The Role of Disciplinary Analysis in Web Science Education

Clare Hooper
IT Innovation Centre
University of Southampton, UK
cjh@it-innovation.soton.ac.uk

ABSTRACT
This paper considers the ways in which Web Science education can benefit from an analysis method used to gauge disciplinary representation. Three key contributions are identified: 1) driving development of the Web Science curriculum; 2) teaching Web Science, i.e. considering its evolution over time and using the method to foster comparisons of Web Science with other like fields; 3) teaching the analysis method itself as an example of a mixed methods, Web Science method.

This paper addresses topic #1 of the Web Science Education activities (Web Science education programmes design).

Categories and Subject Descriptors
K.3.2 [**Computers and Education**]: Computer and Information Science Education – *computer science education, curriculum, information systems education*

General Terms
Human Factors

Keywords
Web Science; Web Science education; disciplinary analysis; methods

1. INTRODUCTION
The interdisciplinary nature of Web Science is both well-known and inherent to this field of research and practice. Previous work has examined the nature of Web Science's interdisciplinarity, most notably via empirical mixed methods analyses of disciplinary representation in Web Science publications [1] [2]. The previous work was motivated by various goals:

a) ground dialogue about disciplinary representation with data, identifying which disciplines are more or less represented,
b) identify problems in need of action regarding discipline diversity (i.e. identify missing types of research or the kinds of collaboration that we may wish to encourage), and
c) improve our communication as a community and our ability to reach out to communities with whom we wish to engage.

This paper asks the question: what lessons can be gleaned for the Web Science Education community? This paper identifies three ways in which disciplinary analysis contributes to Web Science Education. Firstly, insights into which disciplines are more or less present in the Web Science community can help drive the development of Web Science curricula. Secondly, the analysis provides a way to help Web Science students understand not only the current disciplinary make-up of the community, but how this has evolved over time, enabling discussion of the implications of this. Thirdly, disciplinary analysis is itself a sound Web Science method, drawing on mixed methods from multiple disciplines.

This paper responds to the goals of the Web Science Education workshop by discussing the relevance of disciplinary analysis outputs for Web Science Education and drawing on experiences of teaching disciplinary analysis to Web Science students.

Section 2 provides a brief introduction to the disciplinary analysis method, while Section 3 discusses the three ways in which the outputs of this method can support Web Science Education. Section 4 provides conclusions.

2. DISCIPLINARY ANALYSIS
Disciplinary analysis was trialled in 2012 [1] and refined in 2013 [2]. It consists of four stages: gathering data, conducting Natural Language Processing to extract topics, visualising the results, and conducting a substantial expert survey to interpret the results.

1. Data gathering: materials from the field at hand are collected. For Web Science case, sources include Web Science conference proceedings, Foundations & Trends in Web Science, journal.webscience.org, and other key papers.
2. Natural Language Processing: this was done with Saffron, an application to understand research communities [3]. Saffron uses information extracted from unstructured documents with Natural Language Processing techniques. It yielded a set of ranked extracted terms.
3. Graphing and visualisation: a network graph tool, Gephi, is used to build a graph showing links between terms: nodes are extracted terms and arcs are papers that link them. This let us identify clusters of closely related terms. Detected communities can be interpreted as application contexts ranging from technologies (i.e. semantic web) to disciplines (i.e. elearning) and topic areas (i.e. social networking).
4. Expert survey: to map highly ranked terms with disciplines, an expert survey is used. Web Science experts are provided with the top ranked terms, and asked to map these terms to disciplines. This mapping is done as an expert survey to avoid issues of subjectivity and individual bias that impact the mapping if it is done by one or two individuals.

3. DISCIPLINARY ANALYSIS FOR EDUCATION

3.1 Contribution 1: Driving Curricula
Development of the Web Science curriculum has quite rightly been given great attention over recent years, evidenced not only by first class papers [4] but also the prestigious Web Science

Education workshop series, part of the Web Science conference from the first conference in 2009. Production of any curriculum, let alone one for a cutting-edge interdisciplinary area such as Web Science, is no easy task.

The outputs of disciplinary analysis can support Web Science educators by providing insights into two areas: Web Science application contexts and disciplinary representation.

Regarding application contexts, past work has identified four key Web Science application contexts: information retrieval; social networking; semantic web; personalised learning [2]. As can be seen, these range from technologies (i.e. semantic web) to disciplines (i.e. elearning) and topic areas (i.e. social networking). These application contexts suggest key areas to be covered when producing curricula.

Regarding disciplinary representation, recent work has identified the relative prevalence of disciplines [2]. This has yielded a list of more present disciplines (computer science, communications, psychology, sociology) and a list of less present disciplines (economics, philosophy, law). Again, these insights can help educators make informed decisions when building curricula.

It can be seen that the outputs of a disciplinary analysis of Web Science support production of curricula by letting educators assess the correspondence between a) highly-ranked topics and expert-identified disciplines from Web Science publications with b) topics and disciplines that are actively taught. Of course, many factors can and should be considered when building a curriculum, but disciplinary presence in the Web Science community is one of a number of valuable inputs that can inform the Web Science curriculum.

3.2 Contribution 2: Teaching Web Science

The outputs of the disciplinary analysis can also be used when teaching students about the nature and history of Web Science. Application contexts and the current disciplinary make-up of the community provide insight into active areas of research and collaboration.

It is possible to examine trends in disciplinary representation over time, facilitating insights into the evolution of Web Science and discussions of possible future trends. For example, the first three years of the Web Science conference show relatively stable acceptance rates (21%, 26% and 15% in 20009, 2010 and 2011 respectively) and term diversity (only a shift from 609 to 708 of the same top 1000 terms being included in the conferences over all three years). Nonetheless, interesting variations were revealed by an analysis of 'peak' terms, terms to occur in five or more publications that 'peak' in a given year (a difference of more than 5 papers in different years) in both papers and posters. 2 peaks were found in 2009, 10 in 2010 and 1 in 2011 [2]:

- 2009: machine learning; real world
- 2010: available online; information exchange; information retrieval; information sharing; natural language; RDF graph; real time; semantic web; share information; SPARQL query
- 2011: social media.

Finally, disciplinary analysis can be applied in other contexts such as Network Science or Internet Science, enabling Web Science students to compare these fields. This can prompt a greater understanding of Web Science and its broader context.

3.3 Contribution 3: A Web Science Method

Finally, disciplinary analysis is a sound Web Science method of its own right, drawing on mixed methods from multiple disciplines. It combines quantitative Computer Science methods (natural language processing), quantitative and qualitative sense-making methods from Network Science (graphing and visualisation), and non-discipline specific qualitative and quantitative methods (expert survey). Indeed, disciplinary analysis has already been taught to Web Science students studying at Master degree level[1].

4. CONCLUSIONS

Looking ahead, the Web Science Education workshop offers an ideal forum to discuss practical uses of disciplinary analysis and the possibility of conducting a more detailed analysis for insight into the history of Web Science.

Web Science Education is a key part of the Web Science community, essential to the healthy growth and continuation of the discipline. By integrating lessons learned from disciplinary analysis – both in terms of informing curricula and teaching students about the make-up of Web Science – it is possible to strengthen the theory and practice of Web Science Education. Last but not least, disciplinary analysis serves as an exemplar of a mixed methods, interdisciplinary method that is relevant to the heart of Web Science.

5. ACKNOWLEDGEMENTS

The research leading to these results received funding from the EU FP7 EINS under grant agreement No 288021.

6. REFERENCES

1 Hooper, C.J., Marie, N., and Kalampokis, E. Dissecting the Butterfly: Representation of Disciplines Publishing at the Web Science Conference Series. In *Proceedings of ACM Web Science 2012* (Chicago, USA 2012), ACM, 137-140.

2 Hooper, Clare J., Bordea, G., and Buitelaar, P. Web Science and the Two (Hundred) Cultures: Representation of Disciplines Publishing in Web Science. In *Proceedings of ACM Web Science 2013* (Paris, France 2013), ACM, 162-171.

3 Monaghan, Fergal, Bordea, Georgeta, Samp, Krystian, and Buitelaar, Paul. Exploring Your Research: Sprinkling some Saffron on Semantic Web Dog Food. In *Semantic Web Challenge at the International Semantic Web Conference* (Shanghai, China 2010).

4 White, Su, Croitoru, Madalina, Bazan, Stéphane et al. Negotiating the Web Science Curriculum through Shared Educational Artefacts. In *ACM Web Science* (Koblenz, Germany 2011), ACM.

[1] COMP6048: Interdisciplinary Thinking (2013-2014) http://www.ecs.soton.ac.uk/module/COMP6048