**KUHN — Plea**

<0 figures>

**A Plea for a Method-Driven Agenda in the Digital Humanities**

Kuhn, J. and Reiter, N.

This position paper advocates joint efforts towards establishing a modular architecture for ‘deep’ analytical approaches to DH collections such as text corpora.1 The core idea is to augment a typical DH project agenda with an element of *project-independent* and *interdisciplinary* bootstrapping, aiming to identify analytical subtasks beyond standard processing steps, which can be used as building blocks across different complex DH modeling processes (adopting McCarty’s [2005] concept of modeling as a transition through temporary, provisional stages). Our discussion focuses on text-oriented DH projects.

The guiding hypothesis is that despite the fundamentally different characteristics that the overall models may have across projects and disciplines, they can be broken down into modular steps, many of which can be shared and/or developed in a joint effort. For example, the detection of persona mentions (e.g., *the teacher*) in texts is a non-trivial task that is a necessary pre-step for, among others, social network extraction or intertextual relation detection.2 This task is related to standard Natural Language Processing (NLP) tasks (Named Entity Recognition, co-reference resolution), but it targets no inherently linguistic category, and a proper treatment requires insights from the humanities disciplines (here literary studies, but similar concepts underlie network analysis in history/political science). We envisage a collaborative bootstrapping process of refining a persona identification module, taking into account its applicability in different higher-level research questions.

Inspired by a well-established experimental methodology from computational linguistics, annotated test data and evaluation metrics should be fixed as operational prototype specification—possibly starting with a temporary initial specification, whose role in the overall module structure is subsequently adjusted through bootstrapping. Such specifications foster the exploration of technical modeling alternatives on the computational side and provide the basis for critical reflection (cf. Gibbs, 2011) of the analytical possibilities and limitations on the humanities side, which can subsequently lead to revised characterizations.

Although the goal of re-usable cross-disciplinary modules is very obvious—from an engineering point of view—a typical pattern in many DH venues focuses on the application and one-time adjustment of tools, thus underexploiting the true potential of modularization in DH projects. Some prominent interdisciplinary projects building service-oriented architectures have failed to connect with scholars (cf. Dombrowski, 2014), and experience shows that one of the hardest interdisciplinary challenges is to establish a mutual understanding about the strengths and limitations of some modeling component—yet when this has been achieved systematically, the potential for further developments is enormous.

To reach an effective sharing of modeling components, methodological commonalities have to be identified across projects, and the adequacy of an approach has to be assessed in a replicable fashion: we argue that a simple but effective method is to highlight the relevant (and only the relevant) dimensions in **reference annotations** for some dataset(s)—even if some related dimensions are yet ill-understood. This way of thinking is natural to computational linguists, but takes some getting used to for humanities scholars: The initial ‘sharable’ version of a subtask appears trivial and/or insufficiently justified from theory. Our plea is to set these reservations aside and engage in a cyclic process advancing towards a modular toolbox of deeper and more complex notions.

Ideally, this will lead to growing collections of shared reference datasets highlighting particular expectations for submodules; a project with slightly novel working assumptions may well add datasets with deviant characteristics (as long as decisions are well-documented at the meta-level). Methodologically, a rich collection covering a broad, interleaved space is the most helpful and may feed ‘shared task’ initiatives for computational modeling (which have been highly influential, e.g., at the Conference on Natural Language Learning). Based on such data collections, methodological innovations can be tested systematically across projects and disciplines.

# The Issue

At first glance, the situation seems to be perfect: Independently developed tools and models from NLP (integrated, e.g., by European infrastructure initiatives like CLARIN and DARIAH) can be applied to DH text corpora, facilitating the scaling up of research methods beyond the classical ‘analogous’ approach. This scenario comes with a rewarding division of labor: NLP developers can provide and adapt tools; trained scholars in the humanities have full competence to apply the analytical machinery in their research.

Often, however, project constellations are different: Available methods require non-trivial adjustments, data preprocessing poses considerable challenges, and/or analytical questions are addressed for which no tools are available. Although such constellations bear great potential for innovative ideas, practical issues often dominate the collaborative work. The time-consuming preprocessing and tool adjustment steps are often not publishable in the computational communities, and if the adjustments are successful, leading to new algorithmic approaches to scholarly questions, often most of the funding period has passed, so there is no time to work out and advertise the general methodological contribution.

Any predominantly content-oriented characterization of a collaborative DH project (otherwise surely the best justification of an agenda) walks a thin line of sacrificing methodological generality to respond to specific issues. Somewhat unfortunately, aspects that make a project interesting and challenging from a humanities perspective (exploration of new corpora / taking a different perspective than previous work) act against the objectives of systematic method development, which has to emphasize replicable results and modularization.

# Addressing the Methodological Bottleneck

We argue that the potential in advanced computational models and modular task definitions is underexploited in DH. So far, our considerations seem to imply that only dedicated mid- to long-term DH projects can overcome this status quo. However, we’d rather draw a different conclusion that works for smaller projects in a lively community: Rather than characterizing a DH project’s analytical questions top-down and relative to the specific data situation—which implies that the actual implementation cannot happen until after preprocessing and adaptation—characteristic methodological facets are anticipated early on, using prototypical manual annotation of reference data for *preliminary* specification. Ideally, such specifications are based on reference data that are considered well-understood in the humanities disciplines and have known structural properties.

Incidentally, this is the workflow that is behind NLP modules that are now broadly applied: Early linguistic corpus annotation activities in the 1980/90s (which to this date provide the training data, e.g., for part-of-speech tagging and parsing tools) had to make theoretically loaded, controversial decisions, and a long and methodologically diverse process followed. Few people would have anticipated the usefulness of syntactic parsing tools, well beyond NLP standard tasks (and syntactic theoreticians would probably have vetoed many design decisions if they had been in charge).

It is important to note that we do not assume that all ‘modules’ are working fully automatically. The collection of building blocks developed in the DH community may well include manual validation and hermeneutic interpretation steps. This ensures that the method-driven, modularization-friendly agenda that we advocate bears no implication about the characteristics of the composite DH model for a particular project. But for the non-computational building blocks, comparison and critical exchange across disciplines is just as helpful.

# Discussion

To some ears, our proposal may sound like a blunt attempt of computational linguistics to superimpose its purely method-oriented view on humanities scholars, who take their ultimate incentive from questions about content and interpretation. We have no intention to question established DH methodologies. Yet from conversations we had with researchers from a broad range of text-oriented DH initiatives, the picture arises that there is a great interest in tools solving problems similar to NLP problems, but so far there is no established methodological framework within DH for critically reflected development of such tools. The researchers we talked to tend to agree that community efforts towards a sharing of reference datasets and insights about the use of modular components for higher-level questions would be a very good complementation of the established DH methodology.

Seeing the specification of partial tasks for computational modeling as a cyclic, collaborative effort also avoids the detachment of scholars in the role of a mere ‘user’ of custom-tailored tools. Instead the DH scholar is *empowered* to influence crucial properties of the improved model by reacting to experiences from earlier stages.

# Notes

1. We acknowledge that this paper has been inspired by various methodological discussions about DH (including the Dagstuhl Seminar on *Computational Humanities* in July 2014 and the Leipzig workshop ‘Informatik und die Digital Humanities’ on 3 November 2014), involving a considerable number of researchers.

2. We consciously picked a sample task for which the beginnings of the methodology have already been established: Jannidis et al. (2015) present ongoing work on an approach for persona detection that expands the established Named Entity recognition framework. Bamman et al. (2014) is an example of advanced modeling in computational linguistics that takes literary characters into account. Reiter (2014) presents an approach to discover structural similarities over narrative texts that builds on the detection of characters/personae.

# References

**Bamman, D., Underwood, T. and Smith, N.** (2014). A Bayesian Mixed Effects Model of Literary Character. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pp. 370–79.

**Dombrowski, Q.** (2014). Whatever Happened to Project Bamboo? *Literary and Linguistic Computing,* **29**(3): 326–39.

**Gibbs, F.** (2011). Critical Discourse in Digital Humanities. *Journal of Digital Humanities,* **1**(1).

**Jannidis, F., Krug, M., Reger, I., Toepfer, M., Weimer, L. and Puppe, F.** (2015). Automatische Erkennung von Figuren in deutschsprachigen Romanen. *Conference Presentation at Digital Humanities im deutschsprachigen Raum*, Graz, February 2015.

**McCarty, W.** (2005). *Humanities Computing*. Palgrave, London.

**Reiter, N.** (2014). Discovering Structural Similarities in Narrative Texts Using Event Alignment Algorithms. Doctoral dissertation, Heidelberg University.