Corpora Ethnographica Online: Strategien der Digitalisierung kultureller Archive und ihrer Präsentation im Internet, Rostock, 26 – 28 september 2012

Meder, T.

published in
Fabula
2013

document version
Publisher's PDF, also known as Version of record

Theo Meder, Amsterdam

Corpora Ethnographica Online: Strategien der Digitalisierung kultureller Archive und ihrer Präsentation im Internet

Rostock, 26–28 September 2012

From 26 to 28 September 2012, the University of Rostock held an international scientific symposium entitled *Corpora Ethnographica Online*. The purpose of the symposium was further explained by its subtitle, namely developing strategies for the digitization of cultural archives and their presentation on the Internet. The initiator of this conference was Christoph Schmitt, director of the Wossidlo archive in Rostock (city in the federal state of Mecklenburg-Vorpommern). The Wossidlo archive is an ethnographic and linguistic archive that will be fully digitized in the coming years, in the so-called WossiDiA project (DiA = digital archive), thanks to a substantial grant. The digitization project will be executed by the folklore department (Institut für Volkskunde) in cooperation with the university library and the computer science department of the University of Rostock. Representatives of the two last-mentioned departments who were speaking at the symposium were, among others, library director Robert Zepf and information scientists Holger Meyer and Alf-Christian Schering. The conference was attended by representatives from the humanities of several national and international universities, institutes, museums, archives and libraries (for which reason the official languages of the symposium were German and English).

The entire symposium was interlaced by a number of meaningful connecting threads, which will, in succession, play a role of increasing importance in the years to come: the digitization of heritage, the importance of open access, the need for data enrichment, the value of crowdsourcing, the use of portals, the advancement of interconnectivity by means of harvesters and the need for large-scale standardization. Furthermore, international cooperation will have to be sought on an ever-growing scale. And last but not least, the scientific approach should not be neglected in the overall process.

The central topic of the symposium was the digitization of (both material and immaterial) heritage. Barbara Sosič of the Slovene Ethnographic Museum, for instance, talked about the digitization of the folklore collection of Boris Orel (1903–62) and his team. Another example was Risto Järv's presentation on the digitization of the Estonian Folklore Archives in Tartu. Suvi Kivelä explained how the Finnish Saami Archives use digital techniques for the disclosure and the preservation of the language and culture of the Saami people. On a side note, what applies to the digitization of the Saami collection, applies to many more collections in the humanities: the digitization process has only just begun and only a small

Fabula 54 (2013) Heft 1/2
DOI 10.1515/fabula-2013-0011

percentage of the material has been digitized and put online. The digitization of cultural heritage will cost a lot of time, energy and money in the coming century, but this will be well worth the investment, for it will enable us to disclose heritage on a permanent basis and on a global scale, not only to be cherished by the indigenous cultural communities concerned, but also to serve as source material for the educational and scientific community. Moreover, digitization promotes the preservation of fragile objects. For that matter, Germany appears to take the preservation issue highly seriously, for heritage of national importance is also put on microfilm and kept in a secure storage facility in the Black Forest (in the so-called Barbara-Stollen, a four hundred meter deep adit). In fact, the Federal Office of Civil Protection and Disaster Assistance (Bundesamt für Bevölkerungsschutz und Katastrophenhilfe, BBK) was – like the German Research Foundation (DFG) – co-sponsor of the symposium *Corpora Ethnographica Online*. Apparently, digitization in Germany will not imply a halt to the preservation of cultural heritage on microfilm. On the subject of digitization, one short but vital comment deserves to be made: one should constantly take into account the progress of the techniques available. No matter how high the quality of the present photographs of cultural objects may be, the future may require 3D representation of artefacts, enabling views from all sides.

The interest of open access is great: the new, digital museums, archives and libraries will be open 24/7. Access to heritage is of broad public importance. When Barbara Sosič spoke enthusiastically about the digitization process concerning Ore's fieldwork documents, she remarked, with a laugh: "I'm starting to sound like the tourist office now." It is vital to keep in mind that open access is not only intended for the heritage of the in-group, or for the tourist industry, but that it could also serve scientific research purposes.

Matthias Harbeck of the university library of Berlin emphasized that libraries should play a central role in maintaining the stability of the open access policy. Google easily moves and removes Internet documents, publishing companies and JSTOR charge their customers a fee before granting them access to our scientific articles, but universities ought to adopt a more noble cause: they should offer reliable scientific knowledge, which is stable, traceable, accessible and free of charge. As an example, Harbeck referred to the digitization of the collection of Lutz Röhrich (1922–2006) by the Berlin University Library.

Digitization for the mere sake of digitization is not the way to go: the process should have an obvious added value, which is not present in the analogue collections. Enrichment is the magic word here. The digital data will gain in searchability and obtain their added value as soon as we start adding metadata. These may be keywords, or geographical metadata which can visualize the distribution of certain phenomena – Andreas Bieberstedt gave the example of the atlas project on fourteenth- and fifteenth-century East Middle Low German. Another possible issue would be interlinking, which could (re)establish the connections between disparate data. Jutta Weber of the Berlin State Library provided a good example of this. In his day, naturalist and explorer Alexander von

Humboldt (1769–1859) kept up correspondence with distinguished scholars like Georg Forster and Charles Darwin. At the time, there was a network of corresponding scholars, who sent and received letters that have ended up separated from each other, in numerous archives and libraries. Due to the digitization of all this correspondence, the Kalliope project will be able to bring together letters that respond to one another.

During the evening lecture at Rostock City Hall, the Australian researchers Philip Batty (Museum Victoria, Melbourne) and Jason Gibson (Australian National University, Canberra) gave an impressive example of how digitization is capable of regathering (online) Aboriginal heritage from all over the world. The project consists in an elaborate reconstruction of the collection of the Australian anthropologists Walter Baldwin Spencer (1860–1929) and Francis James Gillen (1855–1912), who did extensive fieldwork among the Aboriginal population of Australia between 1875 and 1923. Letters, (written) recordings, photographs, wax cylinders with sound recordings and artefacts are scattered all over the world, in museums, archives and libraries. Their anthropological work was a source of inspiration for scholars like Emile Durkheim and Sigmund Freud. The online regathering of this enormous ethnographic collection must be regarded not only as a virtual return of Aboriginal heritage, but also as a source of great scientific importance, for in its present coherent state, the collection allows an overview that was lost.

The digitization of folklore material could take advantage of the ever-increasing interactiveness of the Internet. In other words, Web 2.0 should be exploited. Ethnology could, for instance, benefit from a phenomenon like crowdsourcing: methods such as using selected, knowledgeable volunteers to transcribe archival materials and suchlike are already being applied successfully, as we speak. In addition, there are the so-called portals that allow large digital cultural collections to be brought together. One such portal, which was frequently mentioned at the symposium, is Europeana (www.europeana.eu). The so-called harvesters take it one step further: these smart search engines are capable of performing searches in several databases on an international scale and of presenting results from several regions, countries, languages or cultures. This form of interconnectivity could be of great value to the folktale databases that are presently being constructed. It would be quite a scientific innovation if folktales could be retrieved simultaneously from several digital databases, like the Dutch Folktale Database, the folktale database of the Wossidlo archive, the Estonian folklore archive and the Georgian database of Elguja Dadunashvili. Other candidates (not represented at the symposium) could join this initiative, such as the Flemish Folktale Database, the Archive of Portuguese Legends and RondCat: Catalan Folktales Search Engine. In order to be able to put a harvester over all the participants, the databases must meet some standard requirements, like those set forth by the Dublin Core Metadata Initiative (DCMI), the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) or CLARIN. Several fields in the databases must be standardized and equalized. For images, there already exists an international classification standard

like Iconclass (which is primarily suitable for paintings, though). For text files, we will need standardization for the keyword system and some other classification categories. In addition, we may ask ourselves the question if our folktales will be sufficiently served by the ATU system and the Motif-Index alone. Ideally, an international group of folktale researchers would take up an initiative like the 'Type-Index of Legends' (a project that was suggested in the past, but never really got off the ground).

In conclusion, we ought to realize one thing. When we are dealing with a digitization operation of considerable proportions, at some point policymakers, politicians and taxpayers will bring up a vital question: before long, all your material will be digitized – but what will happen next? Or, to quote Christiane Cantauw and Jutta Nunes Matias: "Doch was nutzen die schönsten Informationssysteme, die besten Rechtsberater und die ausgefeilteste Quellenkritik, wenn es keine inhaltlichen Fragen gibt, die an die Quellen herangetragen werden?" It will be up to the scientists to confront the source material with research questions, be it old questions that could be asked again, be it entirely new ones. In my own presentation, I suggested some possible research approaches concerning folktale databases. The comparative research into variants in time and space has always taken a central position in folktale research. This research angle will soon enjoy the benefits of both a superabundance of available source materials and a more sophisticated set of techniques to detect variants and visualize their interconnections. We will shortly be able to record the specific mechanisms of the dynamics and variations in oral (and written) tradition, with much more precision than ever before. It goes without saying that the scientists themselves will remain responsible for the interpretation of the variation as far as its function and meaning are concerned. The research into historical and geographical distribution we have long been familiar with may soon enter into a promising liaison with Google Maps or Google Earth. Thanks to the advantages of enrichment, stories, places, objects, persons and historical developments can get interconnected in meaningful ways. And now that we will soon have access to an incredible number of digitized stories, structuralist research can venture a renewed claim on our attention. Before long, it will be possible to take on the analysis of motif sequences as narrative building blocks of fairy tales, legends, jokes and contemporary legends with much larger corpora than Vladimir Propp could ever have imagined in his wildest dreams. By using large folktale corpora, we could research the possible existence of a so-called narrative grammar, perhaps even a universal narrative grammar. It is up to the scientists to make creative use of the possibilities offered by the new digital collections and to continue discussions and collaborations with the specialists in the field of information technology.

In my perception, *Corpora Ethnographica Online* was a highly successful symposium, well worth repeating. After all, there is every indication that the twenty-first century will be one of large-scale digitization and also of computational research in the humanities, for there are signs that in this century, scientific research, too, will develop (at least in part) in the direction of the eHumanities.