Supporting interchange of open hypermedia structures and contents

Kaj Grønbæk & Lennert Sloth
Department of Computer Science, Aarhus University,
Åbogade 34, DK-8200 Århus N, Denmark.
Email: {kgronbak, les}@daimi.au.dk

ABSTRACT
This paper discusses issues in porting hypermedia structures and contents between Open Hypermedia System (OHS) servers. Even though open hypermedia system users are operating in a highly distributed computing environment, there are several situations where open hypermedia structures need to be moved or copied between servers instead of just being remotely accessed. Examples of this are: 1) when users of OHSs wish to publish structure and content developed inside a firewall to a public machine residing outside the firewall; 2) when users wish to download structure and content from a network to a portable machine; 3) when structure and contents needs to be physically moved over the internet due to performance reasons or the like. The paper proposes an approach for providing such interchange by means of XML.

Keywords: OHS, Interchange, XML, Firewall, data model.

INTRODUCTION
Interchange formats have been a research and standardization focus for many individual types of media, such as text and video. Within the area of hypermedia, work on interchange mainly focus on the hypermedia structures and not the contents per se. The Dexter model (Halasz & Schwartz, 1990) made a proposal for an SGML-family interchange format for hypermedia structures. This proposal was used e.g. to interchange hypertexts between two quite different hypermedia systems KMS and Intermedia (Killough & Leggett 1990).

In the open hypermedia systems community (OHSWG 1998) we see needs for developing a similar interchange format. OHS research is focussing on imposing structures on top of standard document formats residing on network file systems or on the Web. There is a need to be able to port such structures between different implementations of OHS servers as well as between the same server implementation sitting on separate networks e.g. on each side of a firewall.

Open hypermedia structures represent a special challenge, since the structures are highly dependent on availability of and consistence with the documents being structured by means of the OHS. In the following we describe an approach which is developed from the Devise Hypermedia Framework (Grønbæk & Trigg 1994).

The Devise Hypermedia (DHM) Framework is used to build an OHS, which integrates WWW documents (Grønbæk et al. 1997). The original framework provided an interchange format, which could export and import structures, called HyperSpaces, in an SGML-style format. To be truly portable the interchange format files started with a listing of all the documents being referenced in the HyperSpace to be interchanged. The documents were listed with a unique ID and a full filesystem path. The hypermedia structures only use the unique ID, whereas a lightweight document management tool, the DocMgr, when requested by the DHM system, looks up the full filesystem path. This DocMgr among other things supports editing the filesystem paths when documents are moved or imported from an interchange file, finally it can produce the output necessary to package and compress (e.g. ZIP) the interchange file together with the individual documents referenced in the HyperSpace being interchanged.

USE OF INTERCHANGE FOR OHS ON THE WWW
In the COCONUT project (http://www.cit.dk/coconut/) we have extended the DHM framework and applications to support a variety of open hypermedia structuring mechanisms which are currently being tested in industrial contexts having to deal with firewalls separating an internal OHS server from a public OHS server.

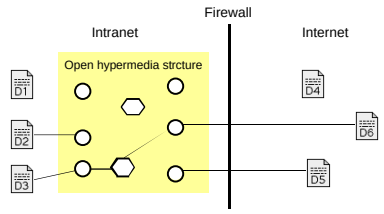

Figure 1: Authoring OHS structures on an Intranet. The structures may reference both local files (D1-D3) and public files (D4-D6)

In both our main cases the authors of OHS structures are developing Web documents on intranets where they use OHS support to create structures, such as links and notes, on top of both local and public internet documents (see Figure 1). When done with the publishing task, the authors need to publish both documents and structures on a Web-server and an OHS server outside the firewall (see Figure 2).

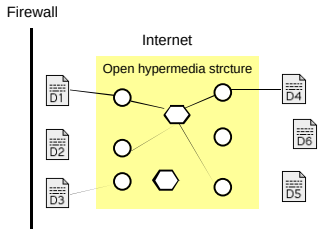

Figure 2: Documents and structure, which have been published outside the firewall. D1-D3 have been copied.

To support this we have updated the DHM interchange format to generate an XML document and to be able to handle local filesystem paths as URLs. This enables the DocMgr to support an easy publish of the structures together with the documents being published, e.g. via FTP or via the HTML editor's publish function which typically preserves all the relative addresses.

STANDARDIZING DATA-MODEL AND INTERCHANGE FORMAT

In the OHS community a standard data-model for open hypermedia structures has been proposed (Grønbæk, 1998), which is used as the source for generating the XML DTD for the interchange. An outline of the currently agreed upon OHSWG navigational data-model is depicted in Figure 3.

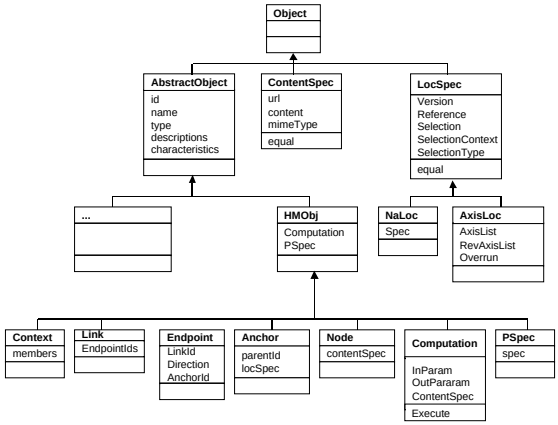

Figure 3: Outline of OHSWG navigational data-model

From this data-model, an XML DTD is derived which in turn allows OHS systems to use standard XML parsers to load the structures into an OHS, which is compliant with the standardized data-model. Excerpts of such XML documents are shown below. The format is specified to list pairs of document IDs and ContentSpecs including URLs at the top in order to make the DocMgr make necessary conversions of all local (intranet) URLs when registering the documents in the OHS server during an import of an interchange file.

```
<DocElm>
<id> 4 912169252  </id>
<contentspec>
<url>http://www.w3.org/TR/REC-DOM-Level-1/</url>
<contents></contents>
<mimeType>Text/HTML</mimeType>
</contentspec>
</DocElm>
```

Links are collections of endpoints as suggested in Figure 3.

```
<link>
  <id>   HS2:000000000000006  </id>
  <type>   ohp-nav/anchoredLink  </type>
  <name>   DOM  </name>
  <pSpecId>   </pSpecId>
  <descriptions>
    <name>default</name><value>DOM specifications</value>
  </descriptions>
  <characteristics>  </characteristics>
  <computationId>  </computationId>
  <endpointIds>
     <endpointId>000000000000010</endpointId>
     <endpointId>000000000000014</endpointId>
     <endpointId>000000000000020</endpointId>
  </endpointIds>
 </link>
```

CONCLUSION

This paper has given examples of the need for a standardized interchange format for OHS structures, and it has proposed an approach for the design of an interchange format based on the OHS data model. Finally, an approach for handling external document references during interchange based on the DHM DocMgr tool is described.

REFERENCES

Bray, T., Paoli, J., Sperberg-McQueen, C. M. Extensible Markup Language (XML) 1.0. W3C Recommendation 10-February-1998. http://www.w3.org/TR/1998/REC-xml-19980210

Grønbæk, K. OHS Interoperability - issues beyond the protocol. In proceedings of OHS Workshop 4.0 held at Hypertext '98 in Pittsburgh, June 20-24, 1998

Grønbæk, K., Bouvin, N. O., and Sloth, L. *Designing Dexter-based hypermedia services for the World Wide Web*. In proceedings of Hypertext 97, Southampton, UK, April 6-11, 1997.

Halasz, F. and M. Schwartz, The Dexter Hypertext Reference Model. Communications of the ACM, 1994. 37(2): p. 30-39.

Killough, R., & Leggett, J. Hypertext Interchange with the Dexter Model: Intermedia to KMS. TAMU-HRL 90-002. Hypertext Research Lab, Texas A&M University. August 1990.

OHSWG. Open Hypermedia Systems Working Group WWW site (Sep 98). Website: http://www.ohswg.org/.