

Semantics, social linking and humanities data in HuNI

Deb Verhoeven¹, Toby Burrows², Alwyn Davidson³

¹Deakin University, Melbourne, Australia, deb.verhoeven@deakin.edu.au

²University of Western Australia, Perth, Australia, toby.burrows@uwa.edu.au

³Deakin University, Melbourne, Australia, alwyn.davidson@deakin.edu.au

INTRODUCTION

HuNI is a virtual laboratory application of great value to anyone interested in understanding Australia's rich cultural heritage. Recent attendees of eResearch Australasia may be familiar with the description and ambition of the HuNI project. This presentation however, will focus on the transition of HuNI from development to a fully operating platform and highlight the challenges of ensuring uptake by researchers and contributing datasets. This paper will analyse, describe and evaluate the full production release of HuNI. It will include a demonstration of the full functionality of the service, and will report on its uptake by researchers and the wider community. We will also discuss the lessons learned from this large-scale project over its two-year lifespan, and the measures taken to ensure the sustainability of the HuNI service into the future.

HUNI

HuNI (Humanities Networked Infrastructure) is a major new digital service for humanities researchers. Developed with funding from the NeCTAR (National eResearch Collaborative Tools and Resources) programme, it aggregates data from 30 different cultural datasets from a variety of disciplines, makes them available for external re-use and provides a set of tools for researchers to work with user-defined collections of data. It is also the single largest aggregation of humanities data in Australia – a new national data service which is of cultural significance in its own right, and accessible to all. At its core, HuNI engages the broad community by enabling them to define the relationships that constitute their field of research through “social linking”.

SOCIAL LINKING

The paper will focus, in particular, on the ways in which HuNI is changing the nature of humanities research in the areas of data sharing, collaboration, community involvement, and the creation of socially linked data. “Social linking” of data is one of the key features of HuNI in which researchers are able to make assertions about relationships between entities represented in the aggregated data. For example by searching the data aggregate, researchers can identify two entities in their result set which are related in some way, and can add a link between the two records and define the nature of the relationship. The linking statement may be drawn from an existing vocabulary of relationships, or researchers can nominate a new relationship using free text. Links can be either symmetric (the same in both directions such as “x is the same as y”) or asymmetric (different depending on the direction of the assertion such as “x is the daughter of y”). The virtual laboratory also allows a researcher to assert that two entities are *not* related, in recognition that this kind of statement is also a key characteristic of humanities research.

To help visualize these social links, each “socially linked” entity has its own network graph, showing up to six degrees of separation, resulting in a growing network of dynamic connections, or the “networked effect”. These “social linking” assertions are visible in the HuNI data aggregate. HuNI users can annotate these links to explain the basis for their assertions. These user-produced links may also appear in virtual collections assembled and “published” by individual users of the HuNI virtual laboratory for other users to see and re-use. Crucially, the provenance of all these “social linking” statements is also captured, enabling subsequent researchers to see who made each assertion (and to therefore make an assessment of the proposed links based on the expertise of the researcher).

Humanities research not only involves making connections between entities; it also involves assessing any changes in cultural flows and network relationships through time. Consequently HuNI is an aggregate with a sense of its own history, in which researchers can trace how records have changed over time. So each HuNI record is time-stamped, meaning that researchers will always see the current view of a record, with its related records and assertions, but will also have the option to view how the record has changed since it was first harvested. The provenance information for each record, together with any curated assertions, is captured, so that researchers can see when the records were harvested and by whom. A link to the originating data record at source is also provided in the user interface.

“Social linking” is a crucial feature of the HuNI virtual laboratory. Instead of relying on a pre-coordinated mapping to a detailed ontology, we are relying on researchers and community users to establish most of the connections within the heterogeneous data aggregate. This enables HuNI to capture the different disciplinary perspectives of users, rather than trying to coerce the data into a single normative framework. It also allows HuNI to encourage its users to share their knowledge and research findings in the form of specific assertions, and to discuss these statements with each other. In the paper, we will report on the way in which HuNI has used the Virtual Laboratory environment to enable a new approach to data sharing in the humanities and we will assess its early uptake and impact.

ABOUT THE AUTHORS

Deb Verhoeven

Professor Deb Verhoeven is Chair and Professor of Media and Communication at Deakin University, Deputy Director of the Centre for Memory, Imagination and Invention and a Chief Investigator in the ARC Centre of Excellence for Creative Industries & Innovation. She is the Project Director of Humanities Networked Infrastructure (HuNI), a two-year project funded by NeCTAR (National eResearch Collaboration Tools and Resources). She served as inaugural Deputy Chair of the National Film and Sound Archive of Australia (2008-2011) and as CEO of the Australian Film Institute (2000-2002). In 2013 Professor Verhoeven initiated *Research My World*, a collaboration between Deakin University and the crowdfunding platform [pozible.com](#) to pilot the micro-financing of university research. Deb is a leading proponent of the Digital Humanities in Australia. Her recent research has addressed the vast amounts of newly available ‘cultural data’ that has enabled unprecedented computational analysis in the humanities.

Toby Burrows

Dr Toby Burrows manages the eResearch Support and Digital Developments Unit at the University of Western Australia. A specialist in digital humanities research, he is currently undertaking a Marie Curie International Incoming Fellowship in the Department of Digital Humanities at King's College London between 2014 and 2016. His main research interests are digital humanities and medieval manuscript studies. He has been an invited participant in workshops funded by the European Science Foundation and COST, and has held visiting fellowships at the Vrije Universiteit Amsterdam, Churchill College Cambridge and University College London. He has been a Chief Investigator on a range of Australian Research Council projects under its Discovery, Linkage, Linkage Infrastructure and Research Networks schemes.

Alwyn Davidson

Dr Alwyn Davidson is an early career researcher specialising in the digital humanities and geovisualisation. Alwyn's research interests include the visualisation of humanities data, the application of GIS and spatial techniques, the analysis of cultural datasets, and working within an interdisciplinary environment. Alwyn is currently a research assistant at Deakin University.