http://dx.doi.org/10.14236/ewic/HCI2018.140

Evaluation of a European Data Analytic Framework

Andrew Boilson
Dublin City University
School of Nursing & Human Sciences
Glasnevin, Dublin 9
andrew.boilson@dcu.ie

Dale Weston
Emergency Response Department
Science & Technology
Public Health England
Porton Down, Salisbury, SP4 0JG
Dale.Weston@phe.gov.uk

Regina Connolly
Dublin City University
Business School
Glasnevin, Dublin 9
regina.connolly@dcu.ie

The multi-national MIDAS (Meaningful Integration of Data Analytics and Services) project is developing a big data platform to facilitate the utilisation of a wide range of health and social care data. The platform will enable the integration of heterogeneous data sources, providing privacy-preserving analytics, forecasting tools and bespoke visualisations of actionable epidemiological data. An evaluation framework starting with a logic model and using the principles of realist evaluation developed working with users, and software developers. The tools used are a series of parallel case studies to address the requirements of stakeholder groups at critical time points during the project to ensure IT systems development is in line with user's requirements. The process includes longitudinal interviews with stakeholders, regular feedback to users and developers, and measurement of stakeholder's attitudes to the project using Q-methodology.

Visual & Data Analytics, Public Health Informatics, Realist Evaluation

1. LITERATURE REVIEW

Health care, like many modern activities, generates very large amounts of data, a proportion of which is stored, in some accessible form as usable information, but rather less of which used to guide practice, planning or policy (Murdoch & Detsky, 2013). Information communications technology (ICT) is a key tool to support this, and to assist with effective decision-making. The need for effective use of data is particularly critical in public health organizations, where it is required to support areas such as epidemiologic surveillance, health outcome assessment, program evaluation and performance measurement, public health planning, and policy analysis (Studnicki et al, 2008).

To take appropriate actions, health policymakers require many different kinds of information. The knowledge translation literature contains many studies on information synthesis methods for producing best available evidence.

However, less attention paid to methods of disseminating such epidemiological data to policymakers (Zakkar & Sedig, 2017). To satisfy this need, more flexible health data representation, analysis, querying, and visualization methods (analytic software tools) are desirable (Tilahun et al, 2014).

These tools can support decision making in several ways. Reducing the cognitive effort required, by structuring the decision, and highlighting factors that merit consideration; making the information more evaluable, actively engaging well-developed human visual capabilities; providing information in a non-linear format, to facilitate its incorporation in decision making deliberations; all to help policy users process and understand information (Chen et al, 2010; Reyna et al, 2008). They serve as cognitive aids in problem solving for policy makers, as users come to rely on these systems to help them solve increasingly difficult problems. Support wider and easier access to health system and epidemiological information has the potential to increase efficiency, reduce errors, and alter professional roles and responsibilities in a manner, which allows improvement in the delivery of patient care (Fonkych et al, 2005).

However, the systematic, disciplined and quantitative standpoint of most software engineering development models does not take into account all the dimensions of software development, in particular the organizational, economic, and human dimensions (Toffolon, 2000; Ilavarsan et al, 2003). Indeed, users differ greatly in experience, and professional background, yet visualization tools and other software platforms seldom designed for a single ideal user (Ziemkiewicz et al, 2012).

There is no silver bullet, and organizations solve software engineering problems they encounter by building satisficing solutions (Brooks, 1995). The effectiveness of knowledge integration in a software system determines the quality of this system i.e. how well it supports business, support, and the decision-making processes within organizations. The gap lies in the difference between the knowledge integrated in a software system and the knowledge owned by the organizational actors who are to use these systems, while carrying out their other activities (Dakhli & Chouikha, 2009). This knowledge gap is the commonest reason for the rejection of a software system by the intended users.

It is therefore critically important to ensure that a thorough evaluation is conducted throughout the development process in order to minimise the potential for software rejection. A rigorous evaluation of information and communications technology is of great importance for policy makers and end users of the technology (Kaplin et al, 2002).

Health informatics evaluations provide an objective measurement of processes and outcomes against expectations, with the intention of identifying strengths and successes, whilst finding means of addressing and improving weaknesses or even failures (Rigby, 2006). There are no gold standard frameworks of evaluation theory and practice (Yen et al, 2017; Rahimi & Vimarlund, 2007). Issues that cause problems for those undertaking these evaluations are partly due to the complexity of the systems,

and partly to the challenge of selecting an appropriate framework (Friedman & Wyatt, 2005; Ammenwerth et al, 2003).

Our process is a health informatics evaluation, designed through realist evaluation, combining logic models and semi structured interviews in a novel way. This is a theoretically grounded evaluation process, which recognises the intrinsic complexity of program development, and applies a structured approach to dealing with this (Pawson and Tilley, 1997).

That is, it seeks to answer basic questions of why (objective of evaluation), who (which stakeholders' perspective are going to be evaluated), when (which phase in the system development life cycle to target), what (aspects or focus of the evaluation) and how (methods of performing the evaluation) (Yusof et al, 2008). It is growing in popularity for health service evaluation (Marchal et al, 2012).These time points will include - pre-post implementation of the demonstration projects, post training sessions, and when end users commence using the software to analyse local and national public health data sets. The findings of this study will expand understanding of these knowledge gaps, which have thus far only been explored for clinical IT systems, but not for users of health policy and public health IT systems.

2. METHODS

2.1 Context of the Study

The objective of this study is to obtain insights into acceptance and use of technology using logic a model and semi structured interviews with stakeholders (software developers and potential end users) for the evaluation of a multi-national data analytics project. The MIDAS platform will enable the integration of heterogeneous data sources, providing privacy-preserving analytics, forecasting tools and bespoke visualisations of actionable epidemiological data (Rankin et al, 2017).

2.2 Experimental Design

Longitudinal semi structured interviews are being performed on three occasions over the duration of the project. This involves stakeholders utilising a novel parallel case study design. The data collection process was developed based on a logic model and uses semi structured interviews with developers and end users to evaluate health analytic software acceptance use gaps at critical time points throughout the duration of the study.

2.3 Semi Structured Interviews

In this first round of technology pre-implementation interviews, the MIDAS platform end users, policy makers, and lead developers have been interviewed in person and via conference software. The same stakeholders are interviewed at several critical time points through the project. For users, the primary objective is to identify their expectations of the software and its uses for the purpose of effective public health decision making and policy formation.

For developers, the focus is on their understanding, evolving over time, of end users' expectations of the software platform. Prior to undertaking the interviews, stakeholders provided with the general themes of the interview questions to assist them to consider their answers in advance. Two topic guides have been developed, one for software developers, and one for end users of the software. Each phase of interviews with stakeholders will inform the next round of interviews as a means of identifying gaps between their expectations of the platform, achievement of the logic model outcomes, impacts and acceptance and use of the technology.

The semi-structured interviews are 30-40 minutes' duration, are recorded with the consent of stakeholders and transcribed verbatim. On completion of each round of interviews, the stakeholders provided with a copy of their transcript for review.

2.4 Transcript Coding

The transcript coding was based on the framework approach to qualitative data analysis (Ritchie & Lewis, 2003) guided by a logic model, which will be refined over time as the interviews proceed. The logic model is a representation of how a program is intended to work and links outcomes with processes and the theoretical assumptions of the MIDAS project. These models are typically diagrams or flow charts that convey relationships between contextual factors, inputs, processes and outcomes (Kellogg Foundation, 2004). They provide a structure for exploring the complex relationships between public health practice and outcomes.

The initial logic model was developed based on the grant agreement and deliverables of the MIDAS project. The model was circulated to the stakeholders several times and refined over a period of six months.

The final model served as a guide for designing the study protocol and semi structured interview framework. The focus was to identify how impacts could be measured for each component of the project. Using the logic model as basis for the coding facilitates constant refinement of themes to aid the development of a clear conceptual framework (Ritchie & Lewis, 2003).

The logic model's outputs, outcomes and impacts provide a good reference point coding framework for evaluating the progress and expectations of both developers and software end users. The themes and categories identified using this coding approach from the logic model in conjunction with Venkatesh et al (2016) multi-level framework will form the basis for developing coherent theories for evaluating acceptance and use of technology for the duration of the project.

3. PRELIMINARY RESULTS

During this pre-implementation phase of this project, software developers have been liaising with potential end users over a period of approximately twelve months. Developers have ascertained potential user's requirements for the visualisation and data analytic tools through online meets, workshops, and mock demonstrations of completed components of the software utilising dummy data.

The interview process with stakeholders focused on discussing issues relating to policy maker's / users expectations, utilisation, priorities, potential benefits and anticipated problems with the software.

The same issues were addressed to the software developers to determine if they understood the potential user's expectations of the platform. The stakeholders (developers / potential users) were given considerable flexibility during the interview process to enable them to identify the issues of importance to them.

The preliminary results of the interviews highlight the importance of individual level contextual factors. Stakeholder's feedback relevant to the main effects of this framework of technology acceptance model will be explored in subsequent rounds of interviews within the context of the logic model outputs, outcomes and impacts.

Seven transcripts (37%) from the first phase of interviews was evaluated. There was general agreement between developers of end users expectations and requirements of the MIDAS platform and end users understanding of the MIDAS tools, completed at this point of the project and a shared understanding of the projects outputs, expected outcomes, and anticipated impacts and themes relevant to technology acceptance and use.

Software developers emphasized the final data analytic platform should simplify routine data evaluation processes, providing end users with specified tools to undertake disease mapping and forecasting which would enable capacity building and assist in policy decision making. The platform should simplify the process of augmentation of existing datasets for more effective data collection strategies, thereby reducing fragmented healthcare data collated and stored in silos. In the intermediate and long run, the platform should demonstrate real economic value, provide access to historical and current end user epidemiological data to drive policy decision making.

As end users of the MIDAS platform become more familiar with the platform cross-pollination of evidence generation and sharing outcomes with their peers both regionally and nationally would improve health policy decision making throughout the European Union. Since the commencement of the project, software developers have not reported any barriers related to designing and programming the platform. The primary issues that have been expressed by stakeholders (developers or end users) have been in relation to constraints (barriers) related to processing and signing of legal agreements that they found time consuming, the completion of various components of the data

analytic platform in accordance to agreed timelines, and potential issues around resources for implementing the platform.

Potential users overall demonstrated they had a very clear idea of how the platform's tools will enhance their needs and productivity within the arena of public health research. Potential users required tools to use augmented real time data to resolve knowledge gaps, enhance policy maker engagement with an emphasis on disease prevention in the arena of diabetes and obesity through cross collaboration with peers, sharing resources and knowledge.
Potential end users expressed hope (but not an expectation) for policy makers to use resources which would eventually be available with the MIDAS tools to have a positive impact on policy making, health practice and service provision.

This process has been achieved through close stakeholder engagement (i.e. conferences workshops, undertaking analytic engine demos to illustrate the value of tools designed so far, fortnightly technical and non-technical meeting with the consortium.

4. CONCLUSION & FUTURE WORK

The design of data analytic platforms, training and information sessions need to focus on ensuring these systems improve medium and long-term outcomes. Each phase of the interviews will inform subsequent evaluation phases facilitating better understanding of health IT use and acceptance, contributing to theory building efforts (Hennington et al, 2007). The MIDAS pre implementation interviews proved very useful for providing feedback to software developers from policy makers and potential end users of the requirements of MIDAS platform. The second phase of interviews will be performed when end users and policy makers start working with the various software components of the platform, and have sufficient knowledge of the system so feedback can be provided to software developers.

5. REFERENCES

Ammenwerth, E., Gräber, S., Herrmann, G., Bürkle, T., & König, J. (2003). Evaluation of health information systems—problems and challenges. *International Journal of Medical Informatics*, 71(2), 125–135.

Brooks, F. P., Jr. (1995). *The Mythical Man-Month* (Anniversary Ed.). Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

Chen, C. (2010). Information visualization. *Wiley Interdisciplinary Reviews*: Computational Statistics, 2(4), 387–403.

Dakhli, S. B. D., & Chouikha, M. B. (2009). The knowledge-gap reduction in software engineering. Research Challenges in Information Science, 2009. RCIS 2009. *Third International Conference on. IEEE,* 287–294.

Fonkych, K., & Taylor, R. (2005). *The state and pattern of health information technology adoption*. Santa Monica, California, USA: Rand Corporation.

Friedman, C. P., & Wyatt, J. (2005). *Evaluation methods in biomedical informatics.* New York: Springer Science & Business Media.

Hennington AH, Janz BD (2007) Information Systems and Healthcare: Physician adoption of electronic medical records: Applying the UTAUT model in a healthcare context. *Communications of the Association for Information Systems,* vol. 19, pp.60–80.

Ilavarasan, P. V., & Sharma, A. K. (2003). Is software work routinized? Some empirical observations from Indian software industry. *Journal of Systems and Software*, 66(1), 1–6.

Kaplan, B., & Shaw, N. T. (2002). People, Organizational, and Social Issues: Evaluation as an exemplar. In R. Haux & C. Kulikowski (Eds.), *IMIA Yearbook,* 91–102.

Kellogg, Foundation (2004) *Logic model development guide.* Michigan: *WK Kellogg Foundation*, pp.71.

Marchal, B., van Belle, S., van Olmen, J., Hoerée, T., & Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, 18(2), 192–212.

Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. *Jama,* 309(13), 1351–1352.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London; Thousand Oaks, Calif.: Sage.

Rahimi, B., & Vimarlund, V. (2007). Methods to evaluate health information systems in healthcare settings: a literature review. *Journal of Medical Systems*, 31(5), 397–432.

Rankin, D; Black, M; Wallace, J; Mulvenna, M, Bond, R; Cleland, B (2017). *The MIDAS Platform: Facilitating the Utilisation of Healthcare Big Data in Northern Ireland and Beyond*. 8th *Annual Translational Medicine Conference*, City Hotel, Derry/Londonderry, Northern Ireland. Clinical Translational Research and Innovation Centre (C-TRIC).

Reyna, V. F. (2008). Theories of medical decision-making and health: an evidence-based approach. *Medical Decision Making*, 28(6), 829–833.

Ritchie J., & Lewis, J. (2003). *Qualitative research practice: a guide for social science students and researchers*. London: Sage.

Rigby, M. (2006). Evaluation–the Cinderella science of ICT in health. *IMIA Yearbook,* 114–120.

Studnicki, J., Berndt, D. J., & Fisher, J. W. (2008). *Using information systems for public health administration.* In G. Benjamin (Ed.), Public Health Administration: Principles for Population-Based Management. (2nd ed.). Sudbury, MA: Jones and Bartlett, 353–380.

Tilahun, B., Kauppinen, T., Keßler, C., & Fritz, F. (2014). Design and Development of a Linked Open Databased Health Information Representation and Visualization System: Potentials and Preliminary Evaluation. *JMIR Medical Informatics*, 2(2).

Toffolon, C. (2000). *The software dimensions' theory.* In J. Filipe (Ed.), Enterprise Information Systems. Springer, 89–98.

Venkatesh V (2016) Unified Theory of Technology Acceptance and Use of Technology - A Synthesis and the Road Ahead, *Journal of the Association for Information Systems,* vol. 17, no. 5 pp. 328–376.

Yen, P.-Y., McAlearney, A. S., Sieck, C. J., Hefner, J. L., & Huerta, T. R. (2017). Health Information Technology (HIT) Adaptation: Refocusing on the Journey to Successful HIT Implementation. *JMIR Medical Informatics*, 5(3) e28.

Yusof, M. M., Papazafeiropoulou, A., Paul, R. J., & Stergioulas, L. K. (2008). Investigating evaluation frameworks for health information systems. *International Journal of Medical Informatics,* 77(6), 377–385.

Zakkar, M., & Sedig, K. (2017). Interactive visualization of public health indicators to support policymaking: An exploratory study. *Online Journal of Public Health Informatics,* 9(2) e190.

Ziemkiewicz, C., Ottley, A., Crouser, R. J., Chauncey, K., Su, S. L., & Chang, R. (2012). Understanding visualization by understanding individual users. *IEEE Computer Graphics and Applications*, 32(6), 88–94.