Valuable Research for Visualization and Digital Humanities: A Balancing Act

Stefan Jänicke

The digital humanities are a platform where scholars with different backgrounds collaboratively investigate cultural heritage research questions with the aid of computational methods. There are plenty of debates [26] to define the scope of digital humanities, discussing what projects and what players are part of the field. In his graduate course, Cordell chooses a different strategy to characterize digital humanities and to apprehend its contours, namely “by studying the theories and methods that undergird it, focusing on its projects and critical publications” [11]. Similarly, Alvarado states that “there is simply no way to describe the digital humanities as anything like a discipline” due to the too multifarious abilities a digital humanist needs [4].

What Is a Digital Humanist? According to Alvarado, a digital humanist requires (1) “to develop the deep domain knowledge of the traditional humanist”, and (2) “to learn a wide range of divergent technologies (including programming languages).” Furthermore, it needs (3) “critical discourses to situate these technologies as texts, cultural artifacts participating in the reproduction of social and cognitive structures.” In fact, Alvarado states that it is unlikely to “master all three, the scope of such a program is simply too vast and variegated.” Indeed, there are a few scholars with academic degrees in both the humanities and computer science who may meet these requirements. But, for example, a computer scientist who never studied a humanities subject cannot gain the necessary profound knowledge like a traditional humanist to fall into the category of a digital humanist. For instance, Ramsay’s viewpoint [36] is that “computer science is a sufficient, but not a necessary condition.” Following that idea, humanities scholars with programming skills would be closer to that definition. For that reason, Ramsay recommends humanities scholars engaged in digital humanities to learn how to program. In his essay On Building [34], he states that a digital humanist even does not really need to know how to program, only the ability to walk the tightrope of generating valuable research for both fields. Derived from my own digital humanities experiences, I suggest a methodology how to design a digital humanities project to overcome this issue.

Index Terms—Digital humanist, digital humanities conference, survey

Abstract—The value of a visualization evolved in a digital humanities project is per se not evenly high for both involved research fields. When an approach is too complex – which counts as a strong argument for a publication in a visualization realm – it might get invaluable for humanities scholars due to problems of comprehension. On the other hand, if a clean, easily comprehensible visualization is valuable for a humanities scholar, the missing novelty most likely impedes a computer science publication. My own digital humanities background has shown that it is indeed a balancing act to generate beneficial research results for both the visualization and the digital humanities fields. To find out how visualizations are used as means to communicate humanities matters and to assess the impact of the visualization community to the digital humanities field, I surveyed the long papers of the last four annual digital humanities conferences, discovering that visualization scholars are rarely involved in collaborations that produce valuable digital humanities results, in other words, it seems hard to walk the tightrope of generating valuable research for both fields. Derived from my own digital humanities experiences, I suggest a methodology how to design a digital humanities project to overcome this issue.
Science and those Humanities scholars who are less willing to accept computing as part of their research tools.” Such a setting enforces an intense collaboration between humanities scholars and computer scientists in order to generate valuable project results. The development of Poemage\(^1\) – a tool that supports the close reading of a poem by visualizing its sonic topology – is an appropriate example project (involving visualization scholars) that underpins Terras’ statement. Despite a difficult project start due to missing, clear project goals and the fear of humanities scholars that computational methods would rather inhibit than benefit their engagement with poems, valuable contributions were published in computer science [29, 30], the digital humanities [10] and also the humanities communities [9, 28]. To build a bridge from the initial question of this paragraph to the Poemage project, in my point of view, all involved researchers in this project – two computer scientists (visualization scholars) and two humanities scholars – can call themselves digital humanists. Aiming to advance their principal domains, they were forced to get acquainted with each others’ tasks, mindsets and workflows in order to evolve a common language. To me, this is the major characteristic of a digital humanist.

**AS NATURAL SCIENTISTS**, computer scientists in general and visualization scholars in particular are trained to systematically observe natural phenomena, to derive rules and formulas explaining those phenomena, and to use their gained theoretical knowledge to solve practical problems [1]. On the contrary, humanities scholars cannot describe their research questions by applying formulas. They explore and interpret the manifestations of human behavior and judge importance for individuals and society, aiming to raise human consciousness. Given these sheer different research approaches, it seems likely that digital humanities projects are not always success stories. It requires time and effort to understand each others approaches and to bridge personalities, which sometimes fails [12]. In our survey on visualizations that support the close and distant reading of cultural heritage texts, we collected the ingredients required to increase the odds for a digital humanities project to be successful [19]. In the following, I outline my own path from computer science to digital humanities.

**My Digital Humanities Story** As I started working on my PhD topic back in the fall of 2009, I would have called myself a visualization scholar with a decent interest in humanistic matters. In my first digital humanities project, europaena-connect\(^2\), I developed GeoTemCo\(^3\), a web-based visualization toolkit to compare geospatial-temporal datasets. Being focused only on contributions in the visualization domain, it was a rocky road to finally publish the comparative design idea [23] as – at that time – I was not working with humanities scholars who served me with a decent motivation in the form of real-world case scenarios. Aiming to get in touch with (digital) humanities scholars, I presented GeoTemCo at the European Summer University in Digital Humanities\(^4\) in Leipzig in 2012. Only able to visualize data items with a single georeference and a single time stamp, I stated in my conclusion that I would require a user to extend GeoTemCo to also being able to comparatively visualize trajectories and/or uncertain datasets. After the talk, my first digital humanities collaboration arose with David Joseph Wrisley,\(^5\) a medievalist also fresh to the digital humanities field with a research question at hand: the geospatial-temporal visualization of placenames with a varying granularity extracted from Medieval French texts with uncertain datings attached. Though we only had a small ALLC funding to organize a small workshop, we were able to present our idea and preliminary results in the form of a short paper [24] at the annual digital humanities conference in summer 2013. This gave me the opportunity to meet a number of digital humanists, to discuss the latest research outcomes, and to detect the potential of my new research topic: the visualization of textual variation. In close collaboration to Annette Gföner, a philologist interested in textual variation among English Bible translations, I developed close and distant reading visualizations, and we presented our interdisciplinary outcomes in the form of long papers at the digital humanities conferences in 2014 [21] and 2015 [20]. Although we published the computational aspects of our visualizations only in rather small visualization realms, especially the interactive design of the Variant Graph visualization TRAViz\(^6\) is well received in the digital humanities community [22]. Another fruitful collaboration I share with Josef Focht,\(^7\) a professor of musicology who moved to Leipzig in 2014 to lead the Museum of Musical Instruments. As we met for the first time in January 2015, he had a clear research question along with a biographical database for musicians. Given a musician of interest, he wanted to discover musicians with similar biographies – a task hardly feasible using traditional methods. With a very frequent communication including face-to-face meetings, telephone calls, and e-mail exchange, at times daily, we found a common language and designed a visual analytics system that supports the given research task of profiling musicians. Successfully published in the visualization domain [18], the idea is yet to be presented in the digital humanities community. But the feedback of the musicologists and the user statistics – around 2,200 profiling computations between August 2015 and July 2016 – underpin the value of the profiling system for musicology. Reviving my collaboration with David Joseph Wrisley, we presented our ideas on visualizing textual variance among medieval vernacular text traditions at this year’s digital humanities conference [25]. Taking into account the many digital humanities collaborations I had (and still have), which pushed my interdisciplinary engagement with humanities scholars, I think I share the above mentioned characteristic that defines a digital humanist. So far, I call myself a visualization scholar and – with a clear conscience – also a digital humanist.

**DEVELOPING VISUALIZATIONS WITHIN DIGITAL HUMANITIES PROJECTS.** I published my research ideas and outcomes at the annual digital humanities conference in the last four years. As digital humanists rather question the methodological approach than the chosen visual encoding, this was a necessary supplement to the feedback of visualization scholars. In my survey on close and distant reading visualizations [19], we emphasized the overall increasing value of visualizations as means of research for digital humanists since 2013, reflected by the number of publications in digital humanities realms. In the following, I take a look at the visualization papers presented at the digital humanities conferences in the last four years. I had two major questions in mind when skimming through the annual proceedings. First, I wanted to observe the ways visualizations were used to gain or to express knowledge about humanities data, i.e., if visualizations

\(^{1}\)http://www.sci.utah.edu/~nmccurdy/Poemage/

\(^{2}\)http://www.europaena-connect.eu/

\(^{3}\)http://www.informatik.uni-leipzig.de/geomteco/

\(^{4}\)http://www.culingtec.uni-leipzig.de/ESU_C_T/

\(^{5}\)http://djwrisley.com/

\(^{6}\)http://traviz.vizcovery.org/

\(^{7}\)https://www.gko.uni-leipzig.de/musikwissenschaft/institut/personal/museum-fuer-musikinstrumente/prof-dr-josef-focht.html
have been newly developed, or existing visualizations applied or modified. Second, I took a look at the qualifications of the corresponding authors to measure the impact of the visualization domain to the digital humanities. To investigate both questions, I looked at contributions published in the form of a long paper, representing substantial and completed research, significant new methodologies or critical discussions. Figure 1 contrasts the number of accepted long papers to the number of visualization papers. In accordance with our survey [19], which included all submission types and the visualization papers of the two major digital humanities journals, the increasing importance of visualization methods for the digital humanities is also reflected by the accepted long papers. In addition to the survey, whose scope ends in 2015, we see that the trend continued in 2016, reaching a peak of 21 long papers containing visualizations as means to communicate humanistic insights. I divided the paper collection into four types: (1) papers presenting a – to a certain extent – novel visualization approach, (2) papers applying existing visualizations, (3) theoretical papers, and (4) papers with unknown novelty. An overview is given in Figure 2. In addition to paper types, I investigated the academic backgrounds of the visualization paper authors by examining their homepages to spot their academic degrees, research interests and published works. This approach brought forth three groups of digital humanists: (1) scholars with a degree in computer science and a focus on visualization, (2) scholars with a prior humanities background and a sufficient visualization experience, and (3) scholars without traceable visualization skills, experience or research interests in visualization. The five authors, for whom I could not assess their academic backgrounds, are contained in the latter group of scholars. The participation of these three scholar groups in the publication process of digital humanities long papers generated (1) papers co-authored with visualization scholars, (2) papers co-authored with scholars having visualization skills, and (3) papers published by scholars without visualization skills (see Figure 3).

Visualization Paper Types at Digital Humanities Conferences

The first eye-catching fact is that more than half of the papers (34 out of 64) only apply existing visualization toolkits, which is in line with Ramsay’s idea of building. In these cases, graphs or timelines have been generated using toolkits like Gephi or SIMILE Timeline, dots are plotted onto maps, or existent D3.js examples were reproduced with own datasets. A closer look reveals that in 29 cases only authors without visualization skills were co-authoring the corresponding papers. In contrast, 11 out of 18 papers presenting innovative visualization ideas involved scholars with visualization skills, 9 visualization scholars among them. The contentual range of these papers includes innovative graph layouts, novel heat map representations, enhanced close readings, explorative visual analysis environments, or enhanced geographical map designs. Three theoretical papers containing critical discussions on the impact and use of visualizations in digital humanities were not presented by visualization scholars. The above mentioned facts show that in the last four years visualizations have been rather used than developed. But, when visualization scholars were involved, in most cases, the result contained also an innovative visualization approach.

Visualization Scholars at Digital Humanities Conferences

I conclude my paper with some thoughts on this problem. Tying up all the above paragraphs, I think that there are far too less visualization scholars who also fall into the category of a digital humanist. From my point of view, the presence of the visualization community at the digital humanities conference could not be worse than in 2016. It seems that we already lose the interest in this field, leaving the assignment to visualize digital humanities data to scholars without visualization skills who just build a visualization by applying existing tool and techniques. But, initiating a digital humanities workshop at the VisWeek is a good first step, but not enough. When we want to entrench visualization as a research subdomain in the digital humanities field, we need to get more in touch with digital humanists and humanities scholars to convince them that visualization scholars in particular and computer scientists in general are not just developers. We require to communicate that it is by far not the same to take any old graph visualization toolkit to generate valuable contributions for both the visualization and the digital humanities fields due to the diverse needs. I conclude my paper with some thoughts on this problem.
Walking the Tightrope On the one hand, as computer scientists we want to develop novel methods having a basic research value for the visualization field. But for us, the digital humanities is an application domain that rather stands for investigating humanities research questions than bringing forth computational innovations. In our survey, we could highlight that the digital humanities indeed provide a number of future challenges for the visualization community [19]. A further crucial issue is that many digital humanists have a prior humanities background, so that the application of complex algorithms and sophisticated visualization techniques – which meet the requirements of a grounded computer science publication – fails as prototypes are likely to be perceived as black boxes [14]. In such cases, the humanities scholar cannot adopt the new research methodology and a visualization might get invaluable for the digital humanities field. Figure 4 illustrates how a digital humanities project should be organized to increase the odds to produce valuable results for both parties. The get together at project start is the most important stage. It gives the opportunity to discuss what is possible and what is not. From my viewpoint as a visualization scholar, this is the best option to assess if a humanities research question has the potential to generate a contribution for the visualization community. In the worst case, indeed, only applying an existing tool might be enough, but my own experiences have shown that humanities scholars are often not aware of what is possible and what is not. The second project stage is an iterative development of the visualization, which is aligned with an iterative evaluation by (digital) humanities scholars. Although this period is not specific to digital humanities collaborations, recurring interdisciplinary discussions strengthen the mutual understanding between both parties, which further helps to improve the shared visualization idea. One of the most important lessons I learned during these interdisciplinary collaborations was that a visualization is typically not the end of a research process in a digital humanities workflow, and that the humanistic gain of knowledge often comes later. Although a visualization is capable of triggering hypotheses generation, a humanities scholar will unlikely proof a hypothesis based predominantly on an image. A statement from a humanities scholar working with a poem visualization, taken from [2], points out this fact by mentioning that “they would not likely look for insight from the tool itself ..., they would look for enhanced poetic engagement, facilitated by visualization.” When approaching a digital humanities project, we should keep in mind that for a humanities scholar, usually, a visualization rather provides a novel perspective on cultural heritage data triggering unfamiliar thought processes than that it represents the solution for her research question.

REFERENCES


