

A Study of Portability in the Deployment of WWW

András Micsik

Several problems are experienced daily in the case of handling compound hypermedia WWW documents. One class of these problems, namely portability is studied in this paper in detail, and possible solutions are given using a new container architecture, the Portable HyperMedia (PHM) format. Structure and operations on PHM are discussed and connections with recent efforts in the field of metadata are explained.

Keywords : hypermedia documents, document-like objects (DLO), WWW, portability, Warwick Framework, Dublin Core, Internet

Introduction

As the World Wide Web [WWW] spread the world from the beginning of this decade, it incorporated more and more powerful tools and formats, and the information served via WWW became more and more complex. The content and layout of WWW pages become competitive with printed material, and in other aspects WWW pages have far more potential than printed documents. The meaning of document in case of the WWW is changing. WWW documents are sometimes more similar to a piece of software than to printed material. They may contain animations or has an annotation facility, and what is most important they are linked together.

Hypermedia documents has rich history. One can think about popular applications like ToolBook or HyperCard. However a large part of the functionality in these software are repeatable with current Internet technologies like WWW, CGI and Java. One significant difference between the two solutions is that while a document prepared with ToolBook can be easily moved to another computer, a similar WWW document may lose most of its functionality when moved to another location. One cause for this is that hypermedia documents on the Internet usually consist of hundreds of files. But very often WWW-based documents are still desired to use as single documents, to send, to store, to manipulate easily.

A new term appeared for such documents: Document-like Objects (DLO) [Ferber96]. DLOs has a complex nature:

- they may contain files in lots of different formats: text, graphics, animation, video, audio and 3D (VRML)
- they may contain executable parts (JAVA and CGI)
- their files and data are interconnected with links.

In this paper a way for evolution of Internet hypermedia documents is shown, using a new container format, the Portable Hypermedia (PHM) coupled with metadata description and automatic learning techniques. The PHM container does not obsolete current WWW usage and file formats, but enhances the manageability of hypermedia documents on the Web. The main advantage of this approach is the seamless integration of this new format into the current Internet without requesting the rewriting of the current tools.

Furthermore the PHM format is also capable for solving problems with the description and cataloging of available data on the Internet, pointed out by the Dublin Metadata Workshop [WGMD95].

In Section 2 common problems in current usage of WWW documents are examined, setting the requirements for current needs (Section 3), and proposing a new container format, the Portable Hypermedia (PHM) for the Internet society. The solutions offered for the problems listed in Section 2 are readable in Section 4.