Easy II, an Electronic Archiving SYstem

Main Author Henk van den Berg

DANS-KNAW, Data Archiving and Networked Services
Po Box 93067, 2509 AB The Hague, Netherlands
EMail: henk.van.den.berg@dans.knaw.nl

ABSTRACT

DANS – Data Archiving and Networked Services – was established in 2005 and is an institute under the auspices of the Royal Netherlands Academy of Arts and Sciences (KNAW) and supported by the Netherlands Organisation for Scientific Research (NWO). The main objective of DANS is permanent preservation of, and enabling access to scientific research data.

With the Easy II project, which is an acronym for Electronic Archiving System, DANS set out on a trajectory to replace the existing Easy-application with a modular, multi-functional architecture and system that could serve as a basis to realize DANS' goals in respect to storage of scientific research data and their dissemination. The field of digital data preservation and access is not new, but still its impact is gathering momentum. We are confronted with new technologies and services and demands from the scientific community gain in complexity and diversity. In order to cope with such needs in a rapidly changing environment, from the outset it was clear that we should aim at a flexible architecture, enabling rapid development and deployment of diverse dedicated applications and services centred on a repository.

Keeping with this preliminary objective, direct requirements of the central web-application that serves as the main entry-point to the system constituted self-archiving of scientific research data, multiple, discipline-dependent metadata formats and (web) forms based on Dublin Core, curating of data and metadata by archivists supported by flexible workflow schemes, search and facetted browsing of metadata and access to data files based on user-defined access restrictions. Besides this public web-application a series of tools should enable diverse administrative purposes such as curating and batch-like ingest of scientific collections based elsewhere and the ingest of results of retro-archiving projects.

The Easy II project delivered a series of reusable, repository centred modules and, based on these, two web front-ends: the general purpose self-archiving system Easy and the dedicated Digital Collaboratory for Cultural Dendrochronology (DCCD) system. The reusable modules and components encompass, besides middle-ware that enable communication with a repository from within a java-environment, visualisation components, user-management, authentication and authorization frameworks. The reusable modules are now under the provisionary name of dans-commons. As implementations underneath these modules we should mention Fedora, as repository system, and Solr, for searching and facetted browsing.

Needless to say we are open for collaboration with other organizations and encourage the reuse of our software. The paper will concentrate on the architecture of Easy and will explain some of its details. Briefly it will memorize some of the organizational and methodical aspects of software development in general. Two of the things we learned: code is condensed knowledge, architecture is a process.