

## Developing Visual Analytics for Text Data

Jonathan WEBSTER<sup>1</sup>, John LEE<sup>2</sup>, and Joe CHAN<sup>3</sup>

Institute for the Study and Integration of Graphical Heritage Techniques

ctjjw@netvigator.com<sup>1</sup>, jsylee@cityu.edu.hk<sup>2</sup>

We introduce a text visualization tool being developed at the Visual Analytics for Texts Group (<http://vistxt.ctl.cityu.edu.hk>) at the Department of Chinese, Translation and Linguistics, City University of Hong Kong. The tool is designed to help users search for and visualize patterns and structures in texts. The key objective of our research is to take text visualization beyond the word-based focus, making it possible to visualize multiple layers of lexico-grammatical and semantic patterning across multilingual, vast collections of text data.

We demonstrate a prototype tool that illustrates this objective by visualizing a collection of religious and poetic texts. Inspired by the “Blue Dots” platform at Berkeley, which allows one to visualize search results in 3-D graphics from the Korean Buddhist Canon, this tool distinguishes itself from its peers in two respects.

- It interacts with corpora richly encoded with detailed lexico-grammatical information, and with multiple languages.
- It also supports a methodology for obtaining rich analytic data, in our case using M.A.K. Halliday’s Systemic Functional Theory.