# Integrating Mixed-Methods for Evaluating Information Access Systems

Simon Wakeling(✉) and Paul Clough

Information School, University of Sheffield, Sheffield, UK
{s.wakeling,p.d.clough}@sheffield.ac.uk

**Abstract.** The evaluation of information access systems is increasingly making use of multiple evaluation methods. While such studies represent forms of mixed-methods research, they are rarely acknowledged as such. This means that researchers are potentially failing to recognise the challenges and opportunities offered by multi-phase research, particularly in terms of data integration. This paper provides a brief case study of how one framework – Bazely & Kemp's metaphors for integrated analysis – was employed to formalise data integration for a large exploratory evaluation study.

## 1 Introduction

The evaluation of an information access system can take many forms. This can involve both system-oriented and user-oriented approaches, the latter falling within the area of interactive information retrieval and the focus of this paper. Kelly [7] has offered one way of conceptualising information retrieval evaluation studies, suggesting a continuum ranging from system to user focus along which different methods can be placed. In attempting more holistic evaluations of information systems, researchers may utilise multiple methods from varying points on this continuum.

We argue that when taken together, such multi-phase evaluations essentially constitute mixed-methods research, an area that has received intense study within the social sciences. As Fidel [5] notes, such research within the IR field frequently does not self-identify as mixed-methods. While this has implications for the quality and clarity of research design and implementation, this paper focuses on the benefits of adopting a more formalised approach to integrating data from multiple research phases. In practice this means identifying the most appropriate and effective strategies for combining results from different forms of evaluation, and identifying areas of research that will benefit most from this integration. We therefore present a case study describing how data from a mixed-methods evaluation of WorldCat.org were integrated according to a framework developed by Bazeley & Kemp [1].

In the following sections we provide a summary of data integration methods for mixed-methods research (Section 2), introduce a case study based on WorldCat.org (Section 3), then demonstrate integrating multiple methods for studying an information access system (Section 4). Finally, Section 5 concludes the paper.


J. Mothe et al. (Eds.): CLEF 2015, LNCS 9283, pp. 306–311, 2015.
DOI: 10.1007/978-3-319-24027-5_33

## 2 Related Work

While the research literature of many disciplines abounds with discussions of mixed-methods research, relatively little attention has been paid to the theory and practice of integrating the results of mixed-methods research [6]. Several authors have noted the prevalence of published works which claim to present integrated results of mixed-methods research projects, but which either fail to adequately assimilate findings from the attendant methodological strands, or do not properly discuss the techniques employed to achieve integration [2], [9].

Perhaps the most commonly cited theoretical underpinning to mixed-methods integration is *triangulation*. The use of the term as a methodological concept dates back to the 1960s, when Webb et al., building on earlier work by Campbell & Fiske, noted that "the most persuasive evidence comes through a triangulation of measurement processes" [8]. Initially this argument was most usually applied to quantitative forms of research, and therefore closer in spirit to the original meaning of the term as a surveying methodology involving the taking of multiple measurement readings. Later however the concept was popularised as a mixed-methods approach by Denzin [3], who outlined four modes of triangulation: data triangulation (capturing data from diverse subjects at diverse points in time and space), investigator triangulation (the use of more than one researcher to collect data), theory triangulation (utilising multiple theoretical constructs to interpret the data) and methodological triangulation (using different methods to collect data).

**Table 1.** Overview of Bazeley & Kemp's metaphors for integrated analysis [1]

| **Complementary Approaches** | **Description** |
|---|---|
| Completion: Bricolage, Mosaics, and Jigsaws | Constructing a "patchy" aggregate based on the available data, or more carefully amalgamating all findings into a unified whole. |
| Enhancement: Sprinkling and Mixing/Stirring | Augmenting meaning by incorporating small data points, or mingling diverse but complementary findings together. |
| Detailing a More Significant Whole: Triangulation and Archipelago | Revealing unknowns through the combination of known points, or reveal a broader picture through snapshots of evidence. |
| **Generative Approaches** | **Description** |
| Exploration Through Transformation Involving Blending, Morphing, or Fusion of Data Elements | Developing new variables, or otherwise transforming or combining data |
| Conversation and DNA as Iterative Exchange | Re-assessing initial interpretations in light of subsequent findings, and identifying and linking "sense strands" |

Of these by far the most influential in social sciences research is methodological triangulation, where it is frequently cited as a justification for and conceptual underpinning of mixed-methods research. A difficulty arises though in the extension of triangulation to a point where it is cited as a model for integrating qualitative and quantitative data. As Denzin himself has noted [4], this interpretation is somewhat beyond the defined scope of his earlier work. The primary purpose of methodological triangulation is to use multiple data sources as a means of validating findings [6], rather than a method of integrating complementary findings.

Bazeley & Kemp's metaphors for integrative analysis [1] offer an attempt to address this deficit. Their work systematically combines ideas taken from a review of the methodological literature into a set of approaches to data integration, which they express as metaphors. These are presented in Table 1. The result is a framework of methods which the authors encourage researchers to interpret imaginatively. The metaphors describe a set of principles and strategies for integrating data obtained through different methods in order to maximize the analytical potential of multi-phase research projects.

## 3 Case Study: WorldCat.org

Bazely & Kemp's metaphors were used as a framework for the integration of data collected during a multi-phase mixed methods research project investigating the use of WorldCat.org. Managed by OCLC, WorldCat.org is a publically accessible online aggregate catalogue of the holdings of OCLCs member libraries. The project sought to address a series of research questions relating to use of the system, including the extent to which it supported record retrieval in a variety of information-seeking contexts. Data were collected and analysed between April 2011 and September 2014. The four phases of the research can be summarised as follows:

**Phase 1: Focus Groups (Qualitative):** 21 focus groups undertaken with users in four countries. The sessions aimed to explore how and why the system is used, and elicit perspectives on system strengths and weaknesses. Transcripts of the focus groups were analyzed using Qualitative Content Analysis.

**Phase 2: Survey (Quantitative):** A pop-up survey was implemented on WorldCat.org, which generated 2,918 responses. The survey asked respondents about their use of the system and its various features.

**Phase 3: Transaction Log Analysis (Quantitative):** Analysis was performed on log files generated from two months of traffic to WorldCat.org, identifying common patterns of user behavior and use of various system features. Some manual analysis of sample sessions was also undertaken.

**Phase 4: Lab based User Study (Quantitative and Qualitative):** A task-based laboratory study in which participants' interactions with WorldCat.org were measured, and compared to the use of Amazon.co.uk. A post-session interview was also conducted, and this interview data analysed using Qualitative Content Analysis.

## 4 Data Integration

While consideration was paid to the potential methods and breadth of data integration at the planning stages of the project, these initial strategies were revised and improved at each stage of data collection as the quality and scope of data became better understood. By the end of the project, seven instances of integration had been identified and completed, some constituting more than one type. These are represented diagrammatically in **Figure 1**.

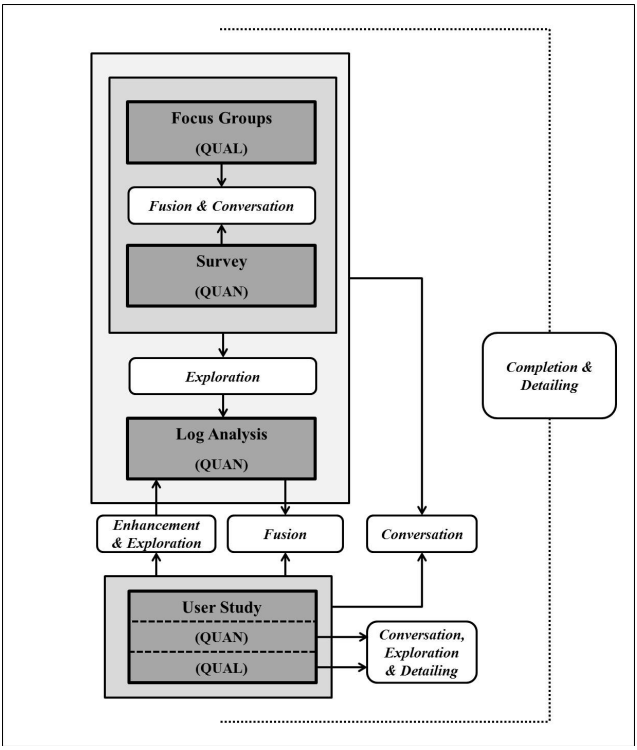


**Fig. 1.** Diagrammatical representation of integration metaphors utilised in the study

The first instance represented the integration of data from the first two phases of the research – the Focus Groups and Survey. Here two of the metaphors were employed, *fusion* and *conversation*. While the focus groups provided a rich source of anecdotal evidence relating to use of the system, they could not provide a reliable means of determining the extent to which participants' use cases were typical, or the degree to which those participants represented larger or smaller user groups. Fusing this data with the survey data allowed for a richer understanding of the relative importance of individuals' perspectives, and the generation of new ideas about the types of tasks that the system need support most. Initial interpretations of the data from each of the first two phases were also re-assessed collectively. In practice this conversational approach

meant both better understanding how respondents were likely to have interpreted questions asked in the survey, and appreciating how users' own limited requirements might have blinded them to beneficial aspects of system functionality for others.

How findings from the first two phases informed the *exploration* of the transaction logs represented the second instance of integrated analysis. Differences in purpose and behavior from users arriving at WorldCat.org via a search engine referral, as opposed to actively seeking out the service, were clearly apparent from survey and focus groups responses, and this proved a fruitful means of approaching the analysis of the logs. Similarly, differences in the reasons for accessing the system were found across different geographical regions. This too informed the approach to the third phase of research, and allowed for a more sophisticated reading of patterns emerging from the logs.

The third occurrence of integrated analysis describes the ways in which results of the user study served to influence understanding of the transaction logs. The rich data collected during the user study for two distinct and clearly defined tasks (one exploratory, one fact-finding) allowed for the creation of unambiguous surrogate log files. These formed the basis for the coding of a sample set of the transaction logs, allowing these samples to be enhanced with codes relating to inferred intent. Furthermore, the computation of *n*-grams and transition probabilities for the user study sessions provided an empirical basis for the *exploration* of the logs according to action sequences. Furthermore, in addition to the user study informing the understanding of the logs, a fourth example of integration comes in the way in which the results of these two phases were *fused* into a whole. This meant obtaining the actual log files for the user study sessions, thereby creating a single data set incorporating both controlled and uncontrolled system interactions.

A fifth instance of integration came in analysis of the user study post-session interviews, and the integration of that interview data with the quantitative data collected from the user study itself. By viewing individual participant interactions in light of the detailed perspectives offered by the interviews, a *conversational* process was enacted which allowed for greater insight into how users behaved and interacted with the system. This was particularly helpful in understanding when users felt satisfied that they had completed the tasks, and in *exploring* how patterns of system interaction represented user frustration. The qualitative data also allowed for a more *detailed* understanding of user engagement with the system, particularly in terms of understanding the extent to which interactions were governed by choice or necessity.

Since the user study represented actual system interaction in a controlled setting, the findings naturally informed the full appreciation of results from the other strands. This sixth example of integration represented a *conversation* between the data collected throughout the first three phases, and the user study findings. The focus groups, survey and log analysis provided a rich picture of who was using the system and why, while the user study investigated how the system supported these tasks. These results could then be viewed holistically to better understand how system functionality influenced perceptions of system performance, and identify ways in which the system could better support its users.

The final instance of analytic integration came at the end of the project. Reviewing data from each phase led to the discovery of *details* beyond the scope of the original project – in this case the differing perspectives of intermediation by librarians and students. The findings from all four phases could also be aggregated into a *complete* whole, with the research questions addressed in a robust and comprehensive manner.

## 5 Conclusions

Evaluations of information access systems that incorporate more than one method can reasonably be viewed as representing mixed-methods research, and offer huge potential for providing a rich understanding of system performance. In order to maximise the potential benefits of these multi-phase studies, we suggest that researchers should attempt to formalise the relationship between each distinct method, and the key modes and strategies of integration. Bazely & Kemp's metaphors for integrated analysis offer a flexible framework for this process which can ensure researchers get the most out of their data.

## References

1. Bazeley, P., Kemp, L.: Mosaics, Triangles, and DNA: Metaphors for Integrated Analysis in Mixed Methods Research. J. Mix. Methods Res. **6**(1), 55–72 (2012)
2. Bryman, A.: Why do researchers integrate/combine/mesh/blend/mix/merge/fuse quantitative and qualitative research? In: Bergman, M.M. (ed.) Advances in Mixed Methods Research, pp. 87–100. SAGE, London (2008)
3. Denzin, N.K.: The research act: A theoretical introduction to sociological methods. McGraw-Hill, New York (1978)
4. Denzin, N.K.: Triangulation 2.0. J. Mix. Methods Res. **6**(2), 80–88 (2012)
5. Fidel, R.: Are we there yet?: Mixed methods research in library and information science. Library & Information Science Research **30**, 265–272 (2008)
6. Greene, J.C.: Mixed methods in social inquiry. Jossey-Bass, San Francisco (2007)
7. Kelly, D.: Methods for Evaluating Interactive Information Retrieval Systems with Users. Foundations and Trends in Information Retrieval **3**, 1–224 (2009)
8. Webb, E.J., Campbell, D.T., Schwartz, R.D., Sechrest, L.: Unobtrusive measures: Nonreactive research in the social sciences. Rand McNally, Chicago (1966)
9. Woolley, C.M.: Meeting the Mixed Methods Challenge of Integration in a Sociological Study of Structure and Agency. J. Mix. Methods Res. **3**(7), 7–25 (2009)