

Sharing Grey literature by using OA-x¹

Elly Dijk

Netherlands Institute for Scientific Information Services, NIWI-KNAW

Abstract

As part of the Dutch DARE (Digital Academic REpositories) programme, NIWI-KNAW is participating in various projects to enlarge open access to Dutch scientific output (including Grey Literature). The Open Source web technology that will be used for harvesting is based on i-Tor, Tools and technology for Open Repositories, developed by NIWI (the Netherlands Institute for Scientific Information Services).

The subject of this lecture is the initiative that NIWI-KNAW has taken for developing the OA-x protocol, a modular extension of the OAI protocol, OAI-PMH (Protocol for Metadata Harvesting). We shall discuss the advantages of OA-x and for what projects it will be used.

In the protocols of the Open Archives Initiative (OAI) currently in use, information is shared by providing metadata of digital files (data providing) that can be read in by someone else (data harvesting). A URL is used to refer to an object in an external site (often a repository).

In certain cases, one needs to go further than just sharing metadata. Certainly in the case of so called collaboratories, it should be possible to transfer the objects themselves from an external website (or repository) to one's own site. And conversely, it should be possible to upload objects to an external site. Even if only the browsing of objects is required, it is necessary to get to the original document in order to be able to index it.

The OA-x project has been set up to enable researchers and administrators of (digital) archives to be able to unlock, edit, supplement, combine and archive metadata and data (objects) in digital repositories. A protocol for harvesting and uploading objects has been developed in this project. There are also several implementations available: OA-x within a CMS, OA-x as extendable OAI data and server provider, and OA-x as repository filler. We have opted for similar names of verbs as are used in OAI-PMH.

The advantages for authors and administrators of (digital) archives are great. It is possible to place articles or other publications on one's own website as full text and easily export them to a repository such as an institute repository. It is also possible to use OA-x to upload publications to electronic journals (e.g. Studies in Mycology) or to a central address where a grey publication will be produced.

With the aid of the i-Tor technology, it was already possible to index PDFs on one's own website as full text and make them searchable via Google. Thanks to OA-x, it is now possible to index PDFs (or other text files) on external sites as full text too.

In collaboratories, it is not uncommon to use collections of images that are split over various sites. With an OA-x implementation, it is possible to make a collection of thumbnails of the images in these distributed collections in one place (as if the images were collected on one site only).

An example of a collaboratory is E-laborate, a virtual joint venture in the alpha and gamma disciplines. OA-x is used by E-laborate to upload datasets to subject-based repositories.

OA-x can also be used to make a (national) electronic depot. Objects can be sent to such a depot or archive from institutional repositories. OA-x makes it possible to not only send the object but also multiple datasets. It is even more important that information about the technical data (e.g. in what version of PDF the object was created) can be sent along with it. These data are of essential importance to a depot because they can be used to see whether an object (and the format of the object) has remained unchanged.

Introduction

As part of the Dutch DARE (Digital Academic Repositories)² programme, the Netherlands Institute for Scientific Information Services (NIWI)³ – an institute of the Royal Netherlands Academy for Arts and Sciences (KNAW)⁴ – is participating in various projects to enhance open access to Dutch scientific output, including "grey" literature. DARE is a joint initiative by the Dutch universities to make their research data accessible in digital form. As well as KNAW-NIWI, its other participants include the Netherlands