White Paper Report

Report ID: 105519

Application Number: HJ-50099-12

Project Director: Katharine Coles (k.coles@english.utah.edu)

Institution: University of Utah

Reporting Period: 1/1/2012-6/30/2014

Report Due: 9/30/2014

Date Submitted: 10/1/2014

WHITE PAPER
REF: HJ5009912
Imagery Lenses for Visualizing Text Corpora

Our role in the project Imagery Lenses for Visualizing Text Corpora primarily involves Objective 4 Under the proposal's "Summary of Research Questions and Objectives":

In this focused study, we will observe how poetry scholars (domain experts) examine small corpora and formulate new hypotheses with and without corpus visualization tools. We will verify the new hypotheses through human analysis, which in general is more reliable than computational verification, hence offering a more demonstrable evaluation that complements Objective 3. In addition, the study also provides tasks in Objectives 1 and 2 with a direct interface between visualization scientists and domain experts.

This description is extended and enumerated in Section 3.5: A Case Study in Poetry:

In particular, we will organize the following human-centred design exercises.
1. The visualization researchers will first observe how literature scholars examine a small corpus, and formulate and verify hypotheses without the aid of visualization.
2. The literature scholars and visualization researchers will work together in developing and refining several taxonomies, which will facilitate different spatial mappings.
3. The literature scholars will examine one or a few poetry corpora using the developed visualization techniques, formulating some hypotheses based on visualization as well as on their domain knowledge.
4. The literature scholars will then verify the hypotheses by close examination of the relevant text documents in the corpora, reporting on the correlation between visual features depicted in the visualization and the actual linguistic features discovered through manual examinations.
5. The visualization researchers will evaluate the effectiveness and usability of the visualization techniques through interaction with the literature scholars and observing the activities in (3) and (4).
6. The activities (2-6) will be repeated for different visualization designs (e.g., Sections 3.1-3.4) as well as for iterative refinement of the visualization techniques.

A. PROJECT ACTIVITIES

We completed our work on DID-funded research in spring of 2014 with the continued refinement of our now-launched web-based visualization tool, Poem Viewer (http://ovii.oerc.ox.ac.uk/PoemVis/), and also through continued work on presentations and papers to make public the ongoing results of our work.

Work performed May-July, 2012

During May 2012, the two literary scholars held preliminary meetings in Utah to enumerate what we considered the most important elements of poetry and those that might lend themselves to visualization. Perhaps to an even greater extent, we were concerned about where visualization might be of most interest to those who are active poetry scholars, most of whom, because of the nature of poetry, are engaged primarily in close reading practices that engage them directly with texts. In what areas, we asked ourselves, might visualization enhance or ease the close reading practice? Our

discussions led us to an initial focus on form and formal elements.

Whole-team meetings in June in Oxford were spent breaching interdisciplinary boundaries and learning from each other about our disparate fields. These meetings led to our decision to focus on sound for our visualizations. Over the following six weeks, we received and responded to three different visualizations. We noted that while there has been considerable interest in visualizing and analyzing repetitions and sound in poetry, little attention has been paid to what most interests poets and scholars: how these elements exist and change through time and how their temporal work is essential to the experience the poem creates and enacts for readers. This is one of the original contributions we are still working to make in the field.

Work Performed August-December 2012

We successfully carried out six design exercises and began implementing our "Poem Viewer" prototype. During this new phase of our research continued to execute #6 as we refined the program in response to beta tester feedback. Having discovered the benefits of an iterative design process, we invited Dr. Miriah Meyer, a visualization scientist at the University of Utah's Scientific Imaging and Computing Institute and expert on design studies, to join our team. Design studies are a newly formalized approach to collaborative visualization development that address domain experts' actual questions and interests by intentionally including multiple software iterations in a process aimed at deepening understanding and enhancing precision at each stage. With Dr. Meyer's help, we approached this implementation phase with an eye toward pursuing new insights even as we fine-tuned our tool. This intense theoretical and practical work repeatedly showed this project's potential for literary, visualization, and digital humanities study.

In early fall our Oxford team-members prepared a questionnaire for us, designed to identify what sonic elements were most important. Of the elements offered, we eliminated only two or three, which left us with thirty variables—a very high number. Some were fairly straightforward (such as assonance and alliteration), while others, like rhyme, were more complicated. In Louise Bogan's poem "Night," for example, a repeating sonic cluster appears in "estuaries," "restless," "inlets," "westward," "set," and "reflects." Poets identify these as rhymes even though the e/s/t sounds occur in different syllables, in rearranged orders, and at different points within lines. One of our insights during this meeting was using the International Phonetic Alphabet to see where in the mouth each verbal sound is located, and to track how various sounds move and change through a poem, something that existing poetry visualization programs cannot do.

This discussion revealed the challenge for our visualization scientists in not only incorporating every variable we require but in creating visual models that are not so cluttered or complicated that they are essentially useless. They addressed this challenge by designating spatial zones around the poem to distinguish phonetic from semantic variables, and to use colors, glyphs, and arcs in each of these zones to visualize specific poetic features. Since users do not need to see every feature at once, a menu allows them to choose which ones to activate at any given time. A paper explaining this strategy in

detail appeared in *Computer Graphics Forum*, one of the top visualization science journals.

In October our Oxford team-members came to Salt Lake City for a second series of group meetings. During these sessions we tested an early software prototype and performed a close reading of a poem. This close reading was longer and more in-depth than the impromptu readings in Oxford that helped us define our initial focus on poetic sound. The hour-long conversation included discussion of sound, line breaks, image, syntax, allusion and other poetic features, demonstrating how two domain experts might approach an individual poem. (We recorded this conversation for our visualization scientists' future reference.) Together we explored possible strategies for capturing such rich data in order to visualize poems as the complex dynamic systems poets and literature scholars understand them to be.

We came to the understanding that, for the purposes of visualization, our team is working to capture and convey poems' multidimensional dynamism through the metaphor of flow, approaching poems as *fluid (or fluids) moving via their linguistic elements, devices, and figures through a (self)defined space.*

Work Performed January-July 2013

Having executed our enumerated objectives, we launched our web-based visualization tool, Poem Viewer (http://ovii.oerc.ox.ac.uk/PoemVis/) while continuing to pursue and extend the research actively. Professor Coles made a research visit to Oxford in May, during which she made a promotional video, gave a talk for Humanities scholars interested in DH work, and worked with team members on refining the tool. Both Professor Coles and Dr. Lein used the tool during this period to engage with and write about poetry, specifically addressing our group's adaptation of the International Phonetic Alphabet to reveal and track sonic movement and change over time, which is innovative and distinctive in important ways and allows us to capture temporal development of multiple textual components, showing whether sonic changes emerge slowly or abruptly, and where such changes intersect with other poetic variables to create what we are calling sonic or poetic "turbulence." This strategy begins to address qualities such as intensity, duration, and interaction among poetic components in ways other existing software cannot. Our strong emphasis on time-dependent variability and reader/user engagement, signified most promisingly in the specific visualization of phonetic mouth-placement, marks our approach as unique. Our writing on poetry during this time reflected the effectiveness of the approach and also helped us to identify areas for future work.

We discovered, moreover, that our general orientation toward technology in the *reading process* is meaningfully unique. Unlike other software that relies on particular textual mark-ups and encodings by individual scholars (or groups of scholars) to represent a snapshot of what is already known about a pre-selected poem, our approach prioritizes the poem itself, along with an enduring openness to new insights to be gained through fresh poetic encounters. Poem Viewer uses computation and visualization to reveal structures, movements, patterns, and relationships that readers might not otherwise be aware of, in essentially any poem or fragment they select and load; it then allows the

readers to decide which patterns and structures are actually interesting and relevant to their specific, changeable concerns. In this way, it actively maintains the uncertainty and openness of the reading process.

Work Performed August 1-December 30, 2013
We worked with the Oxford team to continue to explore groups of poems as they developed alternate interfaces and approaches for this and subsequent tools, and we worked on close readings using the tool in preparation for our 2014 MLA presentation. Meanwhile, as our work on Poem Viewer began to taper, we began to launch a longer-term research agenda with Miriah Meyer. We submitted a successful funding incentive seed grant proposal to the University of Utah to assist in the transition from one project to the other.

Work Performed January-September 2014
In early January, we presented a panel paper at MLA about the insights on poetry we have gained as a result both of the DID work and new work that arose as a result of the DID project. Since January, Professor Coles and Dr. Lein have each also completed work on a major scholarly paper, which we have placed as companion essays in *Western Humanities Review*, a journal devoted to literary scholarship, going into more detail about the insights on poetry that have resulted from our collaborations. We still plan a third paper, which will incorporate results from this work and from ongoing work on a new tool, for which we will seek publication in a journal dedicated specifically to Digitial Humanities research; this will be devoted primarily to theory. As this suggests, and as indicated on the program's home page, we consider this research ongoing. In addition to our own new direction, we continued and continue to respond to new visualizations and refinements from the computer scientists on our original DID team, which has been been testing and verifying alternative methods of visualization and is currently rewriting a paper for submission to TVCG.

B. ACCOMPLISHMENTS
As we set out to do in our Digging Into Data proposal, we have developed a tool for visualizing poetic elements and techniques, Poem Viewer, and have tested and continue to test alternative visualization methods beyond those used in Poem Viewer. As noted in our previous interim report, our group's adaptation of the International Phonetic Alphabet to reveal and track sonic movement and change over time is innovative and distinctive in important ways. First, it helps capture *through non-animated means* temporal development of multiple textual components, showing whether sonic changes emerge slowly or abruptly, and where such changes intersect with other poetic variables to create what we are calling sonic or poetic "turbulence." This strategy begins to address qualities such as intensity, duration, and interaction among poetic components in ways other existing software cannot. Our strong emphasis on time-dependent variability and reader/user engagement, signified most promisingly in the specific visualization of phonetic mouth-placement, marks our approach as unique.

We have discovered, moreover, in the course of our research that even our general orientation toward technology in the *reading process* is meaningfully unique. Unlike

other software that relies on particular textual mark-ups and encodings by individual scholars (or groups of scholars) to represent a snapshot of what is already known about a pre-selected poem, our approach prioritizes the poem itself, along with an enduring openness to new insights to be gained through fresh poetic encounters. Poem Viewer uses computation and visualization to reveal structures, movements, patterns, and relationships that readers might not otherwise be aware of, in essentially any poem or fragment they select and load; it then allows the readers to decide which patterns and structures are actually interesting and relevant to their specific, changeable concerns. In this way, it actively maintains the uncertainty and openness of the reading process.

This work has given rise to thinking that has led to our current work on a new tool, Poemage, will promote these qualities even further both structurally and strategically through our metaphor of poetic *flow*. It uses a query-based interface to allow readers to experience and examine poems as *fluid (or fluids) moving via their linguistic elements, devices, and figures through a (self)defined space.* The breakthrough we have already achieved through this work is a computational framework for analyzing sound—a major breakthrough on the computer science side of our work. Based on this framework, we are working on a tool for visualizing and analyzing sound that we plan to prototype in 2015 and begin to implement in 2016.

C. AUDIENCES

The potential users of both Poem Viewer and the next-generation tools being explored both in Utah and the UK include not only poetry scholars and poets but Humanities scholars beyond this target group—anyone who is interested in considering how sonic and other kinds of figures work across texts. Other potential users include teachers and students working with texts—those of others and their own—within the classroom context.

D. EVALUATION

Evaluation and refinement of Poem Viewer is ongoing. The Oxford-based team has been testing Poem Viewer and the techniques developed within it on groups including but not limited to literary scholars. They are working on a paper detailing the results of this work.

E. CONTINUATION OF PROJECT

The teams in Utah and the UK both continue to work on aspects of the project, both together, as in the Utah team's continued assistance in responding to new visualizations and visualization techniques, and separately, as in the UK team's testing of the techniques developed in this project on scholars in other fields, such as Oceanography. Out of the successes and the lessons learned in developing Poem Viewer, the Utah team is pursuing a new direction with Miriah Meyer, a visualization scientists at the University of Utah who joined our initial team in fall 2012. This new direction will rely on concepts and strategies generated during this DID project but with the aim of designing a tool that

places even greater emphasis on user-directed analysis, aesthetic engagement, and poetic complexity. The new tool we have in mind will be of even greater use to the literary and digital humanities communities and also hold increased potential for other fields requiring time-dependent, contextualized interpretations of complex texts (such as medicine and diplomacy).

This phase of the project was supported in its initial stages by a Funding Incentive Seed Grant from the University of Utah, which has also funded an undergraduate research fellow. We have recently applied to the NEH for start-up funds to develop a prototype tool, Poemage, based on our early investigations. We are hoping to prototype Poemage in 2015 and begin implementation in 2016.

F. LONG-TERM IMPACT

We expect the long-term impact of this work to be significant. In its own right, Poem Viewer represents a breakthrough in tools for visualizing poetry specifically and texts generally, as it allows scholars to load and interact with poems or other texts of their choosing. The development and use of this tool has also led within our group to new theoretical thinking about poetry, the nature of close reading, and what a close reading tool does and does not need to do. This thinking is already leading to the development of new kinds of close reading tools.

In addition to its impacts on poetry scholarship and composition, this project will also soon have impacts on teaching and classrooms. Both Poem Viewer and the tool currently in development will help teachers and students reveal poetic figures and devices "in action," as it were, thus enlivening and illuminating the experience of learning to engage poetry as an active, living art.

G. GRANT PRODUCTS

The primary product of this grant is the tool Poem Viewer, which is available for use at http://ovii.oerc.ox.ac.uk/PoemVis/. Additional products include six conference presentations, including two at the international Digital Humanities Conference 2013 and one at MLA 2014, three papers placed or out in journals, and two proceedings papers.

REFERENCES

Abdul-Rahman, A., J. Lein, K. Coles, E. Maguire, M. Meyer, M. Wynne, A.E. Trefethen, C. Johnson, and M. Chen. (2013). Rule-Based Visual Mappings—With a Case Study on Poetry Visualization. *Computer Graphics* Forum 32, pp. 381-390.

Abdul-Rahman, A., E. Maguire, and M. Chen. Comparing Three Designs of Macro-Glyphs for Poetry Visualization. In Proceedings of EuroVis 2014, Short paper, 2014.

Abdul-Rahman, A., K. Coles, J. Lein, and M. Wynne. Freedom and Flow: A New Approach to Visualizing Poetry. A. Abdul-Rahman, K. Coles, J. Lein, and M. Wynne. Presented at Digital Humanities 2013, Lincoln, Nebraska, July 2013.

Coles, K. "Slippage, Spillage, Pillage, Bliss: Close Reading and Machines. *Western Humanities Review* (forthcoming, 2015).

Coles, K. and J. Lein. "Turbulence and Temporality: (Re)visualizing Poetic Time. *Things My Computer Taught Me About Poems.* MLA2014. Chicago, IL. January 2014.

Coles, K. and J. Lein. · Finding and Figuring Flow: Notes Toward Multidimensional Poetry Visualization. K. Coles and J. Lein. In Proceedings of iConference 2013, pp. 444-448.

Coles, K. and J. Lein. Solitary Mind, Collaborative Mind: Close Reading and Interdisciplinary Research. K. Coles and J. Lein. Presented at Digital Humanities 2013, Lincoln, Nebraska, July 2013.

Coles, K. and J. Lein. A World in a Grain of Sand: Uncertainty & Poetry Corpora Visualization.

Coles, K. and J. Lein. Presented at Big Data & Uncertainty in the Humanities, Lawrence, Kansas, September 2012.

Lein, J. and K. Coles. Seeing the Sonic: Aesthetics, Poetry and Data Visualization. Presented at Aesthetics Re-loaded, Aarhus, Denmark, December 2012.

Poem Viewer. (2013). http://ovii.oerc.ox.ac.uk/PoemVis/ (accessed 31 Jan 2014).