

# Web Content Adaptation Frameworks and Techniques

**Tiong-Thye Goh**

*Victoria University of Wellington, New Zealand*

**Kinshuk**

*Massey University, New Zealand*

## INTRODUCTION

Most Web pages are designed with desktop platform access in mind, but with the proliferation of mobile devices such as Personal Digital Assistants (PDAs) and mobile phones, accessing Web pages through a variety of devices without proper content adaptation can result in an aesthetically unpleasant, un-navigable and, in most cases, unsatisfying experience. This article provides an overview of approaches in Web content adaptation framework and techniques being developed to extend the Web application access to non-desktop platforms. After describing general adaptation techniques, the article focuses particularly on the adaptation requirements of learning systems, especially when they are accessed through mobile devices.

## WEB CONTENT ADAPTATION

Since most existing Web applications are geared towards desktop platforms, they limit the access only to a certain class of users, hence restricting the potential customer growth of the enterprise. With the increasing proliferation of a diverse set of devices accessing the Web under different network conditions, the need for content adaptation is significantly increasing. To circumvent this problem, various commercial products and research prototypes dealing with Web content adaptations have emerged, such as Spyglass (Spyglass, 2001), Intel QuickWeb (Intel, 1998), IBM Transcoding proxy (Smith et. al. 1999), Digester (Bickmore & Shilit, 1997), Mobeware (Angin, Campbell, Kounavis & Liao, 1998), TranSend (Fox et. al 1998a), WingMan (Fox et.al. 1998b) and Power Browser (Buyukkokten, Garcia-Molina, Paepcke &

Winograd, 2000). The types of content adaptation these systems looked into are mostly multimedia rich transformation. In contrast, there are other areas, such as mobile learning, which require the development of Web content adaptations for mobility with respect to user environment and capabilities. These areas have distinct features not yet researched extensively. This article provides an overview of some of the promising frameworks and techniques in content adaptation.

## RE-AUTHORING

According to Bickmore and Shilit (1997), one straightforward method for content adaptation is to re-author the original Web content. Manual re-authoring can be done, but obviously, it is the most ineffective way and requires that the Web pages be accessible for re-authoring. This sometimes poses some practical constraints. However, the underlying principles and questions faced are identical for both automatic and manual re-authoring: What are the strategies used to re-author the pages? What are the strategies used to re-designate the navigations? What presentation styles can be achieved? These questions face any content adaptation process. The underlying principle is to isolate and distinguish the Web content objects, presentation objects, navigation objects and interactive objects for desktop publication and re-map them into other device-capable objects. Figure 1 shows such a re-mapping process. Once the strategies have been defined and the process matured, manual re-authoring can be converted into automated re-authoring through HyperText Transfer Protocol (HTTP) proxy server or server side techniques, such as common gateway interface (CGI), Servlet or client

side scripting. The re-authoring approach can either be mobile-device specific or tailored to multiple classes of devices. For multiple devices re-authoring, transformation style sheets (XSLT) and cascading style sheets (CCS) can also be used.

From another perspective, re-authoring can be viewed along two dimensions: syntactic (structure) vs. semantic (content), and transformation (convert) vs. elision (remove). Syntactic techniques operate on the structure of the page, while semantic techniques rely on the understanding of the content. Elision techniques basically remove some information, leaving everything else untouched, while transformation techniques involve modifying some aspect of the page's presentation or content. The Digester system (Bickmore and Schilit, 1997) used the re-authoring technique that included outlining, first sentence elision and image scaling, and built an abstract syntax tree to provide content adaptation. The Digester system used a proxy-based heuristic approach for its automated re-authoring. This method worked well for small-screen mobile devices. However, the elision process might remove certain content and affect the capturing of a user profile. There is also a possibility of making customization less accurate.

“active transcoding” and is done dynamically without user intervention. Transcoding can be performed in both upstream and downstream directions. An implementation of this technique is MOWSER (Mobile Browser Project, 1996). MOWSER is an Apache proxy server agent written in Perl. MOWSER used proxy to perform transcoding. The incoming HTTP stream is modified by the proxy to include the capabilities and preferences of mobile users. The users' preferences and capabilities are stored in the server. Modification and update of preferences is done by a CGI form on a URL at a Web site maintained by the proxy. The proxy then fetches the files with the most suitable format to the requesting client. This implementation assumed that different formats are available for content adaptation. This is not an issue, as different formats can be created on the fly and cached in the server for future requests. Transcoding of images and videos is done using scaling, sub-sampling or sub key-frame techniques. Transcoding of HTML pages are done by eliminating unsupported tags and allowing users to select their preferences. This implementation, however, did not touch on the aspect of navigation. This technique, therefore, might not work well if adaptive navigation is required.

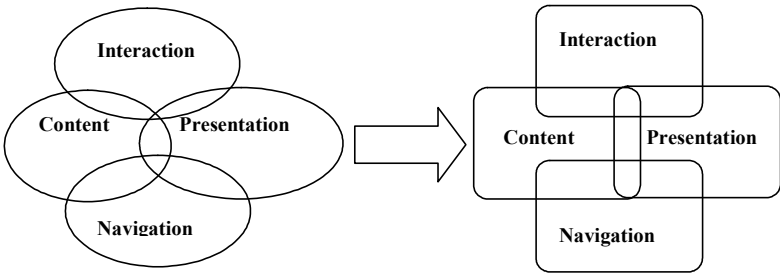
TRANSCODING

According to Bharadvaj, Anupam and Auephanwiriyakul (1998), modifying the HTTP streams and changing its content in situ is called

ANNOTATION-BASED CONTENT TRANSCODING

Annotation is a way to provide hints that enable a transcoding engine to make better decisions on con-

Figure 1. Desktop Web objects re-authored into mobile device-capable objects



6 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/chapter/web-content-adaptation-frameworks-techniques/17369](http://www.igi-global.com/chapter/web-content-adaptation-frameworks-techniques/17369)

## Related Content

---

### Efficient Large-Scale Stance Detection in Tweets

Yilin Yan, Jonathan Chen and Mei-Ling Shyu (2018). *International Journal of Multimedia Data Engineering and Management* (pp. 1-16).

[www.irma-international.org/article/efficient-large-scale-stance-detection-in-tweets/220429](http://www.irma-international.org/article/efficient-large-scale-stance-detection-in-tweets/220429)

### Applying Machine Learning in Optical Music Recognition of Numbered Music Notation

Fu-Hai Frank Wu (2017). *International Journal of Multimedia Data Engineering and Management* (pp. 21-41).

[www.irma-international.org/article/applying-machine-learning-in-optical-music-recognition-of-numbered-music-notation/182649](http://www.irma-international.org/article/applying-machine-learning-in-optical-music-recognition-of-numbered-music-notation/182649)

### Requirements to a Search Engine for Semantic Multimedia Content

Lydia Weiland, Felix Hanser and Ansgar Scherp (2014). *International Journal of Multimedia Data Engineering and Management* (pp. 53-65).

[www.irma-international.org/article/requirements-to-a-search-engine-for-semantic-multimedia-content/120126](http://www.irma-international.org/article/requirements-to-a-search-engine-for-semantic-multimedia-content/120126)

### Unified KS-Code

M. K.A. Abdullah, S. A. Aljunid, M. D.A. Samad, S. B.A. Anas and R. K.Z. Sahbudin (2009). *Encyclopedia of Multimedia Technology and Networking, Second Edition* (pp. 1473-1479).

[www.irma-international.org/chapter/unified-code/17572](http://www.irma-international.org/chapter/unified-code/17572)

### Multicast: Concept, Problems, Routing Protocols, Algorithms and QoS Extensions

D. Chakraborty, G. Chakraborty and N. Shiratori (2002). *Distributed Multimedia Databases: Techniques and Applications* (pp. 225-245).

[www.irma-international.org/chapter/multicast-concept-problems-routing-protocols/8624](http://www.irma-international.org/chapter/multicast-concept-problems-routing-protocols/8624)