Experiences of SGML Standardization:
The Case of the Finnish Legislative Documents

Airi Salminen, Virpi Lyytikäinen, Pasi Tiitinen
University of Jyväskylä
Department of Computer Science and Information Systems
Email: airi@cs.jyu.fi, lyviau@cc.jyu.fi, pti@cc.jyu.fi

Olli Mustajärvi
The Parliament of Finland
Email: olli.mustajarvi@eduskunta.fi

Abstract

The SGML standardization in the Finnish Parliament was activated in 1994 when a project called RASKE was commenced in cooperation with the University of Jyväskylä. The project carried out a careful analysis of document management and designed 15 preliminary SGML Document Type Definitions for legislative documents. The work has been followed by implementation projects where selected companies have developed and implemented SGML solutions for a specific subset of documents, and the Parliament and ministries have redesigned their work processes. The effects of the standardization have concerned documents, document production, archiving practices, information distribution, and inter-organizational collaboration. From the point of view of a citizen, the major effect of the new solutions is in the improved accessibility to legislative information through the Internet. The capabilities offered by the structured form have motivated the Ministry of Justice initiating the design of a database for consolidated legal text.

1. INTRODUCTION

A major portion of information management in public sector consists of document management. Therefore, implementation of better services for the citizens in an information society requires effective implementation of electronic document management.

SGML (Standard Generalized Markup Language) is an international standard for defining and representing documents in an application-independent form. The International Organization for Standardization (ISO) accepted the standard in 1986 [6]. The standard is based on the idea of structured documents: before creating documents their logical structure is first defined in a Document Type Definition (DTD). In the DTD the hierarchic structure of the parts of a document is specified, and possible attributes are associated with those parts to add extra semantic information. In the middle of 1990s, the need for a document definition and representation language especially for information distribution over the Internet was widely recognized. The World Wide Web Consortium (W3C) activated the development of XML (Extensible Markup Language) and published the XML 1.0 specification in 1998 as a W3C recommendation. In parallel with the XML specification activities, the SGML standard was amended and corrected for use on the World Wide Web [5]. XML is a subset of the amended and corrected SGML standard. SGML offers for the purposes of e-Governance a technology for document management by enabling long-term accessibility of information in documents. In the Finnish public administration, discussion about building and supporting e-Governance by means of SGML started in 1994. That year a collaborative project of the Finnish Parliament and researchers of the University of Jyväskylä was established. The project was called RASKE (coming from the Finnish words “Rakenteisten asiakirjastandardien kehitäminen” meaning the development of standards for structured documents). The first results of the project were described in the HICSS conference of 1996 [11]. From