

August 3, 2006

Ian Lumb 8/22/06 5:05 AM

Comment: This is the date of the original blog post.

Google Notebook: A Tool for Annotation

Notes:

- This is an edited version of a post I made to my blog originally on August 3, 2006. The purpose of this document is to
 - Facilitate the editing process
 - Illustrate annotations of fragments using Microsoft Word
- No effort has been made to preserve the hypertext links that exist in the HTML version of the post that appeared online.

Ian Lumb 8/22/06 5:04 AM

Formatted: Bullets and Numbering

Filed under: Internet, Google, xml, Semantic Web, Standards, Global Geophysics, Annotation, Knowledge Representation — Ian Lumb @ 2:16 am Edit This

According to one source, “... annotation means adding information (such as notes, commentary, links to source material, and so on) to existing web-accessible documents without changing the originals.”

Ian Lumb 8/22/06 5:37 AM

Comment: WordPress' bookkeeping of tags, poster, time of post, and a link for editing the post.

Ian Lumb 8/22/06 5:09 AM

Comment: This is where the original content begins.

Wikipedia defines it this way: “Annotation is extra information associated with a particular point in a document or other piece of information.”

Ian Lumb 8/22/06 5:14 AM

Comment: In the online version, this text was a link to <http://www.manning.com/passin/>.

My introduction to the term occurred when I worked for Platform Computing. While I was there, we worked on several projects whose aim was to annotate genomes.

Ian Lumb 8/22/06 5:37 AM

Comment: In retrospect, this is sufficiently vague :-)

Later on, working with Keith Aldridge of York University, I became interested in annotating data from The Global Geodynamics Project (GGP). Briefly, GGP data is collected by instruments called Superconducting Gravimeters (SGs). These instruments are incredibly sensitive to relative changes in Earth's acceleration due to gravity. GGP is in fact a network of SGs that are distributed across the planet. Each instrument, each month provides three data files:

- Gravity and pressure data – The primary observables sampled regularly in time
- Auxilliary data – Complimentary obse_rvables (like groundwater levels) sampled regularly in time

Ian Lumb 8/22/06 5:19 AM

Comment: In this bulleted item, I corrected a spelling mistake via track changes. Because I changed the original, this correction doesn't qualify as an annotation.

- Log data – Other observables (like service events) that occur at irregular intervals

Ian Lumb 8/22/06 5:20 AM

Comment: I'm glad to see that I got the spelling right this time! BTW, this is an annotation.

When Keith and I first started work on introducing an XML-based data model for the GGP, we thought it made sense to automatically represent each of these files separately, and then combine them using XInclude or XSLT. (This XML-based data model is described in detail elsewhere.) After additional exposure to the XML spectrum, XPointer emerged as another option. (I alluded to this in a follow-up paper with Keith.) XML Pointer Language (XPointer) is "... the language to be used as the basis for a fragment identifier for any URI reference that locates a resource ..." It's not by coincidence that the Wikipedia definition for annotation reads similarly to this definition of XPointer from the W3C. XPointer is the W3C's vehicle for annotation.

Ian Lumb 8/22/06 5:23 AM

Comment: I may be guilty of stretching the truth a little here, as on reflection, the Wikipedia definition of annotation is sufficiently vague. For example, in the Wikipedia definition, what exactly does "... a particular point in a document ..." actually mean?

It's still my take that adoption of XPointer is rather slow. For example, the only Web browser that I know of that supports annotation is the W3C's Amaya. When browsing annotated Web pages with Amaya, clickable icons of pencils provide a link to the annotations.

The utility of annotations, taken together with the poor adoption of XPointer in mass-market Web browsers, makes for exciting possibilities for Google Notebook. What follows are a few data points on Google Notebook:

- "Google Notebook makes web research of all kinds – from planning a vacation to researching a school paper to buying a car – easier and more efficient by enabling you to clip and gather information even while you're browsing the web."
- Google Notebook lives in your Web browser
- Google Notebook organizes, in an easily accessible way, your Web findings
- Google Notebook is accessible from any Internetconnected computer
- Google Notebook is another prototype emerging from the Google Labs
- Google Notebook is searchable (of course, doh!)
- Google Notebook is publishable – click here for an example

Ian Lumb 8/22/06 5:39 AM

Comment: Clipping and gathering is useful. However, from the perspective of annotation, referencing Web pages is key.

Ian Lumb 8/25/06 6:28 AM

Comment: This is an example of a comment on a fragment that includes two paragraphs plus one item of a bulleted list. In the corresponding HTML document, the identified fragment crosses two paragraph <p>...</p> tags, plus an item ... in an unordered list XPointer allows such fragments to be identified.

In other words, Google Notebook is a tool for annotation!

Enthusiasm aside, I must note that:

- Google Notebook is not standards based – I assume it's based on proprietary Googleisms, but this requires further research
- Google Notebook is targeted at Mozilla Firefox only – It's enabled as a plugin
- Google Notebook makes use of a simple model of authentication – Username and password form the basis
- Google Notebook doesn't allow for Access Control Levels (ACLs) – Published notebooks are accessible by anyone who picks up on the lengthy, randomly generated URL

On balance, Google Notebook is useful, and allows for an excellent introduction to annotation. I highly recommend investigating Google Notebook, and providing your feedback to Google on it.

Ian Lumb 8/22/06 5:29 AM

Comment: One example of a standard Google Notebook *could* be based on is XPointer. XPointer is an emerging open standard from the W3C. I know there are other standards for annotation, even though I haven't spent any time looking at them.

Ian Lumb 8/25/06 6:29 AM

Comment: This is key. Annotation demands fragment identification. This is a different, albeit complimentary, requirement to clipping and gathering. Google Notebook supports fragment identification to a degree.

Ian Lumb 8/25/06 6:29 AM

Deleted: Google Notebook is not a fragment identifier – The degree of granularity is the Web page itself, not a fragment on a Web page

Ian Lumb 8/22/06 5:36 AM

Comment: In retrospect, this was a bit of a sweeping generalization.

Ian Lumb 8/22/06 5:09 AM

Comment: This is where the original content ends.