

CREATING INTERFACES

Alfalab and practical approaches for connecting researchers, developers and infrastructures

Douwe Zeldenrust

*Meertens Institute, Royal Netherlands Academy of Arts and Sciences
Joan Muyskenweg 25, 1096 CJ Amsterdam, The Netherlands
douwe.zeldenrust@meertens.knaw.nl*

Marc Kemps-Snijders

*Meertens Institute, Royal Netherlands Academy of Arts and Sciences
Joan Muyskenweg 25, 1096 CJ Amsterdam, The Netherlands
marc.kemps.snijders@meertens.knaw.nl*

Rob Zeeman

*Meertens Institute, Royal Netherlands Academy of Arts and Sciences
Joan Muyskenweg 25, 1096 CJ Amsterdam, The Netherlands
rob.zeeman@meertens.knaw.nl*

Joris van Zundert

*Huygens ING, Royal Netherlands Academy of Arts and Sciences
Prins Willem-Alexanderhof 5, 2595 BE, Den Haag, The Netherlands
joris.van.zundert@huygensinstituut.knaw.nl*

Anne Beaulieu

*Virtual Knowledge Studio, Royal Netherlands Academy of Arts and Sciences
Cruquiusweg 31, 1019 AT Amsterdam, The Netherlands
anne.beaulieu@vks.knaw.nl*

Keywords: Alfalab, Royal Netherlands Academy of Arts and Sciences, interfaces, infrastructures, virtual research environments, geographic information system, CLARIN, Open Annotations.

Abstract: This paper presents project 'Alfalab'. Alfalab is a collaborative framework project of the Royal Netherlands Academy of Arts and Sciences (KNAW) and was launched in 2009. The paper explores interfaces for interaction between research questions, tools and data, and infrastructures. It explains how Alfalab supports the very heterogeneous heuristics and data sources used in the humanities research community, while also creating solutions that are integrated into e-science infrastructure initiatives (e.g. CLARIN). One case of interface development, the GISLab virtual research environment, is described in greater detail.

1 INTRODUCTION

The concept of interface is pivotal to project Alfalab's approach to researcher-technology interaction in the humanities. In this paper,

interfaces are defined as the key area of interaction between research question, tool development and data infrastructures. The paper describes one case of interface development within the framework of Alfalab, which was launched in 2009 as a cross-

institutional collaboration (Zundert et al. 2009), with a budget just under one and half million Euros. The six institutes that are involved are: Data Archiving and Networked Services, The Fryske Akademy, Huygens ING, The Meertens Institute, the International Institute of Social History, and the Virtual Knowledge Studio. These institutes are part of the Royal Netherlands Academy of Arts and Sciences (KNAW). The paper also demonstrates how continuous application of the concept of ‘interfacing’, in the broader sense, is fundamental to the successful connection of researchers, software developers and infrastructure providers.

2 PROJECT ALFALAB

Recent investments in e-infrastructures for the humanities include the second European Strategy Forum for Research Infrastructures (ESFRI) roadmap, which earmarked five projects (CLARIN, CESSDA, DARIAH, ESS and SHARE) from social sciences and humanities. These were followed by related national initiatives, such as those around CLARIN. A recent EU report concerning scientific data notes that, while these projects provide an essential layer in a collaborative data infrastructure, appropriate attention should be given to connecting users and data generators (European 2010). According to the report, this can be done via user functionalities, data capture and transfer, and virtual research environments (VREs).

Alfalab has come to similar conclusions through research on the ground, and regards that successful infrastructures must be (1) adopted and adapted by researchers and their fields (Ribes and Lee 2010), and (2) open to the particular research approaches and tools of humanities researchers (Wouters and Beaulieu 2006). Alfalab has therefore focused on setting up an interface that enables these essential mutual adjustments. This interface can deal with a wide range of heterogeneous data, highly specific workflows and diverging user demands.

Alfalab is currently developing three VREs through interdisciplinary collaboration between humanities researchers, computer scientists and science and technology studies experts. The GISLab is a tool chain that combines geospatial research data and analysis. TextLab is a web based online research environment pertaining for (collaborative) analysis for digital textual sources. And LifeLab is a census data- based VRE for historical research.

This interdisciplinarity is made explicit in the project through interactions and interventions (Antonijevic and Beaulieu 2009). It is used as a resource for developing shared understandings of

research agendas, of data and tool coupling, and of researchers’ needs when working in digital settings (Zundert et al. 2010). The combination of disciplines and the creation of a number of pilots and demonstrators in Alfalab have therefore enabled a critical mass of research in digital humanities to come together. This has been the occasion to raise a range of issues in the creation of interfaces, from technical feasibility and standardization to epistemology.

Furthermore, building on the lessons learned from other digital humanities projects (Dormans and Kok 2010; Van Dalen-Oskam 2009), Alfalab seeks to share a number of basic tasks rather than reinvent the wheel. In this respect, CLARIN-NL and DARIAH act as providers of knowledge on standard solutions for infrastructure-related activities such as metadata and metadata publication, persistent identifiers, identity federations and semantic interoperability.

The success of Alfalab’s approach to VREs can be traced back to three aspects of our focus on interface: first, to the focus on the value of interaction between the participants from different fields and the subsequent mutual shaping of Alfalab expertise; secondly to the explicit feedback cycles built in into the development process; and third to the continuous stimulation of critical reflection on the ‘processes of translation’ occurring during the project between researcher, developer and infrastructure-provider. As such, the outcomes of Alfalab will equally be the mechanisms for ‘interfacing’ set up in and around the project, as well as the specific VREs. The continuation of Alfalab in the new e-Humanities group of the KNAW is being investigated.

3 CREATING THE INTERFACES: THE GISLAB VRE

Interfaces play an important role in one of the first deliverables of Alfalab, the virtual environment ‘GISLab’. The GISLab is a web-based geographic information system where researchers can upload their datasets (or use existing ones) and enrich these with further data and tools. Processed datasets can be shared for cross-data research and client-side analysis. The first version of GISLab contains three basic components: a data repository that contains user data, metadata, datasets and plug-ins; a georeferencing tool so users can add geographical coordinates to their own maps or images; and a geo-annotation tool to annotate datasets.

Before and during the design process, researchers were consulted via expert meetings and

interviews. Input from software developers was obtained during technical meetings, interviews with GIS specialists and from the open source domain on the Internet. Five use cases were defined, each coupled to an available dataset: Dutch Rural Microtoponyms (Meertens Institute), Hydronymia Neerlandiae (Gerald van Berkel), Microtoponyms Schagen, Cinema Context (University of Amsterdam) and Early Dutch Cinema (EYE Film Institute). To support the feedback cycle, the use cases and corresponding datasets were selected in order to maximize diversity, while also being specific enough to meet the demands of an individual researcher. Three of these use cases are based on onomastic datasets (Zeldenrust 2005). The other two use historical film screening data. Across all cases, new knowledge arising from use of the VRE will be shared through Open Annotations.

The main challenge was to realize an optimal trade-off between the ability to deal with a wide range of different data, workflows and users on the one hand, and, on the other, to comply with standards required by e-humanities infrastructures such as CLARIN. To meet both demands of standardization and flexibility, GISLab opted for a highly modular system design. With this approach, we were able to restrict infrastructural uncertainties to single, replaceable modules, resulting in highly malleable tools. Tools can be adapted for users' own specific research context. Even at the database level, functions are implemented independently. User data, metadata, application data (indexes, bookmarks, etc.), datasets and dedicated plug-ins will be stored in different databases to ensure performance and high flexibility.

A further requirement was the support of sharing and reuse of data and code as part of our interface. GISLab was therefore developed as open source software. Similarly, we aim to support the establishment of an open research community, so that datasets will be uploadable and will be freely available for (re)use. This approach also ensures that the software components are transparent—a highly valuable feature for scientific research—and that they can be used by others. But this form of openness should not be seen as a sheer technical approach. Further reflective interaction with researchers made it clear that GISLab (or Alfalab in general) should not only provide access to datasets, but should also do so in ways that are meaningful for researchers. To this end, a Rich Internet Application (RIA) is being developed. It will streamline the process of visualizing and publishing datasets. In addition, it will provide researchers with an interface to explore the available data through visualizations.

4 CONCLUSION

Alfalab has developed expertise to implement interfaces for interaction between research questions, tools and data, and infrastructures. As a result, the research-driven, bottom-up approach has met top-down, infrastructure-focused projects. On a practical level, Alfalab supports the very heterogeneous heuristics and data sources used in the humanities research community, while also creating solutions that meet standard frameworks (e.g. CLARIN compliant). Our highly modular technical approach has the added benefit that several data sources can now be used for cross-disciplinary research: databases in GISLab are being used for linguistic, non-geographically-oriented research. Because our interfaces are the outcome of encounters in which diversity is valued, our technological solutions reflect and support this diversity and are especially well suited to the humanities.

REFERENCES

- Antonišević, Smiljana and Anne Beaulieu. "Crossing the Unexpected: Benefits and Challenges of Scholarly Collaboration in a Humanities Labs." Paper presented at the HASTAC 2010: Grand Challenges and Global Innovations Conference, April 15, 2010.
- Dalen-Oskam, Karina van. "Please eLaborate!" *Madoc* 23, 1 (2009): 36-41.
- Dormans, Stefan and Jan Kok. "An Alternative Approach to Large Historical Databases: Exploring Best Practices with Collaboratories." *Historical Methods* 43-3 (2010): 97-107.
- European Union, *Riding the wave. How Europe can gain from the rising tide of scientific data. Final Report of the High Level Expert Group on Scientific data. A submission to the European Commission.* Italy, 2010.
- Ribes, David and Charlotte P. Lee. "Sociotechnical Studies of Cyberinfrastructure and e-Research: Current Themes and Future Trajectories." *Computer Supported Cooperative Work* 19-3,4 (2010): 231-244.
- Wouters, Paul and Anne Beaulieu. "Imagining e-science beyond computation." In *New Infrastructure for Knowledge Production: Understanding E-Science*, edited by Christine Hine, 48-70. London: Information Science Publishing, 2006.
- Zeldenrust, D.A. "DIMITO: Digitization of rural microtoponyms at the Meertens Instituut." In *Humanities, Computers and Cultural Heritage, Proceedings of the XVI international conference of the Association for History and Computing*, 301-307. Amsterdam, 2005.

Zundert, Joris van, Smiljana Antonijevic, Anne Beaulieu, Karina Dalen-Oskam, Karina van, Douwe Zeldenrust and Tara Andrews. "Cultures of Formalization, Towards an Encounter between Humanities and Computing". In *Proceedings of The Computational Turn workshop*, 1-7. Swansea, 2010.

Zundert, Joris van, Zeldenrust, Douwe and Beaulieu, Anne. "Alfalab: Construction and Deconstruction of a Digital Humanities Experiment." In *e-Science 2009, Fifth IEEE International Conference on e-Science*, 1-5. Oxford, 2009.