Introduction

The world produces between 1 and 2 exabytes (exabyte is a billion gigabytes) of unique information per year, which is roughly 250 megabytes for every man, woman, and child on earth. Printed documents of all kinds comprise only 0.003% (SIC!) of the total (www.sims.berkeley.edu/research/projects/how-much-info/summary.html: Jun 2002). This means, that digital archiving and preservation is a very important topic which is currently attracting a great deal attention.

Ever shorter technological innovation cycles replace existing technologies at a breathtaking pace of 2-5 years. File formats and compression schemes are also constantly changing. Therefore, any information stored in digital form may become inaccessible within less than a decade. The case of Domesday Book illustrates the problem very well: “The 1086 Domesday Book, instigated by William the Conqueror, is still intact and available to be read by qualified researchers in the Public Record Office. In 1986 the BBC created a new Domesday Book about the state of the nation, costing £2.5 million. It is now
unreadable. It contained 25,000 maps, 50,000 pictures, 60 minutes of footage, and millions of words, but it was made on special disks which could only be read in the BBC micro computer. There are only a few of these left in existence, and most of them don’t work. This Domesday Book Mark 2 lasted less than 16 years.” (Jackson, 2002).

The above example clearly indicates that access is the key factor: if a resource cannot be used anymore, it is totally pointless to preserve it. Therefore, in the age of digital information the issues of metadata and digital preservation have become of crucial importance.

**Digital preservation principles**

In contrast to physical objects, digital objects cannot be put away and stored at constant humidity and temperature, and then accessed hundreds of years from now. Preserving digital objects calls for constant attention – e.g. methods need to be implemented to ensure bits can be correctly interpreted – and that the interpretation mechanisms are documented, and that scalable methods for migration and documentation of the transactions are employed (Ross et al, 2003).

Preservation over time presents a double challenge:

- to identify pertinent and coherent data objects to be preserved, with correct metadata describing the context of creation, organised in such manner that they remain independent of any tool or system which may have a shorter life expectancy than the data object;
- to preserve those objects over time in manner which preserves their accessibility to users.

The longevity problem can be divided into two questions:

- the lifespan of the medium on which the file is stored;
- the obsolescence of the format in which the file is stored.

Referring to the former, it is significant that low life expectancy (LE) ratings are common for the 20th century media formats. Particularly, digital formats have low LE – less than 100 years, as compared with microfilm which is estimated to have an LE of 500 years. While many paper resources are already older than that. Despite this fact it is currently accepted that it is obsolescence of the format, which poses greater risk than the medium integrity, mainly due to the accelerating technology advancement.

Digital preservation is a multi-faceted problem. It is, however, possible to indicate three overall preservation methods. These methods embrace: refreshing, migration and emulation, and can be additionally sub-divided into short-, medium- and long-term solutions. Short- and medium-term methods include:

- refreshing: periodically copying the data onto a newer carrier of the same type;
- transfer: copying data onto a more stable carrier (e.g. transferring data from a floppy disk to a CD-Rom);
- multiple distributed copies („cloning“): storing identical material in multiple locations and regularly backing up (in order to protect against loss due to media failure or human error);
- format migration: converting documents to newer and probably less volatile standard formats to assist in maintaining access and facilitate later migration (NB the open standards formats, such as TIFF, JPEG or GIF, are far more stable, than those, which require particular software).

While medium- to long-term preservation solutions are as follow:

- technology preservation: preserving entire computer hardware platforms;
- migration: periodically transferring digital material from one hardware and software configuration
to another, or from one generation of computer technology to a subsequent generation;
- emulation: using software that emulates obsolete encoding formats to provide access to programs across different platforms.

More recently a global format with great potential for the preservation of digital information has been developed, namely eXtensible Mark-up Language (XML). XML is a markup language for documents containing structured information. Structured information contains both content (words, pictures, etc.) and some indication of what role that content plays (for example, content in a section heading has a different meaning from content in a footnote, which means something different than content in a figure caption or content in a database table, etc.). Almost all documents have some structure. A markup language is a mechanism to identify structures in a document. The XML specification defines a standard way to add markup to documents. XML has arrived as a key technology for the recent stage of evolution of the Internet. In the beginning, its core characteristics of self-description and ad-hoc extensibility offered the flexibility needed for transport of messages between various applications. Lately, the next generation of XML standards – such as XML Schema – have enabled unification of both document modelling and data modelling.

Another issue related to the digital preservation is metadata. Although Latin and English usage of “meta” refers to something transcendental, or beyond nature, the original meaning of the term “meta” comes from the Greek, and stands for “alongside, with, after, next”. Metadata, then, can be named data about other data. In other words, it is the Internet-age term for information that librarians traditionally have put into catalogues. In this respect, metadata is critical for the effective management of data resources, most commonly the Web resources (Hofman, 2000).

For data to be meaningfully processed, metadata associated with it must be present and accessible. It provides the information required to identify data of interest based on content, validity, sources, physical material, pre-processing, legal aspects or other selected criteria. Metadata gathered during the capture and record keeping processes allows to establish the provenance and authenticity of the digital records as it addresses their content, context, form and structure (Hofman, 2000).

**A new task for schools of architecture?**

The need for digital preservation touches all our academic activities, whether we work on research projects, teach students, deal with administrative data or use a digital camera. In all these instances we create and use e-content and expect that this content will remain accessible to allow us to trace and validate what we have done, or pass a record to future generations.

Computer-aided design has developed into architectural information and communication technology (ICT), to become the main tool of the majority. This was proved by the recent survey by Hannu Penttila aimed at giving an understandable and measurable overview of the current architectural education and its relation with ICT and CAAD.

According to Penttila (2003) one of the key academic activities is “Information management”, namely: storing, saving, archiving information, sharing and finding information.

Each year in schools of architecture students produce thousands works during the CAAD courses, design studios, media labs. And the students’ projects are stored for future use as case studies in educational database systems or web galleries (Dierckx et al, 2002; Heylighen and Neuckermans, 2000). However, in the light of the above, the archiving and management such a mass of digital data comprised of research and students’ work records and files produced every year poses a major problem related to the digital preservation generally defined as managed activities to ensure continued access to elec-
tronic resources. If the majority of architectural work is digitally born it poses the serious threat that it may become irretrievable even in a close future.

In this respect, first general, yet fundamental questions should be posed: How do we manage the digital data and keep them alive for a long term? What methods do we apply? Do we implement standards? There are, furthermore, a number of detailed problems that should be addressed, for example:

- Archiving a great number of students’ works, to become a useful source of information, requires quick searching and accurate retrieval. The use of metadata may be helpful in this task, as it allows to document everything the user needs to make a decision if the resource is usable.
- Various CAAD software have their own formats which are unreadable by other programs. What is more, the most of works make use of the library objects usually downloadable directly from the software with each file opening. Therefore, with the change of software platform used within the curriculum, the students’ works from previous years may become inaccessible. Until recently the most widely advocated solution to these problems has been that of migration. But migration has its problems and disadvantages, for instance: the process of conversion runs the risk of losing data, and moreover it is a time consuming and, therefore, costly process. For these and other reasons emulation may provide a better solution.

The above examples seem to indicate clearly that digital preservation should be adopted by the schools of architecture.

**Conclusions**

The goal of undertaking this subject has been to raise awareness of digital preservation technology and its implementation into a current academic practice. It has been pointed in the previous section that it would be advisable for schools of architecture to pay attention to the issues related to digital preservation. This does not necessarily mean pursuing a dedicated research task from the scratch, since the survey has revealed there are well-established international and interdisciplinary communities dedicated to the field of digital preservation. The schools should, therefore, take advantage of the well-developed and proven technology for digital preservation and customise it to their requirements, instead of reinventing the wheel.

It is, for that reason, necessary to keep up with new technologies and to identify good practice in this field. Good sources of valuable information in the subject area should be, therefore, defined. A good starting point could be expert websites and mailing lists, such as those of the Digital Preservation Coalition (www.dpconline.org/) or the Joint Information Systems Committee (www.jisc.ac.uk/dner/preservation/).

This paper has intended to open a discussion among the schools of architecture on the need for the digital preservation and to exchange of experiences in this subject. To achieve that it would be advisable to undertake a survey (for example in the form of a questionnaire) which would serve the eCAADe members as a source of information, and launch a new forum in the networked society we are.

**References**


Heylighen, A. Neuckermans, H.: 2000, DYNAMO


