

FREEDOM AND FLOW: A NEW APPROACH TO VISUALIZING POETRY

DID PROJECT TITLE: IMAGERY LENSES FOR VISUALIZING TEXT CORPORA

PI: PROFESSOR MIN CHEN, UNIVERSITY OF OXFORD

PI: PROFESSOR KATHARINE COLES, UNIVERSITY OF UTAH

SYNOPSIS

Under an international *Digging Into Data Challenge* program, our team of computer scientists, a linguist, and poet/scholars from the University of Oxford and the University of Utah have been working to create, through computation and visualization, a richer understanding of how poems work: one that relies on computational tools yet embraces qualitative and quantitative components and explicitly engages human readers and perspectives and research needs specific to the humanities in general and to literature, especially poetry, in particular. This new tool, *PoemViewer*, by approaching poems as complex dynamic systems, represents a significant step toward providing literary scholars freedom to explore individual poems, bodies of poetry, and other texts of their choosing in ways traditional scholarship and other text analysis software cannot. In addition to displaying familiar poetic features, such as texts, word frequencies, grammatic classes, and sentiment, *Poem Viewer* provides a unique capability for visualizing poetic sound, including various sonic relationships and changes as they occur in a poem over time.

Poem Viewer provides poetry scholars with a “magic lens” for observing many different attributes of a poem along its textual structure, and for comparing a poem with other poems or texts in such a multi-dimensional attribute space. In the sciences, this is often referred to as *multivariate data analysis*, while in poetry, such an observational and comparative study is commonly conducted in *closing reading*, typically with little help from any digital technologies. There have been several attempts to create such technologies, most notably *Myopia* (Chaturvedi et al. 2012). However, the existing techniques can only depict a few attributes simultaneously, and they commonly require manual encoding of individual features by users. While tools like *Myopia* are useful for literary scholars to disseminate their knowledge about a pre-selected poem pedagogically, they provide limited freedom for scholars to make their own observations or to come up their own interpretations (as Laura Mandell called “my utopia”).

Poem Viewer enables scholars to gain access to an advanced web-based visualization tool designed specially for supporting close reading (i.e., as computer scientists would refer it to as “multivariate data analysis in poetry”). It provides the flexibility for depicting up to 26 attributes with a rule-based user interface for minimizing effort required for specifying complex visual encoding of the attributes. Its capability for identifying and visualizing sonic elements automatically has been found particularly useful for revealing the development of sonic *relationships* in a poem across time. The latest version of *Poem Viewer* features automated translation of a poem to its phonetic representation, powered by *Oxford Dictionaries API*. This allows readers to select and upload any poems they wish for visualization. This represents a significant move forward for literary scholars, making the tool useful not only for lower-level pedagogical applications but for the pursuit of serious research.

In fact, our work has already led the poets on the team to rethink poems and how they work on a number of different levels. Specifically, the need to consider poems as comprising quantifiable elements subject to visualization has led us to think of the poem as *a fluid (or fluids) moving via its linguistic elements, devices and figures through a self-defined space*. At this point, the two poetry scholars on our team are working on an MLA presentation that will focus entirely on the literary insights gained from their engagement with the project and with *PoemViewer*. Katharine Coles is pursuing ideas on sound and image in time, and Julie Gonnering Lein is working in a focused way on sound and time.

Overview of goals and objectives of your DiD project.

The overall aim of this project was to answer a methodological question, i.e., “whether or not advanced visualization techniques can provide a useful interface that enables users (e.g., linguistic researchers, literature scholars ...) to use their enormous domain knowledge dynamically, while utilizing the computational and memorizing capability of computers. In particular, can data visualization help users make new observations and generate new hypotheses?” The proposal defined four objectives, addressing general corpus data visualization as well as an in-depth study on visualization in poetry research. We have achieved the overall aim, carried out R&D to address the four objectives, and achieved an outstanding in-depth study on poetry visualization.

Challenges and lessons learned from international collaboration across disciplines and domains.

There were two major challenges: (i) the distance between Oxford and Salt Lake City, and (ii) the “professional gulfs” between computer scientists and poets. We addressed both challenges successfully by making concrete arrangements for collaborative activities (e.g., visits, seminars, workshops, brainstorm meetings, video recording, joint publications, and software life-cycle) in addition to emails, phone calls and video conferences.

Digital humanities, social sciences and computational based research methods in the context of big data projects.

In this project, using advanced technologies such as visualization, humanities scholars are able to observe and compare visually, many different types of attributes of poems, including non-visual attributes such as the flow of sound and relationships between sound and structures. It is evident that the challenges and opportunities reside not only in the size of the data, but also, often more significantly, in the multi-dimensional attributes of the data.

Indicators of success. Measuring impact. Knowledge dissemination mechanism and tools.

It was a rewarding collaboration. The team jointly delivered two publications and a piece of public-domain software. The project was reported in the JISC Inform online magazine, together with a video interview of Professor Katharine Coles, co-PI of this project.

1. Abdul-Rahman, J. Lein, K. Coles, E. Maguire, M. Meyer, M. Wynne, C. R. Johnson, A. Trefethen, and M. Chen. “Rule-based visual mappings - With a case study on poetry visualization.” *Computer Graphics Forum*, 32(3): 381 - 390, 2013.
2. Abdul-Rahman, K. Coles, J. Lein, and M. Wynne. “Freedom and flow: A new approach to visualizing poetry.” *Presented at Digital Humanities 2013*, Lincoln, Nebraska, July 2013.
3. K. Coles and J. Lein. “Turbulence and Temporality: (Re)visualizing Poetic Time.” *To be presented at Modern Language Association 2014*, Chicago, Illinois, January 2014.
4. K. Coles and J. Lein. “Solitary Mind, Collaborative Mind: Close Reading and Interdisciplinary Research.” *Presented at Digital Humanities 2013*, Lincoln, Nebraska, July 2013.
5. K. Coles and J. Lein. “Close Reading in the Machine.” *Presented at Great Writing: The International Creative Writing Conference*, London, England, June 2013.
6. K. Coles and J. Lein. “Poets in the Lab.” *Presented at the University of Oxford*, Oxford, England, April 2013.
7. K. Coles and J. Lein. “Finding and Figuring Flow: Notes Toward Multidimensional Poetry Visualization.” *Presented at iConference 2013*, Fort Worth, Texas, February 2013.
8. K. Coles and J. Lein. “A World in a Grain of Sand: Uncertainty and Poetry Corpora Visualization.” *Presented at Big Data and Uncertainty in the Humanities: THATCamp Kansas 2012*, Lawrence, Kansas, September 2012.
9. J. Lein. “Seeing the Sonic: Aesthetics, Poetry, and Data Visualization.” *Presented at Aesthetics Re-Loaded*. Aarhus, Denmark, December 2012.
10. *Poem Viewer*: <http://ovii.oerc.ox.ac.uk/PoemVis/>
11. “The diversity of data”, *JISC Inform*, issue 37, page 5, http://www.jisc.ac.uk/inform/inform37/DiggingIntoData.html#.Uhxvtj_GAil

Importance of working with libraries, archives and data repositories.

The availability of British National Corpus (BNC) helped many aspects of this project.

Capacity building and training (students and highly qualified personnel).

This project has enabled two research officers, Dr. Alfie Abdul-Rahman (Oxford) and Dr. Julie Lein (Utah) to develop their research career in visualization and literature respectively. In addition, another PhD student at Oxford, Eamonn Maguire has gained useful experience in collaborative research by contributing to the Poem Viewer software system.