

## **8<sup>th</sup> Göttingen-Hildesheim Workshop, 23/24-9-19, Hildesheim**

### **Multimodality – Abstracts**

#### **Modern multimodal semiotics as a basis for an empirical turn in the Digital Humanities**

*John Bateman (Bremen)*

How the Digital Humanities (DH) will be able to do justice to the diverse disciplines and subjects currently under investigation remains an open question. It is commonly suggested that the DH face new challenges requiring new methods, tools and bridges between disciplines. However, robust progress on such issues will be unlikely without achieving a more foundational view of just what the DH are doing. In this talk, I propose that multimodal semiotics would offer a theoretical and methodological framework sufficiently broad to further many of the fundamental discussions currently underway within DH, with strong implications for both theory and empirical methods. I briefly set out the philosophical parameters within which multimodal semiotics operates and show how methodological guidelines for analysis relevant for DH follow.

#### **Introducing AI2D-RST: A multimodal corpus of 1000 primary school diagrams**

*Tuomo Hiippala (Helsinki)*

The Allen Institute for Artificial Intelligence Diagrams (AI2D) data set contains nearly 5000 diagrams that represent primary school science topics, which have been annotated for their content elements and semantic relations by crowd-sourced, non-expert workers. My presentation introduces AI2D-RST, an alternative, expert-annotated replacement covering 1000 diagrams in the original AI2D data set. AI2D-RST presents a new, multimodally-motivated schema for describing the diagrammatic representations, which accounts for (1) the hierarchical organisation of content elements, (2) the discourse relations that hold between them using Rhetorical Structure Theory, and (3) the connections that are signalled using diagrammatic elements such as lines and arrows. The annotation for each layer is represented using a graph, which use the same identifiers for content elements to enable integrating annotations in different graphs. Conversely, the stand-off annotation allows pulling apart multimodal structures and exploring their respective contributions to diagrammatic representations.

#### **Scientificity and audiovisual formats of knowledge production – outline for an essay-film about the digital image in the Humanities today**

*Stefan Pethke (Göttingen/ Berlin)*

Two starting points:

- a) the outsider perspective;
- b) a SciFi assumption stating that with hegemonic digital culture, audiovisual formats will invade and replace pure text formats in Academia.

Almost three months of audiovisual investigation, centered on interviews with agents on diverse fields of various universities (mainly in Göttingen and Berlin), this cinematic research project intends to put into words AND images something that could be called the state of the art in conveying the digital image, and more specifically the role images play in the systematic attempts to bring computer/data science together with an increasingly wider range of disciplines inside the Humanities.

An outline of the potential structure for that film-to-be is followed by the screening of an approx. 10 minutes sequence of moving images, i.e. an edit of interview passages which shall summarize the outsider's experience in the realm of Digital Humanities, and ground more than one futuristic fantasy in a deeper understanding of what the university stands for today.

## **AdA. A Digital Approach to the Analysis of Audiovisual Images**

*Jasper Stratil and Thomas Scherer (Berlin)*

Audiovisual images in their media specificity, i.e. the confrontation with a fleeting subject matter existing only in time and perception, have posed a challenge for film analysis that has been already addressed in film scores of Sergej Eisenstein or shot tables by Raymond Bellour. Recent developments in the applicability of digital tools have led to significant changes in the possibilities of creating and visualizing annotations in digital frameworks. We will present the current interdisciplinary work of film studies and computational sciences in the junior research group 'Audio-Visual Rhetorics of Affect' as a case study for systematic digital video annotations and their visualization. We want to suggest our video-annotation practice as an open-source research tool for film analysis that enables a scaling of detailed film-analytical approaches for a larger group of films, as well as a possible starting point for the sharing of annotation data among scholars. Our approach combines the detailed annotation and temporal segmentation of video files following a standardized routine including a machine-readable film-analytical vocabulary, with different modes of visualizing and querying this data (Agt-Rickauer, Hentschel, Sack 2018). This method does not focus on individual annotations and their values as information-to-be-processed but focuses on the interrelation of annotations as patterns and dynamics of viewing processes that support a phenomenological informed film analysis (Kappelhoff 2018). On the basis of these methodological considerations we will present first insights of a case study on how a web video from the occupy wall street movement creates an emotional attitude towards the complex societal problem of wealth distribution in the aftermath of the global financial crisis (2007–). We want to highlight how digital tools can be helpful in light of a hermeneutic engagement with audiovisual images and the qualitative analysis of their expressive qualities regarding questions of affect modulation and meaning making.

## **Expanding cinema and 0xdb**

*Sebastian Lütgert (Berlin)*

Sebastian Lütgert & Jan Gerber are two artists and programmers who developed the movie database 0xdb and its underlying software pan.do/ra. The more than 15,000 films in the database are objects that cover films hard to find online. 0xdb ist not just a database for films, but treats film as a veritable digital object, which allows new ways of dealing with films. The project offers a number of special features such as the visualization of the timeline, time-based annotations, additional information and interlinking with other objects and information, and allows for in-depth search. The project stands in the tradition of autonomous archives and other critical media practices and has collaborated with artists and political activists worldwide. The software as well as the movies are available for free.

## **Multimodal Stilometry for Comics**

*Alexander Dunst (Paderborn)*

This talk presents current works on the implementation of a multimodal stylometry for comics. The introduction presents a short overview over existing approaches for visual media und introduces the „Graphic Narrative Corpus“ (GNC). This is a collection of actually round about 270 comic narratives in book length which serves as annotation and training material. In the following two different approaches of combining text and picture measures will be discussed. This shows that an additive method which uses both information channels for the differentiation of authors and genres won't necessarily give better results than pure visual or text based stylometry. From the viewpoint of digital humanities the formalisation of analytical concepts is a promising approach. It will be explained by

means of the concept of stylistic complexity. Concluding, the actual progress at OCR training with help of neuronal networks will be presented. The importance of the close entanglement of picture and text recognition shows here as well.

### **Multi-modal responses to texts:**

#### **Combining lots of low quality data with a bit of high quality data**

*Rebekah Wegener and Jörg Cassens (Salzburg/Hildesheim)*

The observation of individuals reading texts reveals a complex interplay of individual, textual and contextual features. Different aspects of the text, the setting in which it is read and the current mood of the reader all influence the way a reader reacts to a text emotionally or in terms of interest. The reaction of the reader can be observed on many levels, ranging from spoken expressions to physiological measures, facial micro-gestures, reader posture and behaviour. These responses are brought together with aspects of the stimulus texts to reveal the textual triggers for reader expressed affect.

It has been argued that readers' reactions to a text are not just a reaction to the lexical items, but to their organization and the context in which they are embedded. To capture this complexity, we have proposed to model the reading process as a series of layers of contexts. But capturing this level of data complexity limits the scale of any study to the relatively small number of individuals that can be studied in the lab, reducing our ability to generalise our findings or verify results statistically.

In this talk, we explore the option to extend small scale lab based experimental data that is very rich with large scale but relatively poor data. This process involves aligning time-aligned data with text-aligned data. We present both an overview of the project as a whole and the process developed to align the different data types.

### **Development and evaluation of a NER system**

#### **to discern author and character names in literature reviews**

*Marcel Ritzmann (Hildesheim)*

In the days of natural language interaction via dialog systems like Alexa, query auto-completion or question answering systems and many possible applications more, there is a constant need for advancement in automatic text processing. Therefore Natural Language Processing (NLP) and Named Entity Recognition (NER) are always relevant. The purpose of this paper is to develop a Named Entity Recognition system, which can recognize very specific NE types. The domain of the corpus, which is used in this paper, consists of only literature reviews. Usually German NER often classifies Named Entities into 4 types: „PERSON“, „ORGANIZATION“, „LOCATION“ and for everything else „OTHER“. This paper however focuses only on the classification „PERSON“ and tries to discern different subtypes, namely “Author” and “Character”. Thus, the literature reviews in the corpus are very relevant for this paper, as the NE types “Author” and “Character” are found frequently in this kind of text. The NER systems GermaNER (cf. Benikova et al. 2015) and Stanford NER (cf. Faruqui, Padó 2010) were considered for the task of annotating new NE types in a new domain, but ultimately GermaNER was chosen, as it provides better customization options. To start off, GermaNER annotates a test dataset from the new domain, in its basic version. Afterwards GermaNER is trained with a training dataset, which contains only “Author” and “Character” annotations. Then it is tested again. Subsequently the optimal combination of Features is going to be tested through an iterative evaluation process. First results will be presented.

### **Research and Development of Relevance Feedback**

## **Approaches Applicable for Trend Detection**

*Noushin Fadaei (Hildesheim)*

Relevance feedback (RF) traditionally represents an augmented ad-hoc retrieval approach where the activities of the user or a simulation of it are constituted into the relevance model aiming to obtain more relevant documents. The information retrieval system (IRS) exploits the direct feedback of the users on the relevance of the provided documents (explicit feedback) or indirectly collects it following users' behaviour such as clicking to view, viewing timespan or scrolling (implicit feedback). Pseudo relevance feedback (PRF) performs the query expansion considering the top k relevant documents as positive feedback. IRS makes use of relevance feedback to narrow down the gap between the system's assumption of relevancy based on initial query terms and actual users' information need.

In this workshop, I would discuss the development of relevance feedback in information retrieval for general purposes. Therefore, negative, positive and topical feedback over passage marked or complete document would be discussed. I also cover the classic most-used method probabilistic relevance feedback as well as recent methods such as neural relevance feedback. The former allocates higher weights to the terms that are more frequent in the documents selected as positive than the entire collection, while in the latter the weights are obtained through a 3-layer feedforward neural network (NN) trained via back propagation (BP).

Finally I would explain the topic I work on, Interactive Trend Detection in Patents and discuss the applicability of the introduced RF approaches. The goal is to detect trends in the European patent database (until 2012); the trends represent the leading topics which can be a boost in one topic or result of merging topics. Nevertheless, the topic groups which are obtained through clustering methods strongly require to be weighed against the users' need and definitions to result in the most relevant topic lists. Since experts tend to look for different criteria of technology and are interested in a range of details, a simple clustering may still take lots of effort to align patents with their information need.