**BAKER — British**

<0 figures>

**The British Library Big Data Experiment: Experimental Interfaces, Experimental Teaching**

Baker, J., Terras, M., Mohamedally, D., Weyrich, T., Alborzpour, S., Georgiou, S., Stavrou, N. and Wong, W.

Many digital humanities–taught programmes aim to engage undergraduate and postgraduate humanists with computational methods and practices (Hirsch, 2012; Cohen and Scheinfeldt, 2013). It is relatively rare, however, to routinely engage computer scientists with the needs, methods, and worldview of historians, literature scholars, librarians, and related researchers (Spiro, 2012). This poster describes an ongoing collaboration between British Library Digital Research1 and the UCL Department of Computer Science (UCLCS2), facilitated by the UCL Centre for Digital Humanities (UCLDH3), that enables and engages students in computer science with humanities research issues as part of their core assessed work. We demonstrate that CS students can provide an experimental test-bed for developing, exploring, and exploiting technical infrastructure and digital content in ways that may benefit humanities researchers within a library context. Encouraging students to develop skills in a new (and often foreign) domain encourages their critical thinking and provides real-world, complex issues that stretch and develop their technical abilities as well as their understanding of user requirements. Furthermore, from the problems, issues, and potentials such collaborative working raises, we learn more about the nature of computational infrastructure we rely on for research, and perceptions of the institutions’ core business in delivering digital content. As the British Library has a vision for transforming access to and research with its digital collections, the British Library Big Data Experiment forms an important complement to the British Library’s ongoing infrastructure activities through enabling the development of experimental services that offer unconventional engagement with its digital collections (Farquhar and Baker, 2014).

All taught programmes in UCLCS require students to undertake an industry exchange4 where they work in teams as clients to an industry partner. Though UCLCS has experience with developing student projects in partnership with digital humanists (Martin et al., 2012), industry partners have tended to come from the financial or manufacturing sectors. The British Library Big Data Experiment is an umbrella for a series of activities where the British Library is the client for assessed UCLCS project work, allowing for a rolling, responsive program of experimental design, development, and testing of infrastructure and systems. At agreed milestones during the project, the British Library provides access to required data, knowledge of data structures, and project requirements. UCLCS and UCLDH jointly provide technical and academic support to the student teams.

In June 2014 the first British Library Big Data Experiment team was convened with a dissertation project, submitted in fulfilment of the MSc in Software Systems Engineering (Georgiou, Stavrou) and Computer Science (Alborzpour, Wong), using a collection of circa 68,000 17th- to 19th-century digitised volumes to underpin the design of a research-oriented web-based service. Microsoft Azure5 APIs were implemented that functionally scale to the data, whilst the students worked in close consultation with humanities researchers who may wish to use the capabilities of such a system. The final public output (http://blpublicdomain.azurewebsites.net/) represents an attempt to capture the complex and multifaceted needs of humanities researchers whilst offering unconventional services such as bulk download of text based on metadata queries, word frequency lists, and OCR text previews. Following this successful pilot, the British Library Big Data Experiment is undertaking further collaborative work, including machine learning and mobile app development strands in autumn/winter 2014 and a second MSc dissertation project in summer 2015.

Both UCLCS and its students have an appetite for embedding problems faced by memory institutions within CS learning outcomes. In partaking in such truly interdisciplinary project work, students develop new skill sets, question their assumptions about the role of library and humanities scholars, and provide useful experimental design within the institutional context. In addition, having CS students engage with humanities scholars as a routine part of their degree allows humanists to understand their research needs and institutional structures, from a different perspective. We present the British Library Big Data Experiment as a model ripe for reuse and we argue that the benefits of such collaborative programmes outweigh potential risks. The Big Data Experiment is, then, both an experiment in teaching and an experiment in involving and integrating those undertaking advanced study in computer science into memory institutions and humanities scholarship.

**Notes**

1. http://www.bl.uk/ and http://britishlibrary.typepad.co.uk/digital-scholarship/.

2. http://www.cs.ucl.ac.uk/.

3. http://www.ucl.ac.uk/dh.

4. http://www.cs.ucl.ac.uk/business/partnership\_programme/.

5. An open and flexible crowd computing platform for computational processing, https://azure.microsoft.com/en-us/ was used in this project due to UCL’s ongoing relationship with Microsoft.

**References**

**Cohen, D. J. and Scheinfeldt, T.** (2013). *Hacking the Academy: New Approaches to Scholarship and Teaching from Digital Humanities.* University of Michigan Press, Ann Arbor.

**Hirsch, B. D.** (2012). *Digital Humanities Pedagogy: Practices, Principles and Politics.* Open Book Publishers, Cambridge.

**Farquhar, A. and Baker, J.** (2014). Interoperable Infrastructures for Digital Research: A Proposed Pathway for Enabling Transformation. *Digital Humanities 2014*, http://dx.doi.org/10.6084/m9.figshare.1092550%20.

**Martin, W., Abdallah, D., El-Abiary, A., Dalbah, Y., Julier, S., Terras, M., Iliffe, R., Hawkins, M. and Weyrich, T.** (2012). Newton Spectrum: Nonlinear Text Browsing of a Large Corpus. In *Proceedings of Digital Humanities Congress 2012*, University of Sheffield, September 2012.

**Spiro, L.** (2012). Opening Up Digital Humanities Education. In Hirsch, B. D. (ed.), *Digital Humanities Pedagogy: Practices, Principles and Politics*. Cambridge: Open Book Publishers, pp. 331–64.