

Fine-Grained Transclusion in the Hypertext Markup Language

Status of this Memo

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months. Internet-Drafts may be updated, replaced, or obsoleted by other documents at any time. It is not appropriate to use Internet-Drafts as reference material or to cite them other than as a "working draft" or "work in progress".

To learn the current status of any Internet-Draft, please check the `1id-abstracts.txt` listing contained in the Internet-Drafts Shadow Directories on `ds.internic.net` (US East Coast), `nic.nordu.net` (Europe), `ftp.isi.edu` (US West Coast), or `munari.oz.au` (Pacific Rim).

Distribution of this document is unlimited. Please send comments to the author at <xanni@xanadu.net>. This document is intended to become an informational RFC, and its content is designed for adoption in other standards and specifications.

Abstract

The word "hypertext" was coined by Theodor Holm Nelson in his paper "A File Structure for the Complex, the Changing and the Indeterminate", presented at the ACM 20th national conference in 1965. One of the key concepts in Nelson's vision of hypertext is "transclusion" or virtual inclusion, which permits composite documents to be constructed by reference to the original components rather than by copying.

The Hypertext Markup Language (HTML) is a markup language used to create hypertext documents that are platform independent. HTML currently permits the transclusion of various content types using tags which accept a "SRC" attribute, such as the ``, `<EMBED>` and `<APPLET>` tags, but does not provide a mechanism for transcluding textual content. This document proposes markup for text transclusions in HTML and explains its usage.

Introduction

The Hypertext Markup Language (HTML) [RFC1866] is a simple markup language used to create hypertext documents that are platform independent. This draft describes markup for an additional feature needed to support Nelson's vision of composite documents constructed by reference to the original documents. For more information, see <http://www.xanadu.net/xanadu/>

Syntax

The proposal is to add an HTML tag with the following syntax:

```
<TEXT SRC=(URI) {PLAIN} {RANGE=(start),(end)}>
```

Where parentheses () enclose variable parameters and braces {} enclose optional elements.

The SRC attribute is mandatory, and specifies the source document from which text is to be transcluded. (URI) must be the Universal Resource Identifier of a plain text or HTML document. If the source document is HTML and the optional PLAIN attribute is specified, all HTML tags are removed and all SGML entities converted to the characters they represent. If PLAIN is omitted, the source document is transcluded verbatim. In either case, only the contents of the <BODY> element of the source document is transcluded. If the source document is already plain text, the PLAIN attribute has no effect.

If the optional RANGE attribute is specified, only part of the source document is transcluded. (start) and (end) are integers representing character offsets from the beginning of the source document (immediately after the <BODY> tag for HTML) or negative character offsets from the end of the source document (immediately before the </BODY> tag for HTML). If (end) is zero, the source document is transcluded until the end.

Implementation

The intention is to have a facility in authoring programs that permits the author to create transclusions by indicating an insertion point, viewing the document from which they wish to transclude, and marking the region to be transcluded, much in the manner of a traditional "cut and paste" operation except that what is actually pasted is the reference to the transcluded portion rather than the literal text.

Initially, this could be a small editing program purely for adding transclusions to existing documents. It has also been suggested that

people might wish to add transclusions by hand, in which case it might be desirable to have other ways of specifying the start and end of the range besides just the byte offsets, which are inconvenient to determine by hand.

Expires 2 September 1997

[Page 2]

Possible Extensions

The following additional formats have been proposed for the (start) and (end) values:

1. HTML target anchors `` within the source document, indicated by a hash character followed by the anchor name. These can optionally be followed by a positive or negative character offset from immediately before the tag, for example `"RANGE=#start+5,#end-1"`.
2. Paragraphs within the source document, indicated by the letter P and the paragraph number counting from the beginning of the document. These can optionally be followed by a positive or negative character offset from the start of the paragraph, for example `"RANGE=P5+2,P9-3"`. A paragraph break is represented by one or more blank lines in a plain text document and the `<P>` tag in an HTML document.
3. Pattern matches within the source document, indicated by slashes or single or double quotes delimiting the pattern to match and optionally the letter I indicating a case insignificant match. Pattern matches were suggested by Paul Haeberli in his "Merge" script [[MERGE](#)]. These can optionally be followed by a positive or negative character offset from the start of the match, for example `"RANGE=/<ADDRESS>/i+3,'</ADDRESS>'i-5"`. Only the first match is used, and the transclusion is not performed at all if a match can not be found in the source document.

Acknowledgements

Paul Haeberli <paul@sgi.com>
Theodor Holm Nelson <ted@xanadu.net>
Yousuke Igarashi <yousuke@crew.sfc.keio.ac.jp>

References

- [MERGE] <http://reality.sgi.com/grafica/merge/>
- [RFC1866] T. Berners-Lee and D. Connolly, "Hypertext Markup Language - 2.0" ([RFC1866](#)), MIT/W3C, November 1995.

Author's Address

Andrew Pam
Project Xanadu
P.O. Box 26 East Melbourne VIC 8002 Australia

Phone: +61 3 9651 1511
Fax: +61 3 9651 1502
Email: xanni@xanadu.net

Expires 2 September 1997

[Page 3]