

The creation, curation and dissemination of humanities digital resources: The Dutch Dialect Database

Douwe Zeldenrust

Meertens Institute (Royal Netherlands Academy of Arts and Sciences)

The Netherlands

douwe.zeldenrust@meertens.knaw.nl

Introduction

At the past ESSHC conferences, digital and computational methods, such as the use of Big Data and GIS (Geographic Information Systems), play an increasing role. Within this domain the creation, curation and dissemination of humanities digital resources is an essential part. This paper focuses on handling digital data and on connecting the data with European research infrastructures.

Use case: the Dutch Dialect Database

A use case will be presented: the Dutch Dialect Database (NDB) of the Meertens Institute, Royal Netherlands Academy of Arts and Sciences (Oostendorp). The NDB is web based and contains sound recordings of Dutch dialects (more than 1200 hours), typescripts and information about the speakers (in total over 11.000 scans). While constructed for dialect researchers, the NDB offers a rich resource for other interested parties, such as historians and ethnologist. The collection gives a unique representation of the Netherlands that no longer exists. The recordings are about poor living conditions in rural areas, welfare during the Depression, the

Second World War, local customs et cetera. After the initial launch of the NDB, extra datasets have been added, such as Dutch spoken in the USA and Dutch spoken in France. These datasets have also been made CLARIN (Common Language Resources and Technology Infrastructure) compatible and are available via CLARIN (Zeldenrust).

Recently, the web interface of the NDB has been renewed significantly. The latest design principles for web based spoken word archives have been integrated in the NDB. These principals follow the ideas of Annie Murray and Jared Wiercinski. They state that the creation of new knowledge in the humanities 'depends not only on better understanding the role of sound in the work of humanities scholars, but also incorporating this knowledge into the design of sound archives' (Murray). The NDB offers, amongst other search possibilities, the visual interface of the 'speaking map'. The speaking map is a representation of the data using the geographic locations.

Conclusion

To conclude with, this paper will reflect on the creation, curation and dissemination of humanities digital resources. New datasets and tools will open new research possibilities. For instance, adding the collection Dutch Spoken in the USA to the NDB will open the

way for pattern recognition, using the combination of sound and written resources (Sijs).

References

Murray and Jared Wiercinski (2014). "A Design Methodology for Web-based Sound Archives". *Digital Humanities Quarterly*, Volume 8 Number 2.

Oostendorp, M. van (2014). "Phonological and phonetic databases at the Meertens Institute." *The Oxford Handbook of Corpus Phonology*. Eds. J. Durand & G. Kristoffersen. Oxford: OUP. 546-551.

Sijs, N. van der. (2013). "Op de schouders van onze voorgangers. Jo Daan en het Amerikaans-Nederlands". *Nieuw Letterkundig Magazijn* 31.1: 30-34.

Zeldenrust, D.A. (2014). "Access to data: the Soundbites collection of the Meertens Institute and a flexible approach to the curation and dissemination of humanities digital resources". Digital Humanities Benelux Conference (The Hague).

Websites

www.clarin.eu (Accessed March 2, 2015).

www.meertens.knaw.nl/ndb (Accessed March 2, 2015).