

The Arts and Humanities Data Service: the experience of building a research support infrastructure in the UK

Sheila Anderson and Martin Wynne

Arts and Humanities Data Service, UK

1. The AHDS in the UK research support infrastructure

The Arts and Humanities Data Service (AHDS) is a UK national service aiding the discovery, creation and preservation of digital resources in and for research, teaching and learning in the arts and humanities. Currently, the AHDS covers five subject areas:

- Archaeology,
- History,
- Literature, Languages & Linguistics
- Performing Arts,
- Visual Arts.

The AHDS is organised via an Executive at King's College London and five AHDS Centres, covering the five subject areas above, and hosted by various Higher Education Institutions. The AHDS is funded by the Joint Information Systems Committee (JISC) and the Arts and Humanities Research Council (AHRC).

The key functions of the AHDS are advice to funding bodies, advice to resource creators, and the archiving, distribution and preservation of resources.

1.1. Services to funders

One of the key functions of the AHDS is to provide an advisory service to funding bodies such as the AHRC, JISC and the British Academy. The AHDS gives advice to applicants on the technical aspects of their funding proposals, carries out assessments of the proposals for the funding body, offers advice to grant holders, offers advice on strategic and policy issues relating to electronic resources, and offers an archiving service for project outputs. These services are particularly well integrated into the processes of the funding schemes of the AHRC. Recipients of AHRC grants who are creating electronic resources are normally expected to deposit the resources with the AHDS within three months of the end of the project.

The innovative work and the organisational structure of the AHDS have already been taken as a model for the creation of an Economic and Social Data Service (ESDS), and have informed the work of many other new initiatives and structures in the UK.

1.2. Advice to resource creators

AHDS staff offer a free initial advisory service to UK academics who are creating electronic resources. In

addition to dealing with directly with enquiries, various published materials cover data creation, access and delivery, and data deposit and preservation.

The AHDS has created a series of Guides to Good Practice focusing on the practical steps necessary to make a successful digital resource. Some of these are subject-based, while others cover cross-disciplinary topics. The latest title is *Developing Linguistic Corpora: a guide to good practice*. Case Studies review various projects which are creating or completing digital resources in the arts and humanities. Various issues are covered, such as funding, preservation, and using resources in teaching. Information Papers are shorter publications relating to specific technical questions. They are not focussed on any particular subject area but rather deal with the various aspects of a digitisation project (e.g. project management, metadata, XML editors).

The AHDS offers various workshops and events which give further advice to those interested in creating, maintaining and using digital archives.

2. The AHDS and language resources

The centre responsible for language resources is AHDS Literature, Languages & Linguistics, which is hosted by the Oxford Text Archive (OTA) in the University of Oxford. The OTA has been in operation for 30 years and has been the centre responsible for literary and linguistic subject areas since the foundation of the AHDS in 1996. AHDS Literature, Languages & Linguistics benefits from its location in Oxford University Computing Services in close proximity to projects and services including the Text Encoding Initiative, the British National Corpus, the Oxford e-Science Centre and the JISC Open Source Advisory Service (OSSWatch).

The language resource holdings of the AHDS include many historical and literary texts, language corpora, lexical data, and other types of literary and linguistic dataset. As with all AHDS resources, they are catalogued using the AHDS Common Metadata Format, and can be discovered via subject-specific portals and via the AHDS cross-search catalogue. AHDS Literature, Languages & Linguistics also shares discovery metadata with the Open Language Archives Community (OLAC).

AHDS Literature, Languages & Linguistics plays a central role in the UK in promoting good practice in the creation of language resources, in promoting the use of language resources in research and

learning and teaching, and archiving language resources and in developing new and improved ways to deliver them to the user.

3. AHDS and the European Infrastructure

Just as astronomers require a virtual observatory to study the stars and other distant objects in the galaxy, researchers in the humanities need a digital infrastructure to get access to and to study the sources that are until now hidden and often locked away in cultural heritage institutions. Only a fraction of the analogue sources in archives, libraries and museums is as yet available in a digital form. Of course, more and more sources are being digitized, but the permanent and open access to the information they contain is only yet beginning. It is therefore not surprising that the recent US Cyber Infrastructure for the Arts and Humanities report proposed as its grand vision 'access to all surviving humanities and cultural heritage information across all of time and space'.

Since the publication of that report, the National Science Foundation in the US has produced a draft strategy aimed at implementing many of the recommendations contained in the report. The challenge for Europe is to ensure that the development of a research infrastructure for cultural heritage and the humanities that can match, or supersede, that of the US. The infrastructure that is needed for the humanities is indeed very much comparable to infrastructures for the natural sciences such as the virtual observatory, and the kind of organisation and grid-based techniques that are required also show a surprising degree of similarity.

The AHDS is working with partners across Europe to identify and develop the key elements and activities for a Research Infrastructure (RI) that take steps towards achieving this grand vision for European humanities and cultural heritage information, and would provide an infrastructure that eventually could support access to all surviving humanities and cultural heritage information for Europe. Such a Research Infrastructure would:

1. Provide a coordinated infrastructure across Europe that would act as a catalyst to bring together the best efforts of national initiatives, organisations and individuals in order to provide upgraded and enhanced European wide actions, initiatives and services that could not be provided at local or national level.

2. Provide a coordinated infrastructure that would act as both a catalyst and support for the development of national services and digitisation programmes aimed particularly at those European countries without such services and programmes.

3. Provide a coordinated infrastructure that would act as a catalyst to bring together the different sectors involved in cultural heritage and humanities information management and access – education, memory and cultural heritage institutions and organisations, and the commercial sector – in order that they might work together for the benefit of both themselves and the research communities across Europe.

4. Provide a coordinated infrastructure that would act as a catalyst for the enhancement and promotion of digital scholarship in the humanities and arts across Europe, including facilitating cross-disciplinary research and the sharing of content, tools and methods across communities of practice and discipline domains. This would also ensure that the arts and humanities did not work in isolation but took note of developments across the social, physical and medical sciences.