This is the published version:


Available from Deakin Research Online:

http://hdl.handle.net/10536/DRO/DU:30074160

Reproduced with the kind permission of the copyright owner.

Copyright: 2014, The Authors
Linking and sharing data in the humanities and creative arts: building the HuNI Virtual Laboratory

Toby Burrows
Manager, eResearch Support
University of Western Australia
toby.burrows@uwa.edu.au

Deb Verhoeven
Professor of Media and Communication
Deakin University
deb.verhoeven@deakin.edu.au

Abstract:
The Humanities Networked Infrastructure (HuNI) is one of the national Virtual Laboratories that are being developed as part of the Australian government's National e-Research Collaboration Tools and Resources (NeCTAR) programme. This paper examines the methodologies and technical architecture being deployed by HuNI to link and share Australian data in the humanities and creative arts.
Introduction

The Humanities Networked Infrastructure (HuNI) is one of the national "Virtual Laboratories" which are being developed as part of the Australian government’s NeCTAR (National e-Research Collaboration Tools and Resources) programme. NeCTAR aims to integrate existing capabilities (tools, data and resources), support data-centred workflows, and build virtual communities to address well-defined research problems. This is a particularly challenging set of technical requirements and problems for the humanities and creative arts, which cover an extensive range of different disciplines and are characterised by complex and highly heterogeneous collections of data (Burrows 2011). User requirements and use cases are also very varied and complex.

HuNI is being developed by a consortium of thirteen Australian institutions, led by Deakin University. It is bringing together data from thirty different Australian datasets, which have been developed by academic research groups and collecting institutions (libraries, archives, museums and galleries) across a range of disciplines in the humanities and creative arts. These datasets include Design and Art Australia Online, the Australian Dictionary of Biography, AustLit, AusStage, the Dictionary of Sydney, and the PARADISEC linguistics archive (see Appendix 1 for a full list). These datasets contain more than 2 million authoritative records, capturing the people, places, objects and events that make up the country’s rich heritage.

HuNI is ingesting data from all these different Australian data providers, mapping the data to an overall data model, and converting the data for inclusion in an aggregated store. HuNI is also assembling and adapting software tools for using and working with the aggregated HuNI data. Two existing tools are being extended to interface with HuNI: L ORE (Literature Object Reuse and Exchange), developed at the University of Queensland as part of the AustLit service, and Heurist, developed at the University of Sydney.

Fundamental to HuNI’s architecture was the decision to build a central aggregate, rather than designing the Virtual Laboratory functionality (e.g., federated searching, browsing and annotating) to work with the individual data sets in a distributed way. A central aggregate adds significant value to the disparate data sources by maximising the links between them, and by putting them into a much broader interdisciplinary context. It also enables researchers to work with data from a variety of different sources in a much more effective way and on a much larger interdisciplinary scale.

HuNI is part of the rapidly growing global Digital Humanities initiative, which is producing many innovative applications and services aimed at expanding the use of digital technologies in humanities research. In Australia, this saw the formation of the Australasian Association for Digital Humanities (aadH) in March 2011, its formal incorporation in March 2012 and its acceptance into the international Alliance of Digital Humanities Organisations (ADHO). There is a significant overlap in membership between the AADH Executive Committee and the HuNI Steering Committee.