Digital humanities in Estonia: digital divide or linguistic isolation?

- Sarv, Mari
  Estonian Literary Museum
  mari@haldjas.folklore.ee
- Kulasalu, Kaisa
  Estonian Literary Museum
  kaisa.kulasalu@gmail.com

Summary

In this paper, we will introduce the disciplinary developments in Estonian humanities that are intertwined with use of digital resources and methods. The aim of the paper is to compare and contrast the use of digital technologies in Estonian humanities to the developments of the field of digital humanities.

1. Introduction

The paper introduces the disciplinary developments in Estonian humanities that are intertwined with use of digital resources and methods. Estonia forms an interesting case. On the one hand, the post-socialist country is well-known as an example of technological innovation, it was the first to introduce paperless governmental processes, the electronic ID-card is compulsory document and its electronic functions are widely used, to name a few examples.[1] However, at the same time, the developments in the field of digital humanities have not taken place at the same pace with the ones in USA or Western Europe.

The aim of this paper is to compare and contrast the use of digital technologies in Estonian humanities to the developments of the field of digital humanities.

2. Use of the digital technologies in Estonia

In 2012, 78% of the population of Estonia aged 16-74 years used the internet. There are various electronic services, like e-Tax Board where 93% of income tax declarations have been made in 2012. From 2005, it is possible to participate in elections electronically. In October 2013, 21.2% votes in the local elections were given electronically. And what is more, 62% of the inhabitants took part of the 2011 Population and Housing Census electronically.[2] However, the use of digital solutions is not a spotless success story: according to PIAAC (Programme for the International Assessment of Adult Competencies) study in 2011-2012, 30% of 16–65 year old test group refused to take a part of the test online and 13% did not succeed in the simplest computer-related tasks. The results of PIAAC study reveal a digital divide that is related to the linguistic divide: there are no significant differences in gender, education or the social background, but Estonian-speaking group had in general better skills in the use of technological tools than the group of the speakers of other languages.[3]

Apparently, Estonian society is not able to provide all its IT-advantages in the languages other than Estonian enough to integrate the non-Estonians to the Estonian IT-world, thus leaving them isolated to a certain extent. Despite being the wired country in terms of the e-Government, not all fields of life have been equally keeping up the developments in the IT, the humanities being one of these areas. The National Strategy for the Development of Information Society until 2020 addresses the problem of inequality and of moderate specific IT skills in the fields other than ICT.[4]

3. Field of digital humanities in Estonia

Schnapp and Presner[5] have distinguished between two waves of digital humanities: quantitative and qualitative one. According to the divide, the first wave was about digitisation projects and creating infrastructure, the other wave consists of interpretation and research methodologies for digitised and born-digital materials. Taking this approach as a framework, Estonia could be described as being in the first wave of digital humanities. In archives and libraries, there have been large-scale digitization projects from early 2000s onwards, the digital archival systems and repositories have been created and made publicly available for many different collections.

For example, in Estonian Literary Museum, a file repository and archival information system Kivike[6] has been created to manage the collections to store the digital materials and metadata, and to make the collections as much as possible publicly available via Internet. New archival data is being added to the online repository and
described there, among other materials the born-digital sources (photographs, audio and video recordings, as well as “digital manuscripts”) that have been collected systematically since 1994.

This illustrates the tendency of including some of the notions of second wave of digital humanities, because the emphasis on preserving of born-digital materials has also been apparent in several institutions. It is partially connected to the practical needs of document management in a country where governmental practices have moved online. The digital preservation department in Estonian National Archives was founded in 1999 with the goal of “permanent preservation of digital data despite the changes in society and technology”. Web pages are also digitally archived: Estonian National Library is in charge of creating and maintaining a web archive of the web resources that are important for the Estonian culture. For born-digital materials, these approaches still focus on creating an infrastructure rather than using the data. Creating and maintaining the collections have been in the centre of digital humanities projects in Estonia, using them for educational or research purposes is scarcer.

One of the most clearly developed disciplines that uses digital data sets and methodology is linguistics. There has been a constantly developing co-operation of the linguists and computer-scientists since the 1950s in the field of language technology, there has been formed a specific curriculum of computational linguistics at the University of Tartu in the 1997\[7\], large corpora and dictionaries are in constant progress and in 2006 the special national financing program has been founded in order to promote the development of language technology and digital language resources.[8]

Except of the linguistic studies, the use of digital methods in the humanities research has been somewhat sporadic and depends of the interests and skills of individual researchers. In the curriculums of humanities there are no comprehensive courses on digital research methodology (the social sciences with its traditional data analysis are in much more better situation here). In some of the humanities fields digital methods are being taught in specialised courses. For instance, students of archival science need to follow the course on digital preservation.

To sum up the situation in Estonia, the digital infrastructures for the major humanities collections in Estonia have been created by now or are in progress. There are several original web solutions for presenting and/or collecting the humanities data, a couple of successful examples of crowdsourcing have taken place. The creation of the software for the audience in humanities has been more than modest. Because of the linguistic restrictions, and the small size of researchers' community the specific software (e.g. for the analyse of the Estonian folk song melodies) would be tailor-made for the needs of (the group of) individual researchers and is not developed for the public (or multilingual) use. Use of the digital methods in the research has been constantly increasing, but only occasionally, the lack of systematic technological knowledge base seems to be crucial. All this does not apply for linguistics. However, the research mostly takes place within the disciplinary boundaries and the collaboration between the digital humanists of different disciplines has been modest, without signs for the need of the umbrella discipline.

Only in October 2013, first seminar of digital humanities in Estonia took place providing an overview of the projects in progress within the field.[9] Different fields of humanities were represented: linguistics, archeologists, literary studies, folkloristics, arts. In the presentations, digital humanities in general got little attention, most of the presenters focusing on digitizing, maintaining and/or presenting the collections (either digitized or born-digital) rather than interpreting the data. In addition to this a couple of research projects were presented, as well as an technologically innovative film project, two software projects for documenting the artefacts, and a few initiatives of using open data in web projects. It is clear, however, that the seminar did not involve all the digital humanities projects in course. For example there was no paper from Estonian National Archives, although it is serving as the competence center in the field of the digital preservation in Estonia, and there are several IT projects ongoing.

After the seminar an informal network was created with a mailing list and homepage. According to a small survey taken in early 2014, the humanities scholars who work with digital technologies feel the need for collegiality in the field, i.e for wider digital competence among their colleagues both for developing new tools and using the already existing ones.

4. Discussion

Digital humanities became a discipline because the humanities scholars as users of digital resources needed to understand the digital mechanisms and get some academic recognition for using them. One of the crucial
questions of the field is how to bring the technological advantages and knowledge within the reach of researchers so they could develop the tools and environments they need. To code or not to code, that is the question, when being probably the only person in the world who would need a language-specific tool, or if the audience of a web environment would consist of mere ten people. The size of the linguistic community restricts the reasonable amount of work hours spent for the IT solutions or web-pages. On the other hand, the available English-language software is often either not known or not easily applicable for several reasons. This is how the linguistic divide becomes the digital divide even in the country with well-established technological infrastructures. The awareness and the use of various IT solutions in Estonian society with its linguistic divide between Estonians and non-Estonians follows the model of the situation in the digital humanities leaving the peripheral group to the isolation to some extent due to the language barrier. [10]

References