INTRODUCTION

We would like to introduce a new series of topics which would discuss the history of Portals and their role in web developments. There is a question we wish to ask: does the history of web computing and portals development start with introduction of markup languages or could we speculate that the main incentive for developing HTML and web pages came about as a result for example the inherent difficulty to develop graphical user interface with minimum effort using Windows or OS/2 Presentation Manager, or does the web development signify the departure from client – server concept with heavy client footprint or anything else? Or could the web be considered as realization of Vannevar Bushs idea about store of knowledge, “instruments are at hand which, if properly developed, will give man access to and command over the inherited knowledge of the ages” (Bush, 1945).

In this series of papers we would like to explore these hypothetical questions further.

Markup Languages – where they come from?

Historically, electronic manuscripts contained control codes (“macros”) that caused the document to be formatted in a manner called “specific coding.” In addition, there was also the notion of so called generic coding, which can be dated back to late 1960s. Generic coding used descriptive tags for headings etc. Today, we distinguish between two concepts: procedural markup and descriptive markup.

The first approach that has been explored was to define a procedural markup. The procedural markup specifies how to process text, instructs agent what to do with text, it is generally concerned with formatting, presentation and it is “efficient” because requires little further interpretation. IAs the first procedural interpreters appeared, it was soon discovered that procedural markup has some drawbacks such as it is specific to single proprietary processing system, the processing a document focuses on a single purpose (printing on a paper or viewing on a screen).

On the other hand, descriptive markup identifies the logical components of a document, it is generally concerned with what text is, it does not specify what procedures are to be applied to text and therefore requires that other process is responsible for suppling the formatting. Benefits of descriptive markup language include readability by human or machine, identifies information content, rendition type is not important and can be made non-proprietary. The independence of rendition
and non-proprietary nature of the interpreter were decisive features which contributed to the fact that the concept of descriptive markup languages has become the turning point for developing modern markup languages.

There is significant evidence that modern markup languages had their beginnings at a government meeting in Ottawa in 1967. William Tunnicliffe presented his theory on “The Separation of Information Content of Documents from their Format” (Goldfarb, 1996). This concept was further developed into Generalized Markup Language (GML) by a team at IBM, which included Charles Goldfarb, Edward Mosher, and Raymond Lorie. This project was named GenCode and has its roots in work of Tunnicliffe. At the same time, Stanley Rice, a book designer from New York envisaged that a set of labels would allow description of a so-called editorial structure. The process of developing GML was not straightforward. The conclusions of GenCode committee after many meetings stated that “It is impossible to describe all the documents with one set of codes. Markup should be descriptive rather than procedural and Markup should take into account the hierarchical structure of the document”.

Nevertheless, the next generation of GML - Standard Generalized Markup Language (SGML) - was accepted as an ISO standard in 1986. IBM was one of many companies to use SGML for most of their publishing purposes.

William Tunnicliffe was born and raised in Washington, DC. However, he lived most of his life in Winchester, Massachusetts. He graduated from Worcester Polytechnic Institute in 1943 and then continued at Harvard where he received degrees in applied physics and engineering sciences (1951). Tunnicliffe also enlisted in the United States Navy and received the rank of a captain. Tunnicliffe worked in several companies such as Raytheon Corporation and the Courier Citizen newspaper in Lowell, Massachusetts before he founded his own engineering firm - Tunnicliffe Associates - in Winchester. He managed the firm for 10 years before retiring.

Quote from Appendix A of The SGML Handbook, Goldfarb, p. 567, on the Generic Coding Concept (copyright 1989, SGML Users Group) give the credit to Tunnicliffe as father of markup languages:

“Historically, electronic manuscripts contained control codes or macros that caused the document to be formatted in a particular way (“specific coding”). In contrast, generic coding, which began in the late 1960s, uses descriptive tags (for example, “heading”, rather than “format-17”). Many credit the start of the generic coding movement to a presentation made by William Tunnicliffe, chairman of the Graphic Communications Association (GCA) Composition Committee, during a meeting at the Canadian Government Printing Office in September 1967: his topic -- the separation of information content of documents from their format.”

Markup languages use instructions embedded in a document to provide the typesetter with information for visual characteristics of the document. For word processors, markup language is a code included in the document to provide display instructions, and sometime also processing instructions. SGML is a generic “markup” language used by publishers and multimedia companies for organizing and tagging documents. Publishing problems typically arise from incompatibility between text editing, formatting, and database applications. SGML provides the instructions with regards how the text is to be formatted. The computer program such as modern word processor interprets this markup and transforms the information into the desired output format. In modern word processors such as Microsoft Office Word, the markup is hidden from the end user. The markup language used in World Wide Web is called HyperText Markup Language (HTML). HTML is officially defined as the “The authoring language used in the creation of documents for the World Wide Web”.

Not all the credit should be given to the brilliant idea of separation of concerns (as we
would call it today) - document content and its presentation. There was another genius – Bill Atkinson and his HyperCard. In the late 1980s, Bill Atkinson, a programmer working for Apple Computer Inc., developed the application called Hypercard for the Macintosh. Hypercard enabled to construct a series of on-screen 'filing cards' that contained textual and graphical information. Users could navigate these by pressing on-screen buttons, taking themselves on a tour of the information in the process.

Hypercard set the scene for more applications based on the filing card idea. For example, Toolbook for the PC was used in the early 1990s for constructing hypertext training courses that had 'pages' with buttons which could go forward or backward or jump to a new topic. On a background, buttons would initiate small programs called scripts. These scripts would control which page would be presented next. They could even produce some animation on the screen. The concept of linking pieces of information together in some semantic manner was born. Another application based on similar principle named Guide was developed for UNIX and the PC.

Hypercard and the applications based on the same principle became quickly popular tools. However, these packages still had one major limitation: they were confined to a single machine. Hypertext jumps could only be made to files on the same computer. Connections to and information retrieval from computers on the other side of the world were still out of the question. Nobody yet had implemented a system involving text linking on a global scale.


Jana Polgar
Greg Adamson
Editors-in-Chief
IJWP

REFERENCES

