the respective institution. One highly cited reason for this is the lack of concrete mandates - correctly. However, the lack of tools to make self-archiving easier and less costly in time for researchers is also a considerable prohibiting factor.
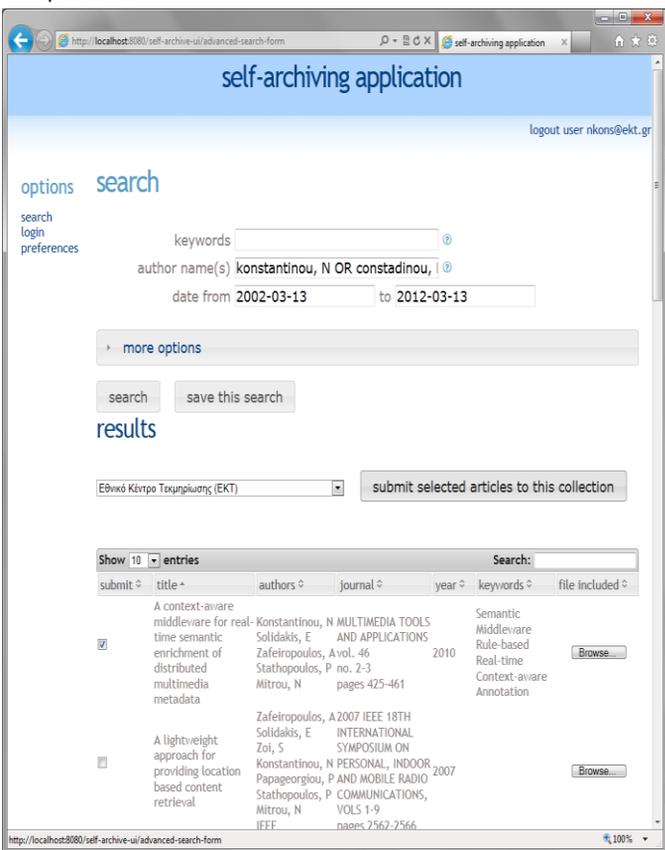
This contribution refers to the development of a tool for proactive, personalized self-archiving, aiming to increase adoption of self-archiving practices among researchers.

The solution / approach

In the proposed application self-archiving is

- *Proactive*: Users can opt in to receive periodical reminders about archiving their published work. If the system notices (e.g. in databases like WoK, Scopus, PubMed Central) that there are new articles published by a specific researcher in the last reference period (e.g. last month), an email notification is sent to the researcher, to encourage him or her to self-archive them – a link is provided for triggering the procedure. Metadata records are retrieved from online databases through queries asking for the author's papers. The results are presented to the researcher for review, potential corrections/additions to the metadata and for attaching the full-text. The processed results can be sent to the author's institutional repository with the click of a button through SWORD.
- *Personalized*: Users can create a profile that is linked with their user account in their institutional repository and the respective record in the repository's authority file for authors. They sign in to the system with the credentials of their repository account. They can create and save custom searches (e.g. "find my own publications") to online databases.

Challenges: Efficient detection of duplicates among input data sources and between data source and target repository. Deletion/merging of records as soon as duplicates are identified. Appropriate mapping / processing of data source records before submitting to the repository. Make user interface friendly so that procedure is quick for end users.

Use case scenario

1. A researcher has already an account and submission privileges at an institutional repository (e.g. the HELIOS repository of NHRF, helios-eie.ekt.gr)
2. He/she logs in the self-archiving application using the credentials of his/her repository account and performs searches in the third-party data sources for publications of interest (e.g. own publications).
3. Optionally, the user can save his/her search and opt to receive periodical notifications with the search results.
4. The results can be submitted to the institutional repository, including user-uploaded files (metadata and manuscripts), after review and processing of information by the researcher.

Implementation details

- The system has been tested currently with a single data source, the Web Services interface offered by Thomson Reuters.
- The submission interface is implemented via SWORD. This enables one-click submission to any SWORD-compliant repository of publication metadata, as well as inclusion of the manuscript itself.
- Java-based technology stack, using Spring/Spring MVC, Javamail, Apache CXF, Velocity.

Ongoing work

- Efficient mechanisms for identifying duplicate records and records already in the repository and perform the appropriate merging without losing information.
- Exploiting information from Sherpa/RoMEO for helping researchers to upload the right version of the full-text
- Integrate more data sources in the system PubMeb Central, Scopus).