

Pandektis: Implementing a repository of greek historical and cultural material with DSpace

Abstract:

The present paper briefly describes the scope and implementation of ‘Pandektis’, an open access repository containing resources of greek history and culture (<http://pandektis.ekt.gr>) that has been developed by the Hellenics National Documentation Center of the National Hellenic Research Institute (<http://www.ekt.gr>). ‘Pandektis’ has been developed using the open repository platform DSpace but has been configured and customized based on the requirements of each hosted digital collection. In the following, after describing the scope of the project, we report the main enhancements or customizations that we performed on the DSpace platform to accommodate the requirements of this specific digital repository implementation. The main issues that have been addressed include context-sensitive presentation and browsing of material, support of custom metadata, linking among digital artifacts and support of range queries in advanced search.

Introduction:

The goal of the “Pandektis” project is to create an open access repository with specific historical and cultural collections containing research output from three humanistic studies institutes of the National Hellenic Research Foundation (<http://www.eie.gr>), namely the Institute of Neohellenic Research, the Institute of Byzantine Research and the Institute of Greek and Roman Antiquity. The repository consists of 11 collections with roughly 20000 images and 45000 metadata records and includes diverse material such as epigraphs, byzantine documents from Mount Athos, heraldic monuments, maps and paintings, among others.

Design and contribution:

Due to the nature of the collections, “Pandektis” is characterized from heterogeneity among its data that requires different ways of representation and visualization of an item in a “detailed view” page and of item lists in a browse or search results page. Another important characteristic of “Pandektis” is the existence of relationships among items of different sub-collections and communities that need to be reflected in the user interface to facilitate navigation of the material. The above requirements led us to certain customizations and extensions of the DSpace repository platform. These extensions are based on the aforementioned requirements and are summarized in the following paragraphs.

Metadata Representation

Dublic Core metadata element set is a standardized vocabulary for describing a variety of digital documents. However, in our case the diversity among “Pandektis”

collections requires an extension of the basic schema. So in order to describe items such as portraits, heralds, epigraphs and maps and make our information discovery system more intelligent, we defined an application profile which consists of data elements drawn from the Dublin Core namespace combined with elements from Qualified Dublin Core namespace and custom elements from our schema.

Advanced Search

DSpace utilizes Lucene which is a high performance search engine library. However, DSpace does not exploit the full range of Lucene functionality such as range queries and proximity queries. In Pandektis there was a strong requirement by the researchers to support date ranges as search criteria (e.g., to enable queries like search all collections for documents dated from 1530 to 1750). This has been achieved by modifying the DSpace code used for forwarding query strings to the Lucene engine as well as the advanced search form, which in our version uses client-side javascript logic to allow range query specification.

Layout Customization

Each collection in “Pandektis” had set different display requirements which can be summarized to the following:

1. Branding content: Preserving the institutes’ brand in the repository. The separate institute identity behind the sources is important and needs to be preserved in the presentation of the repository (e.g., using a custom header and different colours for each collection).
2. Visualizing items: There are items in separate collections which need to be displayed differently, based on criteria like the existence of images and their size and resolution, their copyright limitations and the existence of geospatial metadata. Thus, in paintings or epigraphs a zoom functionality is desirable so that the user is able to view significant details of the images. Furthermore, items containing geospatial metadata can be plotted and browsed on a map.

Linking Digital Assets

For some of the Pandektis collections, there are explicit relationships among two or more metadata records belonging to the same or to different collections. For example, there are certain compound monuments that consist of individual smaller ones (intra-collection links). Another case appears in the travel literature collections that includes records concerning illustrations and separate records (in a different sub-collection) for travel literature books, some of which contain illustrations (inter-collection linking). To provide a better use experience we implemented a mechanism that creates active links to the related records of an item (e.g., each book page has direct links to all illustrations in the book). Furthermore, some items refer to external resources and thus we provided the capability of external linking.

Conclusion

The “Pandektis” project is an attempt to publish different historical and cultural collections from various time periods of Greek history using DSpace as a common

technological platform. To adapt to the special needs of the different collections but also achieve a unified way of presentation and search, several enhancements and customizations of the DSpace platform have been performed, mainly concerning adaptive presentation and browsing of material, metadata representation, linking between digital artifacts and advanced search.