

Stakeholders and their respective costs-benefits in IR evaluation

Cécile L. Paris¹, Nathalie F. Colineau¹,
and Paul Thomas²

CSIRO ICT Centre
¹ PO Box 76, Epping NSW 1710, Australia
²GPO Box 664, Canberra ACT 2602, Australia

firstname.lastname@csiro.au

Ross G. Wilkinson

Australian National Data Service
700 Blackburn Rd.
Clayton, VIC 3700, Australia

ross.wilkinson@ands.org.au

ABSTRACT

Evaluations in Information Retrieval are dominated by measures of precision and recall. Is that enough? Probably not, as it somewhat assumes that all information seeking tasks are equal, and that everyone needs the same thing. In this position paper, we advocate a consumers' guide to systems that aim at supporting information seeking tasks. We propose a method that provides guidance in whole-of-system evaluations, explicitly considering all participants and both sides of the "bang for buck" equation.

1. INTRODUCTION

Evaluations of search engines have mainly focused on measuring their accuracy and completeness in returning relevant information, using metrics such as recall and precision. While important, these measures constitute in our view only a partial view of evaluation. First, accuracy and completeness are only one way to measure a system's impact on the user. Then, there are typically a number of stakeholders involved in any system aimed at supporting information seeking tasks, and we believe an evaluation may need to consider the goals of participants besides than the end-user, who is only one of the stakeholders. Finally, we argue that an evaluation should look at costs as well as benefits, for all parties involved.

Information Science also has a long tradition of evaluation: often taking a wider view, looking at a variety of factors such as the system quality (in terms of response time or data accuracy, for example), user satisfaction, individual impact and, interestingly, organisational impact (asking, for example, questions of cost, investment, return on investment, and productivity). Delone and McLean (1992) attempted to consolidate the work on evaluation in this field, and they introduced a comprehensive taxonomy with six major dimensions, placing previous work within that taxonomy (See also <http://business.clemson.edu/IES/>). In their work, there is a recognition that both benefits and costs have to be taken into account to decide on the success of an information system.

Inspired by Delone and McLean's work and drawing from our own attempts both to evaluate systems and to choose an appropriate approach for a specific situation, we propose a method that provides guidance in whole-of-system evaluations, explicitly considering all participants and both sides of the "bang

for buck" equation. The method we propose is akin to having to write a consumer's guide to a system.

In any consumer report, products are described with a set of attributes and evaluated along a variety of dimensions. These enable consumers to understand, compare and choose, given their own circumstances. A product appropriate for one person might not be appropriate for another. For example, a small two-door car might be appropriate for a single person, but not for a large family. There may also be preferences for some dimensions. For example, someone may put comfort over speed, while another individual will do the reverse. Or, there might be several concerns within the same family, with one member preferring one attribute and another member another feature. Finally, all benefits have to be balanced with costs: while someone might want a sport car, and that is absolutely their preference and desire, they might not want to pay the price it costs and will fall back on something they can afford.

The point here is that there is no such thing as one-size-fits-all, that benefits have to be considered in the context of costs, and that there might be more than one stakeholder to consider. Likewise, we argue that systems that support information seeking tasks must be evaluated along a number of dimensions. This view of evaluation is consistent with ISO 9000, a family of standards for quality management systems—and in particular ISO 9126, developed for software evaluation, which already accounts for attributes of a system such as reliability, usability, efficiency and maintainability. Finally, benefits must be balanced against costs, enabling people to choose what systems best suit their purposes, given their stakeholders.

We believe that one of the compelling attributes of our method is to allow researchers to characterise their system in terms of its strengths and weaknesses, its benefits, costs and impact on all affected stakeholders. Our method provides guidance to think explicitly about the different stakeholders involved in the construction, deployment, maintenance, funding and use of a system.

2. THE METHOD

Typically, a system that supports information seeking tasks involves different actors who have different goals. An evaluation must thus consider all the participants. We have identified four main participant roles:

- The *information seeker*, traditionally the end-user or consumer of the services offered by the system;

- The *information provider*, responsible for the content to be searched, explored and delivered;
- The *information intermediaries*. They can be categorised into two groups: resource builders and exploration partners;
- The *system provider*, responsible for the development and maintenance of the technology.

We realise that not all these roles are appropriate in all situations. For example, general search engines might not want to take into account the goals of all the information providers (i.e., anyone wishing to put content on the web). An enterprise search engine might, however, care about the goals of the enterprise. We believe it is important to think explicitly as to who the stakeholders are.

The costs and benefits of a system are likely to differ for each participant. The main benefits for the information seekers are related to the task effectiveness and their satisfaction in using the system. Their costs relate to the time needed to complete the task, the amount of effort required (i.e., the cognitive load) and, potentially, the necessary learning curve.

For the information provider, the benefits concern mostly the audience targeted – to what extent does the information reach a wide or desired audience? The costs here are the costs of providing the information in a form required by the system.

For the information intermediaries, we consider separately the resource builders from the exploration partners. The resource builders are responsible for creating the appropriate set of required resources (e.g., ontologies). Their benefits can be measured in terms of how easy it is to create the required resources, and their costs are related to the time needed to create them, include them in the system and maintain them if required. For the exploration partners, the benefits include those of the information seekers, i.e., related to the task performance and the quality of search and exploration support. Their costs include the time spent in capturing the information relevant to the information seekers' situation.

Finally, the benefits for the system provider are related to the system usage, with its possible corresponding revenue or corporate value, while costs are the cost of system implementation, maintenance and integration with other systems.

This explicit identification of what might constitute a benefit and a cost for whom (see Table 1) can guide researchers and developers in asking appropriate questions about a system and in identifying the relevant evaluation studies to conduct. This in turn helps understand where the technology fits in a larger picture and evaluate different approaches, characterising their strengths and weaknesses, thus allowing one to choose the approach (or system) best suited to one's needs. It also often becomes apparent that providing a benefit to one participant usually comes at a cost (sometimes to another participant). This is the key "bang for buck" equation. This can raise questions such as: to what extent can we trade the benefits of improved user experience with data and system provision costs?

3. CONCLUSIONS

We have briefly presented an evaluation method aiming at guiding researchers in evaluating their web-based information system, looking at benefits *and* costs for *all* participants. Our cost-benefit method provides the means to evaluate different approaches or systems to make an informed decision as to which costs we are willing to pay to obtain which benefits. We believe that our method also enables the framing of research questions that may not be immediately obvious otherwise. The interested reader is referred to Wu et al., (2009) and Paris et al., (2009) for case studies of this method.

4. REFERENCES

DeLone, W. H. and McLean, E. R. (1992). Information Systems Success: The Quest for the Dependent Variable. In *Information Systems Research*, 3(1):60-96.

Paris, C., Colineau, N. and Wilkinson, R. Evaluating Web-Based Information Systems from a Cost-Benefit Perspective: A Case Study. CSIRO ICT Centre Technical Report 09/165. 2009.

Wu, M., Thom, J., Turpin, A., and Wilkinson, R., *Cost and Benefit Analysis of Mediated Enterprise Search*, Joint Conference on Digital Libraries, June 15-19, 2009, Austin, Texas.

Table 1. Cost-Benefit Assessment Method: identifying all participants, their benefits and costs

Participant	Information Seeker	Information Provider	Information Intermediaries	System Provider
Benefits	Task effectiveness Knowledge gained Accuracy of exploration Satisfaction	Audience reach Audience accuracy Message accuracy	<u>Resource builders:</u> Ease of knowledge creation & context modelling <u>Exploration partners:</u> Task effectiveness	System usage Reliability Response time Correctness
Costs	Time to complete task Cognitive load Learning time	Metadata provision Structured information Currency of Data	<u>Resource builders:</u> Time to create and integrate the resource <u>Exploration partners:</u> Time to capture contextual factors	Implementation hardware & software cost Syst. maintenance Syst. integration