Visualization and the Digital Humanities:
Moving towards Stronger Collaborations

Abstract For the past two years, researchers from the visualization community and the digital humanities have come together at the IEEE VIS conference to discuss how both disciplines can work together to push research goals in their respective disciplines. This process has been both fruitful and challenging, bringing to light how different knowledge is created in the sciences and the humanities, but also how methodological differences can be traversed to produce truly interdisciplinary work. In this paper, we present our experiences as a group and our individual experiences as humanists and computer scientists. While we explore a broad spectrum of ideas, our goals are strikingly similar: to understand the processes and motivations of each other's work better as a way of increasing self-reflection on our own.

1 Introduction

Visualization researchers, always looking for new data into which to gain insight, have acted too soon in trying to map topics from the humanities into the domain of visualization. The prospect of rich data sets, some containing language that is thousands of years old, is tempting to engage in. But, in our excitement at the variety of different datasets available, we have clumped together the work of many disciplines and subdisciplines from the humanities and social sciences under the same rubric, which we simply label 'data'. On account of the different ways that the sciences and humanities create knowledge, the subtleties of humanities subject are often lost when we take them through a visualization design process. That is not to say that visualization specialists neglect their stakeholders, just that something is lost in the translation from the humanities to the sciences. Over the past three years and through two iterations of the Vis4DH workshop, we have come to understand that in these translations of work into data we have missed that these disciplines have as much to teach us about visualization as visualization has to teach them about

their subjects of study.

Humanities disciplines, as they have either come to adopt digital tools or have been subjected to digital analysis, are increasingly adopting the umbrella term ”Digital Humanities” (DH). But even scholars who identify as digital humanists admit that DH contains many different ways of working and thinking. This is both because the humanities themselves are a large domain comprised of many disciplines and because the kinds of digital interventions to which they have been subjected are so variable. Just as the purview of the humanities might be said to be constantly changing, we see the larger group labelled DH also constantly defining and redefining itself. Much DH work has been published in journals such as Digital Humanities Quarterly, Digital Scholarship in the Humanities as well as the ongoing series “Debates in the Digital Humanities.” There has also been a vibrant international conference scene, including, but not limited to, the annual Association of Digital Humanities Organizations (ADHO) conference. What has resulted from this vibrant, international debate is a broad and inclusive definition of the discipline: that any scholar using digital environments to showcase research or applying computation to humanities work is participating in the digital humanities. This ”big tent” notion is both freeing and limiting. Much research in DH embodies a implicit promise of “digitizing” the humanities, thus bringing it into the technological future, and learned societies and scholarly associations have funded that promise. The difficulty is that, seen from the world of computing, what is labelled digital humanities can vary widely: the digital edition of texts, basic studies of frequency in documents, geo-spatial analysis of archival maps, even the application of complex machine learning across corpora. A digital humanities project can participate in anything from digital archives, document digitization, library sciences, history, literary criticism, philosophy, theater history, visual arts, linguistics, and countless other subjects. We see this potential for interdisciplinarity as a strength, but the diversity also makes it difficulty to pin down a set of research directions based on the vastness of the possibilities for inquiry. To complicate matters further, humanities disciplines have an ever-changing abundance of research questions, and in many fields, a majority of scholars do not see the immediate relevance of computing to their work, beyond a basic service role. While the breadth of the humanities has made them an enticing source of data for computer science, it is not clear how computer science can provide authentic learning opportunities to humanities scholars. In order to be relevant to today’s humanities scholarship, we have to do more than simply map data to a view. We have to engage with the ways that the humanities traditionally build knowledge, and through embracing ways of thinking that are new to us, we believe that new shared languages between the disciplines are possible. The task is not an easy one, however. In this paper we present an overview of the past Vis4DH workshops at the IEEE Vis conference and some of what we as an organizing committee have learned about the relationship between visualization and the humanities. We present both accounts of the goings on at the workshop and testimonies from members of both the DH and Viz communities that speak to the wide ranging possibilities of ongoing and future collaborations.

We need to change our approach: Rather than seeing the humanities as a data source, we must begin to think of them as partners in a new hybrid epistemology.

2 Looking Back: Vis4DH Workshops 2016 and 2017

The inaugural workshop, co-located with IEEE VIS 2016 in Baltimore, was the first interdisciplinary event of this kind organized by visualization scholars aiming to bring together researchers from both the visualization and the digital humanities communities. A varied full-day program consisting of paper sessions and panel discussions comprised 17 accepted out of 21 submitted short and position papers. The diversity of digital humanities projects was also reflected by the presented works. Many papers addressed the development of visualizations to support visual text analyses [1]. Close reading visualizations, which focus on a limited amount of textual data, were designed to support text annotations, to facilitate the analysis of poems, or to compare classical text editions. On the other hand, distant reading visualizations, which support exploring large text collections, were presented for analyzing topic modeling results, to browse slave narratives with the help of tapestries, and to explore structure in Shakespeare's plays after topology-based analysis.

In addition, a conceptual work flow of the problem-solving process in digital humanities projects with visual text analytics was presented [2], and, independent of the data type, the benefits of combining Immersive Analytics and Deep Mapping were discussed [3]. Next to those rather application-driven topics, the workshop provided a platform also for critical perspectives on visualization as a method in digital humanities. This included talks given by digital humanities scholars about "slow analytics" as a method of sense-making in literary studies [4], the embedding of close readings to facilitate the interpretation of distant readings [5], and feminist data visualization [6]. Finally, the challenges of digital humanities projects to generate valuable research outcomes for both the visualization and the digital humanities fields were discussed [7]. Representatives of both research fields framed the first Vis4DH workshop with keynote talks. Gregory Crane, Alexander von Humboldt Professor of Digital Humanities at Leipzig University, outlined how visualization has changed the practices of mainstream humanists so far, and he introduced classical philology as a sub- domain of the humanities that provides a broad range of challenges for future visualization research. Min Chen, professor of scientific visualization at Oxford University, shared his team's experience in collaborating with digital humanities scholars, and he characterized the occurring problems as "treasure troves" as they offer "a broad spectrum of challenges and opportunities to the application as well as the advancement of visualization technology." Though both keynote talks pointed out the benefits of visualization for digital humanities research, it was evident that the approaches in how visualization is designed or how it is used differ among the disciplines. While, on the one hand, the (digital) humanities are often seen as a data source for visualization scholars, on the other hand visualization is often seen by the humanities as a mere provider of tools. As Jänicke [7] pointed out, digital humanities realms have not often been places where researchers from both domains meet. The first Vis4DH workshop can be seen as a success simply in providing such a meeting place, since 27% out of 70 attendees characterized themselves as digital humanities scholars.

Encouraged by the positive response to the first workshop, we organized a follow-up workshop co-located with IEEE VIS 2017 in Phoenix. Organized as a half-day workshop, VIS4DH2017 accepted 10 out of 17 submissions for presentation. As in the first workshop, the presented works addressed quite diverse research topics, but geo-visualization as a means to interact with digital humanities data, e.g., to explore language endangerment data, or to analyze relational changes in biography and prosopography data, provided a particular focus. Also, two talks

introduced movie data as a valuable resource for digital humanities projects. One position paper reflected on collaborative practices at the intersection of visualization and humanities research and proposed with a set questions to help researchers involved in digital humanities projects negotiating collaborative roles and project aims [8]. The keynote talk was given by Uta Hinrichs, lecturer at the School of Computer Science at the University of St Andrews, who described her concept of considering visualization in the digital humanities as “sandcastles” that, “in the process of their construction, facilitate the exploration of research questions, and, in their final stage, reflect detours and lessons learned”. She pointed out the value of visualization for digital humanities research to triggering new hypotheses that can lead to changing research perspectives requiring visualization re-design. The concluding audience-driven discussion generated questions like “What is the role of the human in visualizations for digital humanities?”, “How do we design visualizations to communicate data skepticism?” and “What can the humanities community do to help developing theory and methods for improving visualization literacy?” These remain to be addressed by researchers from both fields in the future.

[1] S. Jänicke, G. Franzini, M. F. Cheema, and G. Scheuermann, “Visual Text Analysis in Digital Humanities,” Computer Graphics Forum, vol. 36, no. 6, 2016, pp. 226-250.
[2] M. El-Assady, V. Gold, M. John, T. Ertl, and D. A. Keim, “Visual Text Analytics in Context of Digital Humanities,” Workshop on Visualization for the Digital Humanities (Vis4DH 16), 2016.
[3] J. Roberts, J. Mearman, P. Ritsos, H. Miles, A. Wilson, D. Perkins, J. Jackson, B. Tiddeman, F. Labrosse, B. Edwards, and R. Karl, “Immersive Analytics and Deep Maps - the Next Big Thing for Cultural Heritage Archaeology,” Workshop on Visualization for the Digital Humanities (Vis4DH 16), 2016.
[4] A. J. Bradley, H. Mehta, M. Hancock, and C. Collins, “Visualization, Digital Humanities, and the Problem of Instrumentalism,” Workshop on Visualization for the Digital Humanities (Vis4DH 16), 2016.
[5] B. Schmidt, “A public exploratory data analysis of gender bias in teaching evaluations,” Workshop on Visualization for the Digital Humanities (Vis4DH 16), 2016.
[6] C. D’Ignazio and L. Klein, “Feminist Data Visualization,” Workshop on Visualization for the Digital Humanities (Vis4DH 16), 2016.
[7] S. Jänicke, “Valuable Research for Visualization and Digital Humanities: A Balancing Act,” Workshop on Visualization for the Digital Humanities (Vis4DH 16), 2016.
[8] U. Hinrichs, M. El-Assady, A. J. Bradley, S. Forlini, and C. Collins, “Risk the Drift! Stretching Disciplinary Boundaries through Critical Collaborations between the Humanities and Visualization,” 2nd Workshop on Visualization for the Digital Humanities (Vis4DH 17), 2017.

3 The Vis4DH experiment

Overall the Vis4DH experiment has been a success, but not in the ways we originally thought it would be. We say that because we were thinking initially of our work as inherently providing a service to the humanities, as if any form of visualization intervention would provide added value to their work, but the humanists had different ideas: they wanted visualizations that gave insight into their research problems, not ours. This difference in perception speaks in part to the large gap in epistemological outlook between the two fields, a gap between very different ways of creating knowledge that at times led to misunderstanding and controversy. Some interactions from our first two years have ranged from on-the-floor disagreements sparking outrage to some genuinely interesting interchanges of ideas. The humanists, we came to realize, perceived that we did not understand the traditions and complexities underlying their data; it was as if we were trying to get them to adapt to our methodologies and vocabularies without likewise learning from their processes and workflows. Even the style of delivery of the papers between humanists and visualization people was different, the humanists often read papers prepared just for the event, while the presentations from viz scholars used slides to describe the contents of the submissions. Another difference was the expectation of debate after the paper was presented. Humanists use debate as a way of building knowledge and they expect the opportunity for question and answer periods at conferences. These debates can often get heated. While this was a surprise to the vis community, it seemed quite normal for the humanists in the room. This type of engagement, which is central to their practice, is often frowned upon in visualization. It took us two workshops

to identify and begin to address this important difference, and especially to realize that humanities scholars embrace the idea that truth is a shifting thing. While visualization scholars, on the other hand, are looking for clear answers from data. Since both sides call their results "insight,' it is not surprising that difficulties arise. In direct response to these observations, we relaxed the workshop format to provide space for the bridging of these two disciplines. We have also made room for papers that explicitly bring these different expectation into the fold.

This diversity in approaches has exposed not only epistemological but also methodological differences.

This got us thinking about how to work together and culminated in a workshop paper that discussed the differences between these methodologies. Over time we have learned that our partners in the humanities use argument to build knowledge, much as we in visualization use testing as a path to knowledge. Likewise, though rigor is equally important in both disciplines, it is regarded and even defined differently. A concrete example of these differences is in how Ph.D. theses are produced in both disciplines. In computer science, the work is completed first then written up in a thesis in the form of a report. The social sciences also often complete experiments and report the results as part of their work. In the humanities, on the other hand, the thinking-through-writing of the thesis is the actual work. From the crafting of a long argument emerges a well refined and researched position that moves forward an ongoing discussion on a topic. It makes sense, then, that everything that follows in both fields is different: methods of publishing, styles of conferences, and follow-up discussions. Where an academic paper by a humanities scholar participates in a scholarly conversation that, depending on the field, could have been continuing for 2000 years, a visualization paper means to contribute to a growing body of adequately evaluated knowledge. Visualization papers present technical analyses and statistics to prove their work is valid, and papers demonstrate techniques that can be re-produced in future work, while humanities work is written not to be reproduced but to be followed, understood, and questioned.

In the face of all of these differences, we have been able to come up with ways to address this gap without a *service mentality* from both sides.

First, there needs to be a genuine belief that we can learn from each other. Not that we need humanities data to visualize, but that we can actually learn something about visualization by thinking like humanities scholars and vice versa. This means we need to actively engage with epistemology and methods that are foreign to us and start to ask questions like: **How does doing visualization affect thought in the humanities and how humanities thinking affects work in vis?** Second, for true interdisciplinary scholarship in visualization and the humanities to work, we need to train a generation of scholars at home in both humanistic arguments and big data processing and mappings, a generation capable of exploring new questions with hybrid methods. **We imagine a lab where professors from both disciplines teach arts/science students side by side.** While combined art/sci programs exist, they usually focus on dichotomies like ethics/ hard sciences and not philosophy, interpretation, discourse analysis/ and computation. These programs need to combine technical programming skills, statistics, and mathematics with critical thinking, which includes expert competence in reading and writing.

Thus, problems of bias, interpretation, subjectivity, and ambiguity must be taught alongside problems of scientific rigor, decomposition, and algebra. What we need is humanities work actualized through technology and, conversely, comp sci advancements infused with humanities questions: a combined epistemology providing new ways of knowing. And thirdly, what we learned from the workshops is that there is a need for methodological transparency:

- **How can interactive visualizations support new questions, and new scales of research, in the DH community?**
- **How can we encourage DH scholars to seek out visualizations, or collaboration with visualization researchers?**
- **How does visualization with a DH focus differ from general research in the visualization community?**
- **How can we remove obstacles for humanities scholars wanting to use visual analytics approaches for their research?**

These questions and many others are open questions, but we have come to believe that you can have your focus in one field and work perfectly well with another discipline if and only if you understand the epistemology of the people you are working with. You need to be aware of what other people care about, and develop projects that address both of those needs.

Examples of Research Methods in Humanities and Social Sciences

In humanities and social sciences, many different research methods have been developed to enable the inquests for new knowledge. Some of these methods, such as **Survey Research**, **Usability Studies**, **Case Studies**, and **Focus Group,** have already been widely used in the field of visualization, while some others, such as **Content Analysis** and **Distant Reading**, are extensively supported by visualization and visual analytics tools. However, many research methods in humanities and social sciences, which are not commonly seen as typical visualization research methods, can potentially have a noteworthy role in visualization. The following list comprises a number of examples in alphabetical order.

Conceptual Elucidation --- This research method stems from philosophy and entails the elucidation of concepts. It systematically attempts to formulate and reformulate the exposition, explanation, exemplification, and illustration of a concept and examine its structural relationships with other concepts through decomposition, categorization, integration, and unification.

Critical Discourse Analysis --- Critical discourse analysis (or critique) is a research method of analytical, disciplined, and systematic examination of a written or oral discourse. It ties the discourse to specific contexts (e.g., historical, social, culture, etc.) In visualization, such contexts can in addition be applications, data, users, and tasks.

Ethnographic Research --- Ethnography is a research method for studying people and cultures systematically from the perspective of a group of entities being studied. In the context of visualization, this research method may involve observing behaviours of users and techniques in specific environments and gathering, analyzing and understanding users' interpretations of such behaviours.

Grounded Theory --- Grounded theory is a systematic method for deriving, analyzing, and refining the proposition of a concept, a model, or a theory, It instigates the necessity for a continuing effort to collect and analyze real-world data about the proposition. The phenomena about the proposition in different real world situations are repeatedly observed and systematically tagged with codes. From the codes, concepts, categories, models and theories are formulated and refined.

Historical Research --- This is a commonly-used method to collect and evaluate data about historical facts in order to reconstruct actions and events during a certain period of time as completely and accurately as possible. It facilitates the description, explanation, and understanding of the actions or events of some historical significance, and enables lessons learned from the past, juxtaposition with problems at present, and evaluation of hypotheses and predictions.

Longitudinal Analysis --- Longitudinal analysis is research method for studying short series of observations obtained from many respondents over time. It often involves the analysis of a cross-section of time series, repeated measures, and data collected at multiple spatial, structural or temporal levels.

Narrative Inquiry and **Creative Writing** --- Narrative inquiry studies narrative materials, examining how stories are structured, how they work, who produces them, and how they are consumed, as well as how they are silenced, contested or accepted. Some academics considered creative writing as an extension of literature studies, while some others see it incompatible with the tradition of critical study of literary forms. One supporting argument in favor of the former view is that experimenting with creative writing enables deep understanding of the evolution of languages, literary forms, and narrative techniques.

Participant Observation Research --- This research method differs from research methods based in brief engagements, such as surveys, focus group meetings, and think-aloud discussions. It emphasizes an observer's extended immersion in an environment (e.g., a culture, a family home, a workplace, etc.) and participation in its day-to-day activities. The observation involves the systematic description of artifacts, events, behaviors, relations, emotions, and so on in the environment concerned. In contemporary environments, imagery data and social media data are often captured during the observation.

Quasi-Experimental Design --- This research method is often used in social science studies where some confounding variables cannot be controlled or measured, and random assignment of participants or tasks are not feasible. There are several commonly-used approaches to the design and analysis of these experiments, such as *Nonequivalent Control Groups* design, *Regression-Discontinuity* design, *Proxy Pretest* design, *Double Pretest* design, *Nonequivalent Dependent Variables* design, *Pattern Matching* design, and *Regression Point Displacement* design.

Synoptic Method --- Synoptic method is a research method for studying ideas and concepts in conjunction with the historical evolution and co-evolution of academic disciplines. It entails close examination of the written texts in different historical periods when these ideas and concepts are discussed and critiqued.

WORKS CITED

[1] B. Somekh and C. Lewin, eds., *Theory and Methods in Social Research*, 2nd edition, Sage, 2011.
[2] C. Frankfort-Nachmias, D. Nachmias, and J. DeWaard, *Research Methods in the Social Sciences*, 8th edition, Worth, 2014.
[3] B. Matthews and L. Ross, *Research Methods: A Practical Guide for the Social Sciences,* Longman, 2010.
[4] Edith Cowan University, "Research Methodologies for the Creative Arts & Humanities: Research Methods," Feb. 2018; http://ecu.au.libguides.com/c.php?g=410634&p=2910426.

4 Bridging the Gap: Current and Ongoing Projects

As we have described throughout, we believe the most promising future steps in the overlap between visualization and the digital humanities will be those in which both fields' methods and practices are brought together in a way that allows them to inform each other. There are many ongoing visualization projects currently fostering this sort of two-way collaboration that cover a variety of key topics in DH. One such area is the increasing blend of close and distant reading. Close reading is the traditional practice within the humanities of thorough inspection of an individual work, focusing down to a single passage, paragraph, or even a single word at a time. By contrast, distant reading evokes the process of analyzing an entire collection of texts with a more "zoomed-out" view. Historically, the visualization community has been more closely tied to the practice of distant reading, but there is a growing trend of methods that combine the two. Projects like Serendip [1] and VarifocalReader [2] show how connecting aggregate views of document clusters to lower level close reading can inform high level insight and help build readers' trust in statistical models of text. Alternatively, tools like Poemage [3] and the more recent Metatation [4] are examples of how visualization techniques can open up new lines of inquiry and afford new interactions even at the level of close reading.

Another fruitful collaborative area is in the realm of geospatial data and mapping. Maps and time lines afford the ability to situate information in space, be it geographic data associated with physical places, temporal data associated with particular eras or events, or other types of data that can draw on the metaphor of a mapped position to help uncover relationships. This sort of work can be effective at drawing connections across boundaries. The Linguistic Landscapes of Beirut project has used digital maps to explore the multilingual characteristics of public language in the Middle Eastern country, Lebanon [5]. The Visualizing Medieval Places project has sought out connections between place names found in a corpus of medieval French texts using spatio-temporal visualization [6].

Determining similarity between texts is a common task in large-scale text mining as well as in the analysis of historical documents, especially in investigating parallel or evolving versions of the same document. J¨anicke and Wrisley have recently explored the use of visual analytics to identify, and visualize, variance in medieval poetry [7]. Abdul-Rahman et al. take an image

processing approach to similarity detection, applying it to the study and identification of commonplaces in 18-century literature [8].

There are many other emerging research challenges that can help shape the interaction between Vis and DH. The act of serendipitous discovery, long embraced by the humanities, is still actively informing the field of exploratory visualization–especially taking inspiration from the library sciences [9]. The role of aesthetics in our two fields is constantly evolving and being questioned, as we struggle with parallel questions contrasting the functional role of visually pleasing presentation with the goal of beauty in and of itself.

Finally, there are number of direct efforts to directly combat some of the one- directional impressions that people have of the Vis/DH relationship, through long-term collaboration defining new research areas. For example, the VisArgue [10] project has brought together linguists, political scientists, and visualization experts for over five years, pushing the boundary of semi-automatic argumentation analysis. At the DH Conference in 2017, Hinrichs and Forlini described the research process of visualization as distinct from the final end-products of visualization tools [11]. It is a similar argument that we make to the visualization community here.

[1] E. Alexander, J. Kohlmann, R. Valenza, M. Witmore, and M. Gleicher, ”Serendip: Topic Model-Driven Visual Exploration of Text Corpora,” IEEE Conference on Visual Analytics Science and Technology, 2014.
[2] S. Koch, M. John, M. Wörner, A. Müller, and T. Ertl, ”VarifocalReader — In-Depth Visual Analysis of Large Text Documents,” IEEE Transactions on Visualization and Computer Graphics, vol. 20, no. 12, 2014, pp. 1723-1732.
[3] N. McCurdy, J. Lein, K. Coles, and M. Meyer, “Poemage: Visualizing the Sonic Topology of a Poem,” IEEE Transactions on Visualization and Computer Graphics, vol. 22, no. 1, 2016, pp. 439-448.
[4] H. Mehta, A. Bradley, M. Hancock, and C. Collins, ”Metatation: Annotation as Implicit Interaction to Bridge Close and Distant Reading,” ACM Transactions on Computer-Human Interaction, vol. 24, no. 5, 2017, pp. 1-41.
[5] American University of Beirut, “Linguistic Landscapes of Beirut Project,” 2018; http://llbeirut.org.
[6] D.J. Wrisley, ”Locating Medieval French, or Why We Collect and Visualize the Geographic Information of Texts,” Speculum, vol. 92, no. S1, 2017, pp. 145-169.
[7] S. Jänicke and D.J. Wrisley, ”Interactive Visual Alignment of Medieval Text Versions,” IEEE Conference on Visual Analytics Science and Technology, 2017.
[8] A. Abdul-Rahman, G. Roe, M. Olsen, C. Gladstone, R. Whaling, N. Cronk, R. Morrissey, and M. Chen, ”Constructive Visual Analytics for Text Similarity Detection,” Computer Graphics Forum, vol. 36, no. 1, 2017, pp. 237-248.
[9] A. Thudt, U. Hinrichs, and S. Carpendale, ”The Bohemian Bookshelf: Supporting Serendipitous Book Discoveries through Information Visualization,” Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’12), 2012.
[10] M. El-Assady, A. Hautli-Janisz, V. Gold, M. Butt, K. Holzinger, and
D. Keim, ”Interactive visual analysis of transcribed multi-party discourse,” Annual Meeting of the Association for Computational Linguistics, 2017, pp. 49-54.
[11] U. Hinrichs and S. Forlini, ”In Defense of Sandcastles: Research Thinking through Visualization in DH,” Conference Abstracts of the Digital Humanities, 2017

5 Testimonies

What follows are testimonies by some of the senior members of the workshop team. The contrasting experiences highlight the breadth of perspective that the Vis4DH workshop has cultivated in its first two years.

5.1 The Poem Transformed

– DH Testimony by Katherine Coles

What was life even like before computers? I don’t want to remember. Digital tools get me to archives, searchable databases, facsimiles of handwritten Emily Dickinson poems in high resolution, and citations, all without getting me out of my pajamas. These days, I go to the library

only to visit the Digital Matters Lab or, old-fashioned as I am, to hang out among rare books and letter presses. I even recognize the value of digital textual analysis, and of the tools that visualize patterns in texts or across bodies of texts, though I prefer to analyze texts myself. Therein lies the rub: for me, a reader, digital tools, including visualization tools, are not the point; they are, like machines themselves, only tools, often as annoying as they are useful, sometimes giving, sometimes actually blocking, my access to what I want to read. They give me easier access to the materials on which I work, but they are not themselves my work, nor do they occupy its meaningful center. They provide a useful service, but are not at the bottom of innovation or revelation in my field.

Likewise, these days the creation of digital tools, including visualization tools, or making them faster and easier to use, for the most part requires not innovation but the reapplication of existing techniques in computer science. If, then, such work occupies space at the practical center of the digital humanities, it leaves us to ask where, in DH, original research and discoveries are likely to happen. If we are going to have an authentically innovative digital humanities discipline, in which a scientist who lives to pursue her interest in visualization collaborates meaningfully with a humanities scholar who gets out of bed in the morning eager to engage directly with that deeply human thing, the poem – well, you get my point.

Seven years ago, then, Min Chen asked the wrong poet to join a Digging into Data Grant [1] to work on poetry visualization. When I declined, he refrained from telling me why I should be interested in machines, and instead asked what I was interested in, which is reading and writing poems.

For someone fascinated by machines and what they can do, especially what they can solve, a poem may not in itself be interesting. Still, it presents enormous challenges in negotiating ambiguity and complexity, and so may open opportunities to work with the machine. It may even, as with the POEMAGE project I undertook with Miriah Meyer, Nina McCurdy, and Julie Gonnering Lein [2, 3, 4], provide opportunities to address open problems in computer science [5], which suggests that computer science may gain from working with poets and poetry. For example, machines so far cannot negotiate the thing-it-is to thing-it-is-and-also-isn't comparison that occurs in metaphor, even in basic metaphors, of the kind that represent the barest tip of any poem's iceberg and that any marginally experienced reader can identify. As neuroscientist Gary Marcus puts it, the computer lacks "a theory of the world and how it works," the ability even a toddler has to see the "difference between the reflection and the real thing" [6]. According to Marcus, "integrating that sort of knowledge of the world may be the . . . prerequisite to grander projects," including the development of AI. If in metaphor we have the equivalent of that reflection, and if poets and scientists worked to teach a machine enough about metaphor to visualize its operations, might we inch the machine that much closer to developing a worldview?

On the other hand, if a computer scientist is interested in what machines can do, a poetry scholar will interest herself in what poems do, and will not care for machines or visualizations that intrude themselves unproductively between her and the poem at hand. Unlike the visualization scientist, the poetry scholar pursues not solutions or clarifications but ever deeper ambiguities, questions leading to still more questions. If anything, an adventurous reader remains alert for, even actively desires, the moment in which what seems to be a stable reading is cast instead, by a

single moment of ambiguity, into dissolution.

My current work in poetics both revels in this reverse "aha" moment and has its foundation in and began with talks and essays I wrote out of work with visualization scientists. The book I am writing emerged precisely because I undertook this work and did so skeptically, and because the scientists respected and addressed my doubts, and tolerated my drive toward ambiguity, throughout the process. My collaborators led me, gently but firmly, to think about the operations of poems – first the question of what they do and then the question of how – in ways that were newly and excitingly technical and precise. Most of this thinking arose not through my actually using the machine to visualize poems but through thinking about how to get it to do so in rich, interesting, and potentially useful ways. Out of this shift in the way I was reading came new, higher-level theorizations about metaphor, sound, image, and how poems manage their temporalities. Even my poetry colleagues who remain profoundly skeptical about the value to their own work of reading through or even alongside a machine find my theoretical work (not the machine part, but the poetics part) valuable and interesting. They don't care how I got to my readings, just that I did.

Thus, my work with visualization scientists has been intensely valuable and satisfying for me, even career-changing, not because it has led to new applications in visualization (though I am glad it has), nor because it has substituted machine judgment for my own, or solved anything, or saved me time (anything but), but because it has returned the poem to me transformed, newly visible, not on the screen but on the page. In this, ironically, it is difficult to determine the work's usefulness within a paradigm that values quantitative measures; beyond counting talks and publications, you will have to take my word for it.

[1] A. Abdul-Rahman, J. Lein, K. Coles, E. Maguire, M. Meyer, M. Wynne, A.E. Trefethen, C. Johnson, and M. Chen, "Rule-Based Visual Mappings—With a Case Study on Poetry Visualization," Computer Graphics Forum, vol. 32, no. 3pt4, 2013, pp. 381-390.
[2] N. McCurdy, J. Lein, K. Coles, and M. Meyer, "Poemage: Visualizing the Sonic Topology of a Poem," IEEE Transactions on Visualization and Computer Graphics, vol. 22, no. 1, 2016, pp. 439-448.
[3] J. Lein, "Sounding the surfaces: Computers, context, and poetic Consequence," Western Humanities Review, vol. 68, no. 3, 2014, pp. 84-109.
[4] K. Coles, "Slippage, spillage, pillage, bliss: Close reading, uncertainty, and machines," Western Humanities Review, vol. 68, no. 3, 2014, pp. 57-83.
[5] N. McCurdy, V. Srikumar, and M. Meyer, "Rhymedesign: A tool for analyzing sonic devices in poetry," Proceedings of Computational Linguistics for Literature, 2015, pp. 12-22.
[6] G. Marcus, "Artificial Intelligence Is Stuck. Here's How to Move It Forward," New York Times, 29 July 2017.

5.2 Developing a Critical Lens on Visualization

– Vis Testimony by Christopher Collins

Over the two iterations of co-organizing the Vis4DH workshop my perspective on what it means to say "Vis4DH" and the mechanisms for hosting inclusive and productive workshops has changed. I started to see disciplinary divides everywhere, even in the workshop name: we separated the fields and said one was "for" the other. I observed a tension between technological sophistication and domain expertise. The workshop experienced a language divide: the very definition of insight, or visualization, and truth changed depending on which side one stood. And

we saw an expectation divide about the ultimate outcomes of a visual analytics process: solid answers, uncertainty, fuel for argument and opinion. There was also a rich diversity of how the disciplinary divide manifested: Humanities researchers as partners, participants, users, critics. These differences did cause some discomfort, but from the experience I have come to see discomfort as fertilizer for innovative thinking.

Indeed, a key learning outcome for me occurred during a tense moment in the first iteration of the workshop, when an attendee from the humanities challenged the very idea that visualization was a tool to operationalize humanities insight, and the implicit assumption that visualization researchers were helping humanists but not the other way around. I have come to agree that the interchange has to be bidirectional. We in computer science often (but not always) approach visualization as something objective, something that presents truth and helps with the analysis process. We evaluate accuracy and speed of insight, and we often approach cross-domain collaborations as opportunities to reduce the time-to-insight for those domains we "serve". However, as we saw in the workshop, humanists are not often interested in faster insights, and do not accept the single pointedness of an objective "view from nowhere" on the data [1]. Interpretation in the humanities is the work, and it cannot be replaced by a false notion of objectivity. We may hide behind the screen of science as visualization researchers, but indeed we are also designers who are creating experiences as well as algorithms. These are curated experiences, no matter how carefully we design the color scheme or how objectively we think we have chosen the visual encodings. This applies in all our work. The workshop series demonstrated to me that this is not a failing of visualization, but something we should embrace. We can be both computer scientists interested in developing and evaluating techniques, tools, and algorithms and designers welcoming the plurality and "beautiful mess" often involved in interpreting a visualization [2]. In earlier work lead by Drk [3] we challenged visualization designers to disclose data sources, support a plurality of possible interpretations, provide contingency for multiple ways of experiencing the data, and empower users to annotate, remix, share the visualization. However, our views remain outside the mainstream of practice in visualization research (though, others are considering these issues, such as the rhetoric of narrative visualization explored by Hullman and Diakopoulis [4]). Building on this, one of the most eye-opening papers from the workshop was Feminist Data Visualization by D'Ignazio and Klein [1]. This paper invites us to rethink the binaries in our science. Beyond the obviously reductive binary male/female gender distinction which appears often in data sets but ignores trans and gender-fluid people, the authors draw attention to the rich possibilities when we allow for visualizing data which does not fit into pre-defined categories. Interesting research and design challenges arise when we acknowledge that pre-conceived notions of the cardinality of a data dimension or even it's type may change or be open for debate. How can we visualize gender as multidimensional, continuous, and dynamic, for example.

In a feminist approach to data visualization, a plurality of views are welcomed, and the position of the visualization designer is acknowledged. D'Ignazio and Klein invite us to consider, for example, voices that are not represented on the design team but might be important in the project. By bringing this sort of reflection to our research process in science, we will certainly enable deeper insights, multiple viewpoints, discussions, and engagement of a wider diversity of people.

Building on these ideas has made me think twice when creating new visualizations and teaching information visualization. Does a color choice reinforce a negative stereotype? Does a map ignore the Indigenous history of a place? Does a scatterplot reduce the stories of suffering in a war to points on a screen? Does a visual encoding choice force a particular point of view on the viewer? The recent wide call for computer scientists to consider the ethics of the research we do and the systems we build [5] clearly applies to visualization research. But to consider the ethical implications skillfully, we need to learn from and collaborate with the humanities as well as provide training in these skills to the next generation of visualization researchers and practitioners.

In attending and planning this workshop we even experienced a divide about what it means to have a workshop. In computer science we often use a workshop as a venue for presenting shorter works, giving short talks heavily structured based on slides, with a minute or two for questions. Our humanities colleagues told us that delivering a pre-prepared talk, by reading it, is the norm, and that much more time and openness for argument, discussion, challenge, and reflection should be provided. I was skeptical; who reads a talk word by word? When I experienced those talks, however, I stopped my multitasking. I put away my laptop and hung on every word. It opened my eyes to the possibility of engagement when a talk is carefully crafted and delivered as an argument. I don't think I can do it. And that's ok. The Vis4DH (DH4Vis?) workshop series has allowed for the space of collaboration and connection, and for the realization that while we can drift toward the other discipline, it can also be fruitful to look across divides without trying to bridge them. I don't want to be a humanities scholar, but I do think I have a lot to learn from them which will help me improve the way I do my scientific work.

[1] C. D'Ignazio and L. F. Klein. "Feminist data visualization." Workshop on Visualization for the Digital Humanities (VIS4DH), Baltimore. IEEE. 2016.
[2] N. McCurdy, J. Lein, K. Coles, and M. Meyer, "Poemage: Visualizing the Sonic Topology of a Poem," IEEE Transactions on Visualization and Computer Graphics, vol. 22, no. 1, 2016, pp. 439-448.
[3] M. D¨ork, P. Feng, C. Collins, and S. Carpendale, "Critical InfoVis: Exploring the Politics of Visualization," Extended Abstracts on Human Factors in Computing Systems, 2013, pp. 2189-2198.
[4] J. Hullman and N. Diakopoulos, "Visualization Rhetoric: Framing Effects in Narrative Visualization," IEEE Trans. on Visualization and Computer Graphics, vol. 17, no. 12, 2011, pp. 2231–2240.
[5] D. Gotterbarn, A. Bruckman, C. Flick, K. Miller, and M. J. Wolf, "ACM Code of Ethics: A Guide for Positive Action," Communications of the ACM, vol. 61, no. 1, 2018, pp. 121–128.

5.3 A Poetics, not a grammar

– DH Testimony by David Joseph Wrisley

As we write this article, we are witnessing a particular ebullience about the mutual discovery of ways of thinking in and between visualization and the humanities. There is a body of literature emerging on the intermingling of scientific and humanities thought and the kinds of spaces (like the collaboratory) that encourage interdisciplinarity, but in this article we are not asking how the humanities intersect with *all* of the experimental or empirical sciences. Rather, we are specifically reflecting on how specific materials, the "data" of the humanities as well as evolving conceptual frameworks in the humanities, might find intentional and pleasing visual expression on a screen. This work can only take place in collaboration with our colleagues in visualization

who understand best the practical and theoretical stakes of such moves.

The meeting point of visualization and the humanities might be productively thought of as one of translation across media. Translation studies, a robust field of literary and linguistic research that finds its roots in the mid-twentieth century, provide a powerful toolkit for theorizing how we can be said to move from one domain to another, from one language and its ways of thinking to the next. “How” in the previous sentence could also be replaced by other adverbs: ”why,” ”when,” ”where,” or ”to what extent” we translate. Common sense approaches to translation claim that the best translators are the ones who are ”fluent” in both languages, and so understand in a deep experiential sense that a “house” and a “casa” are at once the same and different, but if we look at our collaborations, who do we know who is fully ”fluent” in subfields of the humanities and in research frontiers of visualization?

We might also stop to ask which theory of translation best fits the Vis4DH - DH4Vis scenario? Would we focus on the purpose of a humanities scholar wanting to translate her work to the screen, as Skopos theory did? In our blend of method should we think about translation of different cultures, either domesticizing or exoticizing difference in the source domain? Is the translation of visualization about semiotics, that is, passing from language-based modes of communication and argument to non-language based ones? It is too early to embrace only one of these frameworks, but we can say with certainty that translation is not a neutral process: it changes the source object. Much more reflection in this domain would be useful in order to construct that (ex)change as one of mutual value.

Two short reflections point us in that direction. First, about data. When humanities researchers come to a DH training camp and are introduced to digital tools and methods for the first time, the possibility of visualizing their subject material can be accompanied by a flash of initial euphoria. When, on the other hand, we are asked about what we want to visualize, or worse yet, what kinds of data we already have, for some the experience can resemble the simile of the cave in Plato’s Republic (514a–520a). We might realize in that moment that we don’t know that we don’t know, but we continue to sit chained, unable to turn our heads. We might even retreat from the possibility of new ways of seeing. Some brave souls rally to learn about how to build datasets, and, even like myself, to build them, but even then the data literacy gap can be quite wide. By data literacy, I do not mean only how to use the data already created in the world, but rather how to translate our materials into something rich that can be things given (in Latin, data) to others to work with. There is a whole domain that remains untheorized about how the materials of, and questions in, the humanities can be translated into data-rich form.

Second, from data to conceptualization. It is important not to persuade the visualization community to imagine importing materials and methodologies from an unchanging humanities. The word ”humanities” is used in the plural, not in the singular! Instead to work well together we need to understand how researchers in the humanities themselves blend disciplines together and enact theory within the practice(s) of their discipline. The humanities are, after all, already an interdisciplinary domain with a full gamut of praxis, ranging from the empirical to the artistic. One of the cornerstones of knowledge is the way that concepts are adopted, elaborated, combined and explained through the accretion of argument. Take Mieke Bal’s Traveling Concepts in the Humanities: A Rough Guide that argues for using concepts to understanding cultural texts not

simply because we can, but because concepts help us to understand those works better [1]. What are the concepts in visualization that match those in the humanities: absence? overlap? saturation? color? How does a listening browser parallel (or not) the rhetorical moves of reading or argument? It is possible to use available visual vocabularies and codes to go beyond their normal meanings?

Working with a researcher in visualization has been for me more of an exercise in poetics than one in grammar. My collaborator has often asked me if I wanted to experiment with this or that functionality, and, as with unknown words that I was learning for the first time, I slowly began to feel that I could embrace them. Bored with the dull out-of-the-box visualization that abounds in DH, with time, I wondered if I could begin to deviate, as an artist does, beyond what conventional uses of semantics allowed. This is, of course, known as ”poetic license”—not usually a mode allowed to non-native speakers of a language.

If we take the example of contemporary art that is ”doing” much more than what we see at first glance, I propose that humanities visualization collaborations revert to small explanatory statements—we might call them ”visualization licenses,” that allow for the complex weave of humanities and visualization thinking to be included in our practice. This might look like a small companion page entitled ”how to read this visualization” or textual snippets appearing with conventional hover over mouse behavior. It could take on more essay-like format, as Montfort and Strickland did in their piece ”Sea and Spar Between,” embedding the theoretical writing within commented out fields in the code or perhaps even computational notebooks that articulate the dance between theory and practice [2].

[1] M. Bal and S. Marx-MacDonald, Traveling Concepts in the Humanities: A Rough Guide, University of Toronto Press, 2002.
[2] N. Montfort and S. Strickland, “Sea and Spar Between,” 2010; https://nickm.com/montfort strickland/sea and spar between/.

5.4 Learning from Humanities and Social Sciences
– Vis Testimony by Min Chen

Scientific investigation in the field of visualization involves not only data, mathematics, algorithms, techniques, and systems, but also users and their tasks, knowledge, cognitive abilities, interaction with computers, collaboration with each other, and so on. It is not yet feasible to model all these human-centric aspects using mathematics or computer programs, even if it is possible in the distant future. The more we work with humanities scholars and social scientists in helping them analyze and visualize their data, the more we realize that many of the research methods humanities and social sciences can be deployed in visualization research.

For the past three decades, there have been many arguments articulating the usefulness of visualization, ranging from conveying meaning to amplifying cognition, and from recognizing hidden structures to gaining insights. There are also arguments that bad visualization may result in inconsistent and biased interpretations. Clearly, the value of visualization depends on the judgment of proper or improper visualization, which in itself raises many scientific questions that demand an in-depth investigation. Such investigation may benefit from the research method

known as Conceptual Elucidation in a detailed and structured analysis of proposed concepts (see also the Panel on Examples of Research Methods in Humanities and Social Sciences). It may also benefit from the Synoptic Method in reconstructing and describe the evolutionary path of, and interactions between, these arguments.

In order to generate “proper” or “effective” visualizations, several hundreds of guidelines have been proposed in various books, research papers, and blogs on topics of visualization. The applications of the majority of these guidelines are likely subject to different conditions, and their effectiveness may depend on the variations of data, users, and tasks. Some guidelines conflict with each other, while others are too vague for practitioners to follow [1]. In humanities and social sciences, the Grounded Theory has been extensively used to address problems of a similar nature. For example, consider a guideline as a concept. The grounded theory instigates mechanisms such as collecting data about the guideline (e.g., writing diaries), using the guideline in different conditions, identifying and categorizing repeated causal relationships, and so on. The approach of the grounded theory can facilitate more comprehensive understanding about each guideline and its continuing refinement. Such discourse and refinement can be systematically studied using the Synoptic Method.

In the field of visualization, we often report our observation of the users of visualization systems as part of an application paper. We rarely employ the Ethnographic and Longitudinal methods in humanities and social sciences to study our users. In humanities and social sciences (similarly in psychology and neuroscience), studying uncommon phenomena often yields significant understandings. While we all enjoy the insightful books about unusual neurological phenomena by the neurologist Oliver Sacks [2,3], perhaps some of us can also study visualization users with such enthusiasm, profoundness, and supererogation. The field of visualization can benefit from Participant Observation Research that would deliver articles entitled, for instance, “The analysts who refuse to learn parallel coordinates plots” and “The girl who can hear scatterplots.”

There are some obvious connections between research topics in visualization and research methods in humanities and social sciences. For example, storytelling in visualization can benefit from the methods of Creative Writing and Narrative Analysis. Studying the historical development of visual cognition, visual representations, visual metaphors, and visual literacy can benefit from methods in historical and anthropological research. The critique of visualizations is not trivial as the judgment of properness and effectiveness often depends on data, users, and tasks. Conducting critiques that are holistic as well as comprehensive and emancipative as well as empathetic can benefit from the disciplined and systematic approaches in Critical Discourse Analysis.

Visual representations can be considered as visual languages, while visualization images are instances of writings in visual languages. Our understandings about these visual languages at the lexical, syntactic, and semantic levels are not in anyway as sophisticated as our understandings about commonly-used spoken and written languages. Perhaps using the research methods in linguistics and literature, we will one day reach the same level of sophistication in understanding visual languages as we do today for commonly-used spoken and written languages.

[1] M. Chen, G. Grinstein, C. R. Johnson, J. Kennedy, and M. Tory, “Pathways for theoretical advances in visualization,” IEEE Computer Graphics and Applications, vol. 37, no. 4, 2017, pp. 103-112.
[2] O. Sacks, The Man Who Mistook His Wife for a Hat, Picador, 1998. [3]
O. Sacks, The Mind’s Eye, Picador, 2010.

6 The future of the Vis4DH workshop

The Vis4DH workshop has had two iterations with a third soon to come. The lessons we have learned from bringing together researchers in the digital humanities and visualization communities continues to both challenge and inspire. As a group, we have had to learn how to speak each other’s language, understand each other’s motivations, and ultimately find a middle ground that satisfies the needs of both communities. This has lead to a hybrid epistemology (research that satisfies the requirements of scholarship in both fields) and a call for training for hybrid researchers. We acknowledge that this is an ambitious goal, but the success of joint projects presented at the workshop has given us confidence that this is possible. Future directions for this group and these projects need to also include a larger dialog with both communities. The sheer size of disciplines and sub-disciplines in the humanities and in computer science make this a fluid process and even with the type of training that we are calling for, each new project will come with its own set of challenges. We embrace this as a way to push research in new directions and will continue to investigate the possibilities of these partnerships as the workshop matures and continues to move forward.

Acknowledgements

We would like to acknowledge all current and former OC and PC members for their contributions to the Vis4DH workshop: Miriam Butt, Michael Correll, Catherine DeRose, Angus Forbes, Valentin Gold, Markus John, Daniel Keim, Steffen Koch, Uta Hinrichs, and Florian Windhager.

More information on the Vis4DH workshop can be found at http://www.vis4dh.org/

Authors

Adam James Bradley is a Post-Doctoral Research Associate in the Visualization for Information Analysis Lab at the University of Ontario Institute of Technology. adam.bradley@uoit.ca.

Mennatallah El-Assady is a Research Associate and Ph.D. Candidate in the group for Data Analysis and Visualization at the University of Konstanz, Germany. mennatallah.el-assady@uni-konstanz.de.

Katharine Coles is a Poet and a Professor of English at the University of Utah. katharine.coles@utah.edu.

Eric Alexander is an Assistant Professor of Computer Science at Carleton College. ealexander@carleton.edu.

Min Chen is a Professor at the University of Oxford. min.chen@oerc.ox.ac.uk.

Christopher Collins is an Associate Professor of Computer Science and Canada Research Chair in Linguistic Information Visualization at the University of Ontario Institute of Technology. christopher.collins@uoit.ca.

Stefan Jänicke is a Post-Doctoral Researcher at the Image and Signal Processing Group at Leipzig University, Germany. stjaenicke@informatik.uni-leipzig.de.

David Joseph Wrisley is an Associate Professor of Digital Humanities at New York University Abu Dhabi. djw12@nyu.edu.