

Semantic Blogging



Steve Cayzer

Semantic Blogging for Bibliographies



- is a Hewlett-Packard Laboratories project in support of the European 'Semantic Web – Advanced Development' project
- Is intended to showcase the power of a new technology (the semantic web) in the context of a successful application (blogging)
- Can be summarised by the phrase

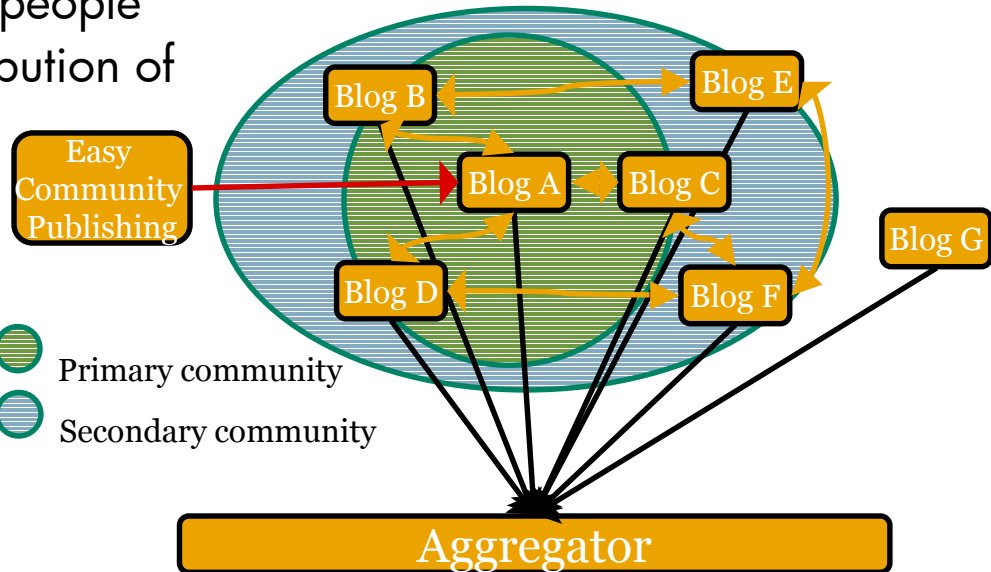
“Blogging is cool - but it could be even cooler”

Why is Blogging Cool?



Easy Publishing

- 'Push button publishing for the people'
- Encourages creation and distribution of useful information.



Community Formation

- blogrolling, aggregators, portals for coarse grain
- Linking, comments and trackbacks for fine grain

Aggregation

- RSS for a description of your published items
- Aggregators for a personalised 'mix & match' subscription mechanism

How could it be cooler?



- **Semantic View**

- More ways to view a set of blog items
- over and above date and topic based views
- eg top 10 rated items, summary tables by author...

- **Semantic Navigation**

- More ways to navigate around the blogosphere
- over and above unlabelled links
- eg 'who agrees/disagrees with me on this?'

- **Semantic Query**

- More ways to find relevant information
- over and above free text based searching of aggregators
- 'who is blogging on a semantically related topic?'
- 'who else has blogged this item?'

Semantic Blogging in a nutshell



- Our approach involves attaching *semantics*, or meaning, to web markup (*metadata*).
- To understand this process, it is helpful to consider an everyday example. English speakers tend to attach semantics to the word "cat", whereas the Viennese equivalent "katze" might simply be seen as an sequence of letters; an opaque symbol.
- The *semantic web* defines standards for attaching meaning to symbols in a way that computers can process. The symbols are mapped to concepts and relationships, which are formally described using an *ontology*.
- Using these principles, we are building blogging tools which associate symbols with appropriate concepts, and which use this understanding to provide enhanced services.

This is what we mean by *semantic blogging*

Semantic Web in a nutshell



The Semantic Web involves adding *semantic metadata* to pages, and using machines which process that markup to do cool stuff.

- **What's semantic metadata?**

- *Metadata* includes simple markup tags (like author & title). We give this metadata a well defined meaning (*semantics*) using *ontologies*.

- **What's an ontology?**

- A way of formally representing knowledge. A simple example would be a topic hierarchy.
- These ontologies are created & shared in a decentralized way – a common trait for web paradigms, such as blogging.
- But in the semantic web we are also able to share *meaning*.

Markup with Meaning



So, the semantic web:

- **uses metadata**

- markup (or metadata) is information about a blog item.
- eg author, title, topic...
- Using meaningful vocabularies (RDF).

- **adds meaning - by removing ambiguity**

- items (and ontology terms) are given a unique identifier (URI) so if we are talking about the same thing, we can pool the information
- eg I say 'blog', you say 'weblog': but we use the same URI

- **adds meaning – by enabling implicit mechanisms**

- inference using rules.

`"Blojsom is a subcategory of BloggingSoftware, therefore a search for articles on Blogging Software should include articles on Blojsom"`

The Semantic Blogging meme is spreading



- **People are encouraging the creation and use of blog metadata**
 - Enrich those RSS feeds!
 - Assisted markup using CSS, auto-categorization, NLP...
- **People are using blogs for rich information sharing**
 - k-logging, community formation, ThreadsML ...
- **People are talking about semantic blogging**
 - A number of blogs (including mine) specifically devoted to it, plus countless discussions in other fora.

Baby steps towards Semantic Blogging



- **We already have blog metadata**
 - But we need to progress beyond XML to semantic metadata
- **We can already use metadata to navigate**
 - Eg by (hierarchical) topic
 - But semantic navigation is much more than this...
- **We already have emergent ontology sharing**
 - How do we know we are talking about 'the same topic'?
Answer: The [TopicExchange](#).
 - But the semantic web vision also incorporates decentralized ontologies

The SWAD-E Semantic Blogging Project



- use the semantic web stack to enrich both blog contents (publication of structured data) and blog metadata
 - structured content in blog enables publication of data (e.g. event times, ratings, bibliographic info) which can then be aggregated and machine processed
 - channel and item classification metadata based on shared ontologies/thesauri
 - semantically tagged links between items (*extends, replaces, agreesWith* ..)
 - applied to small group bibliographic management

SWAD-E Project Background



- Semantic Web Advanced Development - Europe
- EU Framework 5 project, 30m, started May 2002
 - Members: W3C Europe, U. Bristol (ILRT), HP, Rutherford Appleton Lab, Stilo
 - Aims to put semantic web on a firm footing with current developers

More details at:

- SWAD-E: <http://www.w3.org/2001/sw/Europe/>
- Me: http://www-uk.hpl.hp.com/people/steve_cayzer
- Project Blog: <http://jena.hpl.hp.com:3030/blojsom-hp/blog>



i n v e n t