10th International Research/Expert Conference
"Trends in the Development of Machinery and Associated Technology"
TMT 2006, Barcelona–Lloret, Spain, 11-15 September, 2006

INFORMATION PROCESSING AND INFORMATION SYSTEMS – DIGITALISING LINE

Angela Repanovici, Marian Repanovici, Luciana Cristea
***Transilvania* University of Brasov,**
500036 Brasov, B-dul Eroilor 29
Romania

ABSTRACT
The digital information has dramatically influenced the development of the humanity. The article herein reveals a brief history and the influence of the digital information occurrence, it analyses the content and the support and storage techniques of the digital information.
In the end there are presented the types of digital information used in libraries.
Keywords*:* information, digital, digitizing information systems

1. INTRODUCTION

The explosive growth of communications networks after 1990, particularly in the scholarly world, has accelerated the establishment of the "virtual library". At the leading edge of this development is public-domain information. Residing in thousands of databases distributed worldwide, a growing portion of this vast resource is now accessible almost instantaneously via the Internet, the web of computer networks linking the global communities of researchers and, increasingly, non-academic organizations. Internet resources of electronic information include selected library catalogues, collected works of the literature, some abstracting journals, full-text electronic journals, encyclopaedias, scientific data from numerous disciplines, software, archives, demographic registers, daily news summaries, environment reports, and prices in commodity markets, as well as hundreds of thousands of electronic-mail and bulletin-board messages.
The vast inventory of recorded information can be useful only if it is systematically organized and if mechanisms exist for locating in it items relevant to human needs. The main approaches for achieving such organization are reviewed in the following section, as are the tools used to retrieve desired information.

2. STORAGE STRUCTURES FOR DIGITAL-FORM INFORMATION

Digital information is stored in complex patterns that make if feasible to address and operate on even the smallest element of symbolic expression, as well as on larger strings such as words or sentences and on images and sound. From the viewpoint of digital information storage, it is useful to distinguish between "structured" data, such as inventories of objects that can be represented by short symbol strings and numbers, and "unstructured" data, such as the natural-language text of documents or pictorial images. The principal objective of all storage structures is to facilitate the processing of data elements based on their relationships; the structures thus vary with the type of relationships they represent. The choice of a particular storage structure is governed by the relevance of the relationship it allows to be represented to the information-processing requirements of the task or system at hand.
In information systems whose store consists of unstructured databases of natural-language records, the objective is to retrieve records (or portions thereof) based on the presence in the records of words or short phrases that constitute the query. Since there exists an index as a separate file that provides information about the locations of words and phrases in the database records, the relationship that are of interest (e.g., word adjacency) can be calculated from the index. Consequently, the database text itself can be stored as a simple ordered sequential file of records. The majority of the computations

use the index, and they access the text file only to pull out the records or those portions that satisfy the result of the computations. When relationships among data elements need to be represented as part of the records so as to make more efficient the desired operations on these records, two types of "chained " structures are commonly used: hierarchical and network. In the hierarchical file structure, records are arranged in a scheme resembling a family tree, with records related to one another from top to bottom. In the network file structure, records are arranged in groupings known as sets; these can be connected in any number of was, giving rise to considerable flexibility. In both hierarchical and network structures, the relationships are shown by means of "pointers" (*i.e*., identifiers such as addresses or keys) that become part of the records.

3. ANALYSIS AND DESIGN OF INFORMATION SYSTEM

The building of information systems falls within the domain of engineering. As is true with other engineering disciplines, the nature and tools of information systems engineering disciplines, the nature and tools of information systems engineering are evolving owing to both technological developments and better perceptions of societal needs for information services. Early information systems were designed to be operated by information professionals, and they frequently did not attain their stated social purpose. Modern information systems are increasingly used by persons who have little or no previous hands-on experience with information technology but who possess a much better perception about what this technology should accomplish in their professional and personal environments. A correct understanding of the requirements, preferences, and "information styles" of these end users is crucial to the design and success of today's information systems. The methodology involved in building an information system consists of a set of interactive activities that are cumulatively referred to as the system's life cycle .

The principal objective of the systems analysis phase is the specification of what the system is required to do. In the systems design phase such specifications are converted to a hierarchy of increasingly detailed charts that define the data required and decompose the processes to be carried out on data to a level at which they can be expressed as instructions of a computer program. The systems development phase consists of writing and testing computer software and of developing data input and output forms and conventions. Systems implementation is the installation of a physical system and the activities it entails, such as the training of operators and users. Systems maintenance refers to the further evolution of the functions and structure of a system that results from changing requirements and technologies, experience with the system's use, and fine-tuning of its performance.

Many information systems are implemented with generic, "off-the-shelf" software rather than with custom-built programs; versatile database management software and its nonprocedural programming languages fit the needs of small and large systems alike. The development of large systems that cannot use off-the-shelf software is an expensive, time-consuming, and complex undertaking. Prototyping, an interactive session in which users confirm a system's proposed functions and features early in the design stage, is a practice intended to raise the probability of success of such an undertaking. Some of the tools of computer-aided software engineering available to the systems analyst and designer verify the logic of systems design, automatically generate a program code from low-level specifications, and automatically produce software and system specifications. The eventual goal of information systems engineering is to develop software "factories" that use natural language and artificial intelligence techniques as part of an integrated set of tools to support the analysis and design of large systems.

4. HIGH-QUALITY DIGITISATION SOLUTIONS

For a high quality digitisation, i2S DigiBook proposes a large range of monochrome and colour, manual and automatic scanners. They perform digitisation formats from A5 to A0 and optical resolutions up to 800 dpi.
The Book Restorer™ software is delivered with three external modules – BKR Automaton, BKR Watcher and BKR Control. These modules allow the automation of Book Restorer™ treatments – in a batch mode or continuously during the digitisation (workflow) – and the quality control of restored images.
The Book Restorer ™ modules guaranty your productivity without compromise of the image quality.
Book Restorer™ can create scripts by selecting appropriate treatments. A wizard application can help the operator to define the scripts according to the work kind.

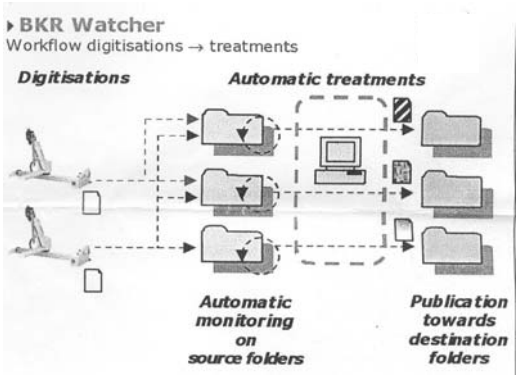

Figure 5. BKR Watcher

5. DIGITISATION LINE PROPOSED AT THE LIBRARY FROM *TRANSILVANIA* UNIVERSITY OF BRAŞOV

In a DigiBook scanner, the digitisation process includes a document scanning phase (see step ❶ on 1st timing chart form fig. 6), followed with an image recording phase, to the local disk or the network (see step ❷). While the camera head and its embedded lighting system are scanning the document, the image acquisition and display are done simultaneously. At the end of the scanning period, the image is available in the processor memory, and is saved on the disk. So far, a new scanning step could not be launched before both of the following conditions were realised: the end of the previous image saving and the end of the page turning (see step ❸). The new DigiBook application software allows now to execute in parallel the scanning and the saving step. Consequently the scanning operations can be run more quickly, without waiting for the end of the image recording. Whatever the format of image used, and with very rare exceptions, the image archiving period is always lower than the one of a scan. The productivity is now directly dependent on the scanning time to turn the page or to change the document (see 2nd timing chart).

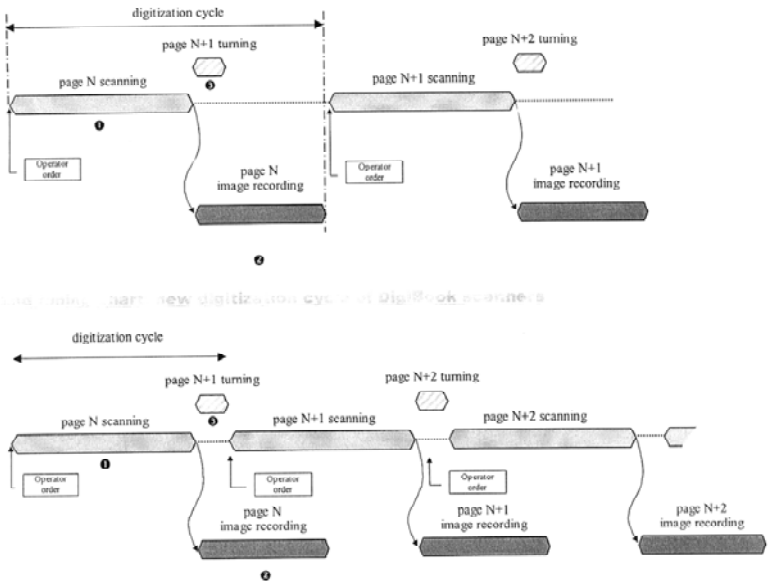

Figure 6. The digitisation process

This release is particularly interesting for large size and high resolution colour formats, where recording times are long and critical for the productivity. The attached chart compares the productivities measured on a machine equipped with a 10,000 pixel colour camera head, between two recent versions of the DigiBook software (V5.22 and V5.3), the last one integrates this release. These productivities take into account an average time of page turning equal to three seconds. The gain of productivity can reach almost 30%. This release will soon equip all the A0 and A1 DigiBook scanners, except the DB5600 model. To upgrade already delivered machines, the technical departments mist be contacted to check that memory and processor configuration inside the machines are sufficient.

Format	Resolution	Previous theoretical productivity	New theoretical productivity	Gain
2*A3 Colour	600 dpi	131 p/h	170 p/h	+29,7%
2*A2 Colour	400 dpi	148 p/h	189 p/h	+27,7%

6. CONCLUSIONS

By analyzing the projects developed so far and by realizing the fact that there are no developments of this type in Romania, under the specific financial conditions, the research team wishes to obtain obtimization being considered the cost and time related to the quality of acces and of the image of digitized documents, by elaborating an automated digitized line and by optimizing various parameters on different types of documents. There will be developed a research didactic base within Transilvania University Library. The automated digitizing line aims at creating, in the short run and with maximum efficiency, on-line bibliographies on various fields updates at each beginnig of the university year. The digitizing line will be made up of equipment that we already own and that were used for other projects and the new equipment will be assembled in the research laboratory.

The digitizing line will effectuate researches in order to obtain maximum results regarding: scanning spped, which involves increased productivity; quality of the scanning process; assuring the performance for a large category of documents of 2×A2 format; maximum resolution; adjusting the image vizualization and archiving the ideal one; protection for the digitization process for small, thick and tightly bound books; minimizing the physical contact with the user; using the metadata as research instrument, creating a metadata service; analyzing a set of DublinCore elements.

7. REFERENCES

[1]. Alger Jeff : Effective DATABASES for Text & Document Management, Issues in Science and Technology Librarianship, 2004.

[2]. Ayre Catherine, Muir Adrienne: The Right to Preserve: The Rights Issues of Digital Preservation, D-Lib Magazine, 2004.

[3]. Bakelli Yahia, Benrahmoun Sabrina: Long-Term Preservation of Electronic Theses and Dissertations in Algeria, LIBRI: international journal of libraries and information services 53(4), 2003.

[4]. De Stefano Paula: Moving Image Preservation in Libraries, Library Trends 52(1), 2003.

[5]. Dean John F: Digital Imaging and Conservation: Model Guidelines, Library Trends 52, 2003.

[6]. Digitization, Conservation and Quality Control * XML as a Preservation Strategy * The Long-term Preservation of Databases, ERPANET Newsletter, 28 October 2003.

[7]. Eason Ken : Towards a User-Centred Approach to Digital Libraries, Ariadne, Issue 38, 2004.

[8]. i2S-DIGIBOOK: http://www.i2s-bookscanner.com/pdf/digibook_mag_no6_en.pdf

[9]. Reilly Jr, Bernard F.: Preserving America's Printed Resources: The Roles of Repositories, Depositories, and Collections of Record, Library Link News, 2004. www.emeraldinsight.com/lm.htm

[10]. Williams Mark: What the Resource Discovery Network is doing for Further Education, Ariadne, Issue 38, 2004.