Empirical Methodologies for Web Engineering

Briony J. Oates, Gary Griffiths, Mike Lockyer, and Barry Hebbron

School of Computing
University of Teesside
Middlesbrough
TS1 3BA, UK
`B.J.Oates@tees.ac.uk`

Abstract. We review a range of data generation methods and empirical research strategies of potential usefulness to web engineering research. The various strategies do not all share the same underlying philosophy about knowledge and how it can be acquired. We therefore explain two contrasting philosophical paradigms: positivism and interpretivism. We suggest that empirical web engineering should use a plurality of research strategies and data generation methods, and recognise the potential usefulness of both positivism and interpretivism. Finally we discuss the implications of such a plurality.

1 Introduction

The majority of web engineering research, like software engineering research, concentrates on design research i.e. developing new concepts, models, methods or instantiations [1]. However, there have been strong criticisms of software engineering researchers for under-usage of empirical studies and failing to validate their research ideas [e.g. 2]. There is therefore increasing interest in empirical software engineering. If web engineers are to avoid similar criticisms, they must be able to both perform empirical studies and also assess the empirical research findings of others. We therefore review a range of empirical strategies and data generation methods of potential usefulness to web engineering research and summarize two contrasting philosophical paradigms: positivism and interpretivism. We argue that both philosophies and all the strategies and methods are relevant to web engineering research.

2 Data Generation Methods and Research Strategies

Data and data analysis can be either quantitative (i.e. numeric), or qualitative (i.e. textual, verbal or visual). Data generation methods available for gathering evidence include questionnaires, interviews, observations and documents (which include multimedia 'documents': non-textual artifacts such as photographs, videos and screenshots). Research strategies are the ways in which data generation methods are used and combined. One strategy can use many data generation methods, although

particular strategies may be associated with particular methods, and typically one research strategy addresses one research question.

Strategies for using and combining data generation methods include experiments, surveys, case studies, action research and ethnography. Table 1 below summarizes each strategy, provides references where more information about each strategy can be found, and gives examples of its use in web-related research.

Table 1. Summary of research strategies for web engineering

Strategy	Brief description	Examples
Experiments	Use observations to look for evidence of cause and effect, so can confirm or refute a hypothesis [3]	[4, 5]
Surveys	Systematic gathering of information from a large sample, looking for general trends or patterns via statistical analysis [6]	[7-9]
Case studies	Rich account of particular experience, event or situation, often longitudinal view [10, 11]	[12, 13]
Action research	Developers research iteratively into own practice, with twin aims of contributing to practical concerns of people in a situation and to the goals of science [14]	[15, 16]
Ethnography	Researchers immerse themselves in lives of the people under study, experience the same as them, and place phenomena they observe in their social and cultural context [17]	[18, 19]

3   Philosophical Paradigms

The strategies in Table 1 are based on different philosophical assumptions about the nature of 'reality' (i.e. ontological assumptions) and about the nature of 'knowledge' and how it can be obtained (i.e. epistemology). These are summarized in Table 2.

Positivism underlies the scientific method, which has been developed by the natural sciences [e.g. 20]. Many people know of only this approach to research, and our modern daily discourse is frequently based, often unthinkingly, on a positivist worldview, with politicians and journalists demanding 'proof' and 'the truth'. Interpretivism [eg. 21] has been developed by the social sciences, and recognises that the social world has few equivalents to the 'laws of nature' in the physical world. For example, there is no guarantee that two people joining in with the life of a web development department as ethnographers would gather the same data and interpret it in the same way to draw the same conclusions.

Table 2. Summary of positivism and interpretivism

	Positivism	**Interpretivism**
Strategies	Experiments and surveys	Ethnography, most action research and many case studies
Ontological assumptions	Physical and social world exists independently of humans; exists 'out there' to be studied, captured and measured. Researcher 'discovers' this world by measurement, modeling and observations.	Whatever 'reality' is, it can only be accessed through social constructions such as language and shared meanings.
Aims	Generalizations – irrefutable objective facts and fundamental laws	Understanding, how people make sense of their perceived worlds, and how those perceptions change over time and differ from one person or group to another
Researcher	Must be neutral, objective detached	Can never be neutral: their assumptions, beliefs, actions inevitably shape research process and affect situation.
Epistemology	Empirical testability of hypotheses and theories, leading to verification or refutation, and a search for universal laws or principles	Studying people and practices in their natural social or work settings
Evaluation criteria	Internal and external validity, reliability and replication	Plausibility and cogency of the reasoning and the evidential data

4 The Need for Plurality

Our literature searches have confirmed the findings of Bahli and Di Tullio [22]: most empirical web engineering research has used surveys or experiments and a positivist perspective. In Section 2 we had to suggest web research examples for some strategies from the information systems, social sciences and education disciplines.

Members of the web engineering community *could* decide that only positivist research is appropriate to their discipline. Or they could decide that an interpretive case study is only appropriate as an exploratory method of investigation prior to a more 'scientific' approach. We suggest, however, that web engineering should accept both positivism and interpretivism and recognise a wide range of research strategies [cf. 23]. Web engineering is dependent on the people and the environment in which it

is practised, making it difficult or impossible to design and implement carefully controlled and repeatable experiments. Where such experiments can be achieved, a positivist approach could provide some truths on which to build the discipline. Where this is not possible, interpretivist approaches such as ethnographies and case studies can help us to explore particular situations and contexts, in order to understand better how people understand, engineer and use web-based artifacts in the real world. Rich and detailed understanding from a series of case studies of web engineering might, but not necessarily, gradually accumulate into a generally applicable body of knowledge.

5   Implications

We are not proposing that everyone has to abandon positivism and adopt interpretivism. But our argument for plurality does mean that researchers and reviewers should not automatically reject a as 'unscientific' work which does not fit the positivist paradigm. On the other hand, researchers and reviewers do not automatically have to accept qualitative, interpretive evidence. They should know enough about the tenets of interpretivism to accept or reject the validity of such evidence on its own terms.

Web engineering practitioners may also need educating in the different types of data generation methods and strategies and underlying philosophies of empirical research, especially if they are to be persuaded to adopt new practices on the basis of qualitative, interpretive evidence. Their own educational background might have only introduced them to the positivist scientific method, and, as we noted earlier, our modern daily discourse is frequently unthinkingly based on a positivist worldview.

By using a wide range of strategies and data generation methods, and both positivist and interpretivist approaches, empirical web engineering research can increase our knowledge and understanding of how to develop, deploy and maintain high quality Web-based systems and applications.

References

[1]   S. March and G. Smith, "Design and natural science research on information technology," *Decision Support Systems*, vol. 15, pp. 251-266, 1995.
[2]   M. V. Zelkowitz and D. Wallace, "Experimental validation in software technology," *Information and Software Technology*, vol. 1997, pp. 11, 1997.
[3]   B. A. Kitchenham, S. L. Pfleeger, L. M. Pickard, P. W. Jones, D. C. Hoaglin, K. El-Emam, and J. Rosenberg, *Preliminary Guidelines for Empirical Research in Software Engineering*: National Research Council of Canada, 2001.
[4]   C. Rumpradit and M. L. Donnell, "Navigational Cues on User Interface Design to Produce Better Information Seeking on the World Wide Web," presented at 32nd Hawaii International Conference on System Sciences (HICSS-32), Hawaii, USA, 1999.
[5]   S. Grazioli and A. Wang, "Looking without seeing: Understanding unsophisticated consumers' success and failure to detect Internet deception," in *Proc. 22nd International Conference on Information Systems*, 2001, pp. 193-203.

[6]   S. L. Pfleeger and B. A. Kitchenham, "Principles of survey research, Part 1: Turning lemons into lemonade," *Software Engineering Notes*, vol. 26, pp. 16-18, 2001.

[7]   M. Lang, "Hypermedia systems development: A comparative study of software engineers and graphic designers," *Communications of the AIS*, vol. 12, pp. 242-257, 2003.

[8]   D. Gehrke and E. Turban, "Determinants of successful web site design: relative importance and recommendations for effectiveness," in *Proc. 32nd Hawaii International Conference on Systems Sciences*, 1999.

[9]   A. S. Huarng, "Web-based information systems requirement analysis," *Information Systems Management, 2003*, vol. 20, pp. 49-57, 2003.

[10]   R. K. Yin, *Case Study Research. Design and Methods*, Third ed. Thousand Oaks, Calif.: Sage Publications, 2003.

[11]   R. K. Yin, *Applications of Case Study Research*, Second ed. Thousand Oaks, Calif.: Sage Publications, 2003.

[12]   E. Mendes, N. Mosley, and C. Steve, "Web metrics - Estimating design and authoring effort," *IEEE Multimedia*, vol. 8, pp. 50-57, 2001.

[13]   C. Hine, "Web pages, authors and audiences. The meaning of a mouse click," *Information, Communication & Society*, vol. 4, pp. 182-198, 2001.

[14]   R. L. Baskerville and A. T. Wood-Harper, "A critical perspective on action research as a method for information systems research," *Journal of Information Technology*, vol. 11, pp. 235-246, 1996.

[15]   R. Vidgen, "Constructing a web information system development methodology," *Information Systems Journal*, vol. 12, pp. 247-261, 2002.

[16]   G. Griffiths and B. J. Oates, "Lecture-free teaching for systems analysis: An action research study," in *Proc. INSITE Informing Science and Information Technology Education conference, 24-27 June*. Pori, Finland, 2003, pp. 355-365.

[17]   J. Van Maanen, *Tales of the Field: On Writing Ethnography*. Chicago: University of Chicago Press, 1988.

[18]   N. Hayes, "Boundless and bounded interactions in the knowledge work process: The role of groupware technologies," *Information and Organization 11 (2001) 79-101*, vol. 11, pp. 79-101, 2001.

[19]   C. Hine, *Virtual Ethnography*. London: Sage Publications, 2000.

[20]   K. Popper, *The Logic of Scientific Enquiry*. London: Harper, 1959.

[21]   N. K. Denzin and Y. S. Lincoln, "Handbook of Qualitative Research." Thousand Oaks, Calif.: Sage Publications, 1994.

[22]   B. Bahli and D. Di Tullio, "Web engineering: An assessment of empirical research," *Communications of the AIS*, vol. 12, pp. 203-222, 2003.

[23]   R. Dawson, P. Bones, B. J. Oates, P. Brereton, M. Azuma, and M. L. Jackson, "Empirical methodologies in software engineering," in *(Under review)*: Springer Verlag, 2004.