Chapter 9
Building for Success?
Evaluating Digital Libraries in the Cultural Heritage Domain

Vivien Petras
Humboldt-Universität zu Berlin, Germany

Juliane Stiller
Humboldt-Universität zu Berlin, Germany

Maria Gäde
Humboldt-Universität zu Berlin, Germany

ABSTRACT
In the cultural heritage field, heterogeneous materials and multicultural, multilingual user groups and their varying needs pose a challenge for information system design and its evaluation alike. Cultural heritage information systems can be evaluated from a system-centric or a user-centric perspective. This chapter discusses evaluation methods in digital libraries with a particular focus on cultural heritage collections and their distinctive features, interaction patterns, and challenges. It describes state-of-the-art evaluation methods illustrating them with examples from Europeana, the European portal for access to digital library, museum, and archive collections, and other projects from within the cultural heritage domain.

INTRODUCTION
The announcement of Google to digitize several big university library collections in 2005 triggered a movement in Europe to create a counterbalance to the commercializing of cultural heritage. The idea of Europeana, the European portal for access to digital library, museum, and archive collections was born (Purday, 2009). Since 2009, this European cultural heritage portal provides a single access point to digital and digitized content collected by galleries, libraries, (audiovisual) archives, and museums. The content providers differ in scope, reach, and in user groups they

DOI: 10.4018/978-1-4666-2991-2.ch009
Users differ in culture, language, access, and interaction requirements (Cousins, 2006). Bridging the gap between these different contexts and requirements is one of the biggest challenges for digital libraries. Evaluating these efforts and assessing when a library is “successful” is equally demanding.

Heterogeneous materials and their descriptions, multicultural and multilingual user groups and their varying needs pose a challenge for information system design and its evaluation alike. Traditional and innovative patterns (search, browse, comment) for user-system interactions cater to different user goals. Different approaches for testing the functionality and quality of individual system features, with respect to both the target audience and the target results, are necessary. Cultural heritage information systems (like all digital libraries) can be evaluated from a system-centric or a user-centric perspective. System-centric approaches have a long-standing tradition in information retrieval research and provide comparable quantitative measures for search effectiveness. However, their results are based on fixed assumptions about user-system interactions. User-centric approaches are more flexible in their assumptions. Commonly qualitative in nature, they provide deeper insights into user motivations and problems, but are also more labor-intensive and do not necessarily provide representative or comparable results for system evaluation. Some new system functionalities like personalization or collaborative features do not fit traditional use cases or evaluation set-ups. Similar to the design or implementation challenges for these features (e.g., user interface design, privacy, or data quality issues), new evaluation methods need to be developed and tested.

This chapter discusses evaluation methods in digital libraries with a particular focus on cultural heritage collections and their distinctive features and challenges. It describes state-of-the-art evaluation methods illustrating them with examples from Europeana and other projects from within the field. The chapter is organized as follows: the first section introduces general conceptual frameworks for evaluating digital library systems. The next section describes cultural heritage information systems, common system components and interaction patterns and their particular challenges. Four case study projects in the cultural heritage domain are introduced. The next sections describe first system-centric, then user-centric evaluation approaches and experimental set-ups from these case studies. The chapter concludes by discussing the challenges arising from innovative system features and interaction patterns and constraints posed by methods and project administration.

FRAMEWORKS FOR EVALUATING DIGITAL LIBRARY SYSTEMS

All evaluation efforts and experiments have to make decisions with respect to the evaluation object (elements/components to be evaluated), the evaluation aims and objectives, the evaluation context (viewpoint/framework for evaluation), the evaluation criteria and measures, and the evaluation methodology (for data collection and analysis) (Saracevic, 2000). In the research literature, reports on practical implementations of systems and their subsequent evaluations have been accompanied by discussions about general concepts, models, and frameworks that summarize work on a more theoretical level. Some of these general frameworks are described here.

Saracevic (2000, 2004) distinguishes seven approaches for digital library evaluation that address different goals or aspects:

- **The Systems-Centered Approach:** Performance, effectiveness, and efficiency.
- **The Human-Centered Approach:** Behavior, information needs, and requirements.