Digital libraries - creating information space excellence: is it already time for benchmarking?

Radovan Vrana, M.Sc.
Department for information sciences
Faculty of philosophy Zagreb
rvrana@ffzg.hr

Abstract

As they are gradually becoming mature, digital libraries are being put more frequently under a magnifying glass in order to be evaluated. Some of the world’s leading information scientist like T. Saracevic (Rutgers University, USA) claims that it is too early to evaluate quality of existing digital libraries as they are not yet taken their final shape and there is much to be expected from such young and important information institutions. The idea of evaluation of services of digital libraries has come from scientists and professionals measuring quality of the existing library services. Some scientists also consider problem of persistence of digital libraries as an their information sources as a prerequisite for design of evaluation tools for measuring various aspects of digital library services. Literature review in this field also shows major concerns for comparison of digital library content representation and categorization, advanced tools and processes in traditional libraries and their counterparts in digital libraries. Aim of this paper is to give a view of current perspectives in this area.

Introduction

In the first decade of their existence, digital libraries have made a solid start giving a promise about well-organized and accessible resources and user profiled services. It is often said and written that information technology holds the potential for the improvement of the quality of life and work. However, connecting people and vast amount of information objects in digital resources is not always a flawless process. Therefore, digital libraries and information scientist have a great perspective in building most advanced information resources in human history, but they also have a big responsibility in managing different types of resources.

Two topics often discussed in literature and conferences related to the development of digital libraries are digital collections and services. Since digital libraries are the next evolutionary step in the development of libraries, it is expected that it will be soon be possible to determine whether existing digital collections and accompanying services have really become pillars for digital library evaluation process. At the present moment, this may be a problem, since there is a lack of standard metrics across systems, which would give precise results. Having this in mind, this paper will try to give a perspective regarding issues surrounding evaluation of digital collections and services.

Understanding framework for the creation of evaluation tools for digital libraries

Development of modern information institutions such as digital libraries requires a broader conceptual overview of changes influencing their maturing process. This overview introduces conceptual differences to the approach of development processes of resources and services in traditional and modern automated libraries in contrast to the development process of digital resources and services in digital libraries. It also shows influence of new modes of communication already present in libraries as well as many new types of interactions. There are also some shifts in theoretical concepts or rather in paradigm of library and information services:

<table>
<thead>
<tr>
<th>Modern libraries</th>
<th>Postmodern libraries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed, permanent, formatted text collections</td>
<td>Fluid and transient multimedia resources</td>
</tr>
<tr>
<td>Static library facilities with fixed stacks</td>
<td>Free, flexible and virtual information spaces</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Uniform sources, citations, references</th>
<th>Customized annotations/transient works</th>
</tr>
</thead>
<tbody>
<tr>
<td>Services provided to individual readers</td>
<td>Tailored services to collaborative teams</td>
</tr>
<tr>
<td>Standard reference services</td>
<td>Personal consulting and analysis</td>
</tr>
<tr>
<td>Professionally provided services</td>
<td>Integrated service provision</td>
</tr>
<tr>
<td>Locally owned permanent collections</td>
<td>Holistic, integrated networked systems</td>
</tr>
<tr>
<td>Centralized collections and services</td>
<td>Distributed, decentralized global access</td>
</tr>
<tr>
<td>Hierarchical organizational structures</td>
<td>Participative and collegial relationships</td>
</tr>
<tr>
<td>Discipline specialization</td>
<td>Inter-, multi-, cross-disciplinary studies</td>
</tr>
<tr>
<td>Generic user service offerings</td>
<td>User-/use-specific relevant services</td>
</tr>
<tr>
<td>Formal publication acquisition</td>
<td>Integration of informal with formal</td>
</tr>
</tbody>
</table>

Apparently, postmodern libraries are quite different in their nature from their modern counterparts. Traditional libraries with their collections based on formatted text evolved into modern libraries, which took advantage of the process of automation by intensively introducing computers and telecommunication devices to libraries in the second half of the 20th century. Even so, modern libraries still have their collection mostly in formatted text but offer information technology based services.

Even the postmodern libraries such as digital libraries are still depending on printed text collections, which are now combined with new digital material. Despite the differences between modern and postmodern libraries, this example of library collections illustrates that no transition between two generations of crucial information institutions can happen suddenly and over night. Previous experiences in the development and management of collections and services in traditional libraries are often being considered for implementation under new and usually improved conditions. This also applies to standards, policies and best practices necessary for flawless library operation.

Traditional and modern libraries have developed numerous quantitative and qualitative evaluation tools and methods in order to evaluate their holdings and services. They are measuring quality of patron’s benefit from library services and their satisfaction with specific services. Most of them are rooted in stability of library as an institution. Stability has enabled information scientists to develop adequate measurement tools and to implement them in order to observe quality trends in library collections and services and to suggest necessary improvements. Persistence and stability are the most desired qualities of digital libraries. If solved, they will boost the development of digital library services based on features of postmodern library.

Process of determining the level of quality of a particular resource considered for inclusion in library collection is a function of the desired depth of collection. The depth of the collection denotes the level of inclusion of different quantity of material into a library collection. The strength of a library collection (and library holdings) is therefore determined by number of items included in the collection.

The more items (presumably paper-based for the modern libraries, and digital for postmodern libraries) included in the library collection the stronger the collection. One of the ways to measure the quality of a paper-based collection is its use. The frequency of use (circulation) of library items (books etc.) can show positive sides (popular items in collection) as well as its weaknesses (least used items). This and other values are called library performance indicators and are covered by several ISO standards.

Use of digital material included in digital libraries can’t be measured in the same way as paper-based material in traditional library collections. Some issues surrounding digital library use have already been mentioned. Inclusion of information technology introduced a need for different approach in measurement of library performance. A popular approach for measurement of use of digital material is machine generated usage statistics. For instance, a digital library system, which requires identity authentication, is typically able to collect personal data about use of digital resources, which, later, can be used for analysis of user behavior during the access to the system. Usage logs are very popular on the Internet but they do not show motives or incentives users might have when they decide to use digital library material.

In both cases, librarians are very keen on using user survey or questionnaires (as well as interviews, focus groups, demographic studies, user commentaries etc.). Questionnaire is a popular method, which can help understand what drives users to use library material. Traditional and modern libraries use paper-based questionnaires distributed to users in libraries and real and potential users outside libraries. Digital libraries are

---

usually left to use the electronic equivalents of questionnaires in aspiration to find out more about their users and why they use certain digital collections. As it was said in the introduction, some information scientists think it is too early to employ evaluation methods on digital collections in existing digital libraries. If this is true, then there is another possibility to impose quality before users start using digital libraries more intensively.

This possibility is related primarily to the development of digital collections. If a digital collection is being developed to a certain standard, its items can be compared and evaluated across formats and easily exchanged with other collections. Digital collections are usually planned and developed by assistance of project handbooks, technical guides, non-technical articles, guidelines, standards and exceptionally popular best practices.

Collections

Generally speaking, collections are the heart of libraries of all types (traditional, modern-automated and digital). In time, their concepts changed. Traditional and modern automated libraries are based on library collections of printed material. Digital libraries are based mostly on digital objects, both tangible and intangible, which may have physical presence but can also be present on other media, and can be accessed locally (directly) or at a remote location.

Digital collections are the core or the hub of all activities in digital libraries. As it was stated, collections of printed material were in focus of attention of librarians developing services in traditional libraries, and now digital collections attempting to attract the same attention of users in digital libraries to help librarians develop services.4

Along with attracting attention, digital libraries require additional empirical studies, (especially of user expectations of what content and services in digital libraries in Croatia will provide to users) which will create standards not only for widely asserted preservation, but also for making information resources accessible and interoperable i.e. discover what are the obstacles in physical access to information resources in digital libraries and what are parameters for sharing of resources set on different locations in different libraries. The second part of this paper will give some examples of simple application of a benchmark for digital collections.

Despite results achieved in the development of digital libraries, there are still many uncertainties regarding the current state of development of digital libraries in general because most of them were created at different times and had different levels of financial support, sometimes insufficient to progress from the state of a pilot project. Because of that and versatility of approaches in development of digital collections, it is hard to achieve consensus about the development of methods for evaluation of quality of digital collections.

One of the world's leading information scientist Tefko Saracevic points out that it is probably too early in the development of digital collections to establish formal evaluation efforts. This would put focus on details and distract attention from what can be actually measured as a part of the bigger picture in the process of development of digital libraries.5

In her publication about strategies for building digitized collections, 6 Abby Smith also suggests that it's too early for many libraries to have long-term considerations about digitization, as this is a period of experiments and building skills. Furthermore, while many libraries wish to have some kind of a digital project, there is no clear vision of purpose such projects will ultimately serve. As a consequence, it is difficult to expect that digital libraries will soon have developed evaluation tools (metrics) when there are no clear purpose for building digital collections and models of their use.

Services

4 Attracting attention of users is the most important task for digital libraries today. Only concrete use of concrete collections and emerging services can provide us precise insight into human-computer interaction and human behavior in digital libraries.
In accordance to information resources (library material) available in libraries, librarians offer a number of services such as reference service, document retrieval, interlibrary loan, lending etc. These well-known library services now seek ways for alteration and adjustment to the new, electronic environment. Judging from the global experiences in the development of digital libraries, they will offer both paper-based material as well as digital material with support of information technology. It is also expected that this trend will continue since it is quite impossible to digitize all existing human knowledge and there is no need for doing that. We can say that digital libraries will continue the work of hybrid libraries and offer digitization on demand as a new service for material explicitly requested by users (and which doesn’t fit into category of mandatory preservation – this material will undergo constant digitization).

It is expected that some of the services now existing in modern automated libraries will be modified to meet new user needs for print materials and will eventually be included in digital libraries enabling users to access information resources from previous types of libraries by use of services in digital library. Of course, new and more personalized services based on user profiles will be created for users demanding digital material. This will be the most exciting area of library development in the next decade.

Theoretical grounds are already provided. In their article about library and information services, Paul Kantor and Tefko Saracevic established taxonomy of user values for library services and preconditions for the carrying out the investigation based on those values. Both authors place great value on user feedback about library services. They suggest three steps in clarification of the value of library and information services:

1. Identification of attributes or dimensions of value; they should be organized in some rational structure
2. Development of procedures for appraising value according to each of these dimensions
3. Collecting and analysis of data following the dimensions and procedures identified in steps 1 and 2.

One of the methods digital libraries will most certainly use in their evaluations of future service is user feedback about the value of services, which helped them to get the requested document. Evaluation of the concrete user actions and outcome of those actions during the access to digital collections using multiple features will produce a corpus of scientific data about something that hasn’t been recorded yet or hasn’t been recorded extensively enough.

Derived from this and other citations, we can conclude that there is still no clear notion of the definitive shape of services in digital libraries. What we know is that services will include acquisition, storage, preservation, access, retrieval, delivery, and different aspects use of digital material.

A very similar multi featured viewpoint can be found in paper on measurement of electronic services in libraries written by Peter Young. He points out a number difficulties in measurement of use of resources in electronic environment on multiple levels: definitions for electronic media and services, lack of standards for quantitative measurement of digital objects, problems in developing standard definitions for measuring the collaborative and cooperative activities between libraries and other institutions in providing electronic media and services etc. He suggests use of measurement of information technology-based media and services in libraries based on a combination of several approaches:

1. Transaction-based measures: counting of interactive sessions, downloads, hits, patrons, domain and host addresses, images, and files and recording them and measuring by sampling or by transaction logs
2. Time-based measures: measuring service hours, session length/duration, system/server peak level
3. Cost-based measures: reports about cost/expenditure for telecommunications/bandwidth, hardware equipment, staff, training, maintenance, site licenses
4. Use-based measures: measuring user activities, anticipated demand, simultaneous users, group use, hits/patron, user satisfaction, local or remote/off-site use.

---

7 These are so called hybrid libraries – offering both material on paper (and other media) as well as digital material. All services are supported by information technology.
Due to the complexity of new information systems that evaluation of systems, it is obvious that resources and services will require multi-sided approach when being evaluated.

Kaplan & Nelson suggest another method of evaluation of new services. They investigated impact of publication in digital libraries. The result of their investigation is a proposal for retrieval analysis as a possible evaluation method for the services in digital library. This method would include total number of retrievals (or disseminations) a publication receives in a digital library. While this method of retrieval analysis is indeed a realistic approach to the problem of evaluation of digital library services, other metrics like citation analysis of printed and digital resources as well as investigation of the purpose of use of the retrieved digital object (its impact on further use by user in his/her area of interest) should be also taken into consideration.

Some authors go a step further and suggest scenarios for the development of services for digital repositories: the development of services will take place gradually, in phases. In phase I, digital repositories will offer rudimentary tools and processes to preserve and access materials. Services in this phase offer an equivalent of current services in libraries. Phase II will offer more advanced tools and processes. It is expected that these services will appear over the next decade. In the last phase, phase III, very advanced tools and processes will appear. Services in this final phase will follow after the next decade.

All this will be possible if digital objects i.e. digital library material will be produced in compliance with standards, which will provide assurance of quality. Then, digital objects will provide grounds and conditions for development of comparable and interoperable services. An example of one such large-scale project of digitization and accompanying services that has been available for several years can be found at http://memory.loc.gov (The American memory project – USA).

Where it all begins: digital collection development

There are three main methods for a digital collection development:

1. Digitization (conversion into digital format) of text, photos, manuscripts on paper or other medium
2. Transcription of a text or handwritten material
3. Acquisition of material already in digital form (so called digitally born material).

Digitization is the most popular method for conversion of text and image material already existing and stored on paper or other medium. Some of the reasons for employment of digitization are: making material more accessible by digitizing it, preservation especially for material for which there are standards for storage and retrieval, for educational purposes, for saving content from deteriorating medium etc.

Text, manuscripts and photos are of interest to the topic of this paper since they are the most popular and most available content formats and in the greatest need for application of standards, guidelines or benchmarks, which will provide assurance for quality of development of digital collections. The greatest application of digitization is in science i.e. digitization is used for data generated in all phases of scientific research and organized by digital libraries in a way that researchers can always access information regardless of their location.

There are, again, two major issues related to the use of benchmarks: preservation and access. Preservation is one of the most important reasons for the digitization of the material on various media. Good preservation strategy must not be depending on technology since technology changes too fast. In order to make original and digitized material independent of technology, we must use a set of benchmarks. This also helps digitized material to become more accessible.

---

12 LeFurgy, William G. Levels of Service for Digital Repositories http://www.dlib.org/dlib/may02/lefurgy/05lefurgy.html (18.6.2002.)
What is a benchmark?

Benchmarks are standards, especially resolution standards, necessary to meet requirements for authenticity, long-term value of surrogate copies, and their usability. When we speak about digitization of library material, we speak about three types of files on which benchmarks can be and should be applied.

1. Digital masters. Digital object optimally formatted and described with a view to their quality (functionality and use value), persistence (long-term access), and interoperability (e.g. across platforms and software environments). Such files should be stored on a stable medium and should remain in a controlled environment.
2. Access files for daily use. Used for access on Web pages (rather good quality)
3. Thumbnails: Small files for use in databases or Web pages.

Use of benchmarks: experiences

For each type of file there should be precise parameters for digitization. Different material waiting for digitization needs different approach and accompanying parameters. There are many digital publications available on the Internet presenting parameters for text, manuscripts, maps, drawings, black and white or color photos etc. Most of them are based on the actual project experiences and are result of them. For the purpose of this paper, we will choose only one such set of benchmarks, in this case, for textual material.

<table>
<thead>
<tr>
<th>Image Type</th>
<th>Printed Text</th>
<th>Damaged Printed Text</th>
<th>Handwritten Manuscripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master</td>
<td>Scan Type: Bitonal Resolution: 600 DPI Format: Uncompressed TIFF</td>
<td>Scan Type: 8-bit grayscale or 24-bit color Resolution: 400 DPI Format: Uncompressed TIFF</td>
<td>Scan Type: 8-bit grayscale or 24-bit color Resolution: 600 DPI Format: Uncompressed TIFF</td>
</tr>
<tr>
<td>Access</td>
<td>Type: 8-bit grayscale/24-bit color Format: JPEG Compression: Medium Spatial Resolution: Resize to 1024 x 768 Pixels</td>
<td>Type: 8-bit grayscale/24-bit color Format: JPEG Compression: Medium Spatial Resolution: Resize to 1024 x 768 Pixels</td>
<td>Type: 8-bit grayscale/24-bit color Format: JPEG Compression: Medium Spatial Resolution: Resize to 1024 x 768 Pixels</td>
</tr>
<tr>
<td>Thumbnail</td>
<td>4-bit grayscale/8-bit color Format: GIF Spatial Resolution: Resize to 150-200 pixels across the long dimension 72 DPI</td>
<td>4-bit grayscale/8-bit color Format: GIF Spatial Resolution: Resize to 150-200 pixels across the long dimension 72 DPI</td>
<td>4-bit grayscale/8-bit color Format: GIF (or JPEG) Spatial Resolution: Resize to 150-200 pixels across the long dimension 72 DPI</td>
</tr>
</tbody>
</table>

Table 1. Example of a benchmark for the text based and manuscript material ready for digitization


Benchmarks serve as control devices during the digitization process i.e. development of digital collections. We can use them in a form of a printed table, or as an interactive Web service. An example of interactive benchmark

---

14 It is also the first reason for digitization in many institutions. Access and use come second.
16 Benchmark for digital reproductions of monographs and serials as endorsed by the DLF http://www.diglib.org/standards.htm (18.6.2002.)
Web service is "Image quality calculator" available at http://images.library.uiuc.edu/projects/calculator. Based on given input parameters, the calculator is capable of making a calculation of recommended resolution necessary to capture all the details in the digital image.

As an example of use of this interactive benchmark, we have chosen a small format Croatian – German dictionary by Ivan Filipović, (printed in 1911. in Zagreb). The title page was scanned using a flat bed 600 dpi scanner and then compared to values obtained by another scanned text image done according to the set of benchmarks for printed text master file in Table 1.

The interactive Web service required entering several parameters: image height (text image), image width, quality index (small-high), dimension of the smallest significant printed character (in order to make him legible on the scanned image) and bit depth.

<table>
<thead>
<tr>
<th>Input parameters:</th>
<th>Results:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image Width = 9.2 cm</td>
<td>Recommended Resolution = 153 dpi</td>
</tr>
<tr>
<td>Image Height = 13.5 cm</td>
<td>Approximate File Size = 0.421 Mb</td>
</tr>
<tr>
<td>Quality Index = 8 (high)</td>
<td>Horizontal Pixel Dimension = 805 Pixels</td>
</tr>
<tr>
<td>Smallest Significant Character = 2 mm</td>
<td>Vertical Pixel Dimension = 548 Pixels</td>
</tr>
<tr>
<td>Bit Depth = 8-bit</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Example of use of interactive benchmarking tool

In the case of interactive Web service, the recommended resolution was 153 dots per inch, which is significantly smaller value when compared to the usual value of 600 dpi for the master file recommended in benchmark from Table 1. (source: Digital imaging and media technology initiative at University of Illinois at Urbana-Champaign).

The resulting size of the text image (Image 1.) was very close to the estimation given by the interactive image quality calculator: 0.433 Mb. The second text image (Image 2.) was scanned according to parameters for the first category (printed text) from the Table 1. Its size was: 0.878 Mb, twice as big as Image 1. What is most important, the quality of the first image was obviously inferior when compared to the second image file due to the smaller resolution used during the scanning process. Both images were saved as uncompressed TIFFs.
This simple example shows how the same specimen can produce different results according to two different benchmarks. Choosing the benchmark may be a relatively difficult task if there isn’t at least a coordinating body, which will make guidelines or recommendation regarding the existing benchmarks and modes of their use.

Application of benchmarks during the process of digital collection development is extremely important because it will enable digital objects in collections to be compared with other digital object in other collections and in general to enable creation of compatible digital library services that will be interconnected. When mentioning services, many of them, current and future, will be depending on the persistence of digital objects. Persistence is the next sine qua non of digital libraries. Without persistence of digital object, services will have nothing to be based on. Many institutions in the world (archives, libraries etc.) are concerned about long-term medium durability. There is no evidence that CD-ROMs will outlast paper and libraries are worried about other media such as privately made 35mm camera films that entered library collections and are exposed to deterioration. There is no general solution for this problem.

In his article about levels of services for digital repositories William Le Furgy discusses the question of preservation of digital materials. Only a fraction of digital objects will meet necessary conditions for optimal preservation and use, and other material that deviates from these conditions will be included in digital collections but their location and utilization will be difficult if not impossible. To ensure digital objects persistence, material must adhere to strict conditions regarding their construction and description (by use of proper metadata). In this way digital object will be successfully managed independent of specific technology.

Conclusion

There is no doubt that the diversity of digital collections available in this moment around the world will create substantial problems to users who will try to use them and for information specialist who will want to evaluate them. For instance, digital objects in some digital libraries can be accessed only by use of proprietary software, while other digital collections can be accessed by use of a Web browser. Some collections have been developed with users in mind (user centered), some are (sadly) just showcases, which will probably be just that without any proper service in future. In both cases (and in many others) it is important to know that there are benchmarks, which can be used to help developers and to guide them while they are trying to assure quality of digital objects. Only quality digital collections with clear preservation policies will enable digital libraries to develop new services based on quality digital collections. Is it already time for benchmarking? In case of collection development, the answer is a definitive yes. In case of quality evaluation of services, let us wait at least phase II in services development and if the definitive yes for quality of digital collection development prevails, we will have solid grounds for evaluation of digital library services too.

References

Benchmark for digital reproductions of monographs and serials as endorsed by the DLF http://www.diglib.org/standards.htm (18.6.2002.)


---

18 This may be of special interest to scientists wishing to access several information resources at once using one referring service. Examples of such services are SFX (www.sfxit.com) and X-Refer.

19 LeFurgy, William G. Levels of Service for Digital Repositories http://www.dlib.org/dlib/may02/lefurgy/05lefurgy.html (18.6.2002.)


LeFurgy, William G. Levels of Service for Digital Repositories http://www.dlib.org/dlib/may02/lefurgy/05lefurgy.html (18.6.2002.)


Web Content Accessibility Guidelines 1.0 http://www.w3.org/TR/WAI-WEBCONTENT/ (8.7.2002.)
