

Text Analytics for Detecting Dutch, German, and Allied Perspectives on Events at the End of the Second World War in the Arnhem-Nijmegen Region

Antal van den Bosch (Radboud University Nijmegen, the Netherlands)

Marten Düring (University of North Carolina at Chapel Hill, USA)

We present MERIT, a project that brings together researchers in history, memory studies, and computational linguistics with the goal to undertake pioneering research developing new methods for the aggregation and analysis of digitized historical sources, specifically with the goal to (1) identify and group related documents across thousands of sources and (2) compare the structure of narratives and study how they have evolved over time. These technologies are developed with general applicability in mind, so that they may be used for the analysis of other digitized texts such as source editions, newspapers or interviews.

The project's historical component will focus on the end of the Second World War in the Arnhem-Nijmegen area, between the fall of 1944 until the spring of 1945. The liberation of the Netherlands led to frequent encounters between civilians and soldiers in the war zones. These interactions were marked by a high degree of uncertainty, differing anticipations and violence. Drawing on sources made available by the Airborne Museum (Oosterbeek), we will concentrate on micro-historical reconstructions of this epochal period of transition. We ask how memories of individual encounters have been shaped within and across social communities as well as over time, as the collection's time range is from wartime to the near present (post-hoc memories). For a multi-perspective reconstruction of these encounters and their commemoration we rely mainly on digitized sources and metadata on eyewitness reports of all sides, oral history interviews, and secondary sources.

In the collection obtained from the Airborne Museum we will first identify named entities using standard NER tools for English, German, and Dutch in order to obtain a higher level perspective on the geographical space our sources cover, on military and civilian organizations that feature in them, and on frequently occurring individuals. In a second step we will cross-reference the entities we have identified to find matches between organizations, locations and individuals co-occurring in different sources. This approach corresponds to MERIT's concept to move historical source analysis beyond keyword searching and browsing of archival contents towards finding specific patterns of interest. In this case the patterns we are interested in are the different perceptions of particular historical events by Allied soldiers, Dutch civilians, and Wehrmacht soldiers. These multi-perspective renditions of events e.g. utilized in combination with tourism or education allow us to connect places and history but more importantly, to open up access to the raw, conflicting, incomplete and messy side of history that is so often hidden away by polished, factual and - let's face it: boring plaques on sites of historical interest.