

## **Kenneth Benoit: “QUANTESS: Quantitative Text Analysis for the Social Sciences”**

### **Abstract**

QUANTESS is aimed at developing innovative methods for the quantitative analysis of textual data in the social sciences. These methods would be sharply distinguished by more traditional content analysis schemes for analyzing texts – whether computer-assisted or not – by their explicit treatment of words as pure data, from which inductive statistical procedures may be used to estimate latent traits. Besides unlocking features of the texts not possible through interpretative methods, the “text as data” approach also allows rapid analysis of huge volumes of text in any language, providing a means for researchers to deal with the ubiquitous textual data now available. Existing statistical methods for textual data analysis exist, but these are still primitive in their development, relying on untested assumptions and unproven applicability, based on only short “proof-of-concept” demonstrations. In addition, there exists no single book-length work explaining the field of textual data analysis for the social sciences. Finally, software tools for applying textual data analysis techniques, particularly the advanced scaling models, are poorly maintained and documented and not accessible to users lacking a high degree of programming ability. QUANTESS will deliver on all three fronts: methodological innovation, dissemination of knowledge uniting all existing knowledge in a graduate-level text (plus a website, short courses, and instructional materials including videos), and creation of powerful yet accessible free software (see <http://github.com/kbenoit/quanteda>) to be used for all analysis from the project and the resulting books and articles.