

## Institutional Repositories and Desktop Silos

The well known IT columnist, Jon Udell once wrote that “The Document is the Database” [1] and demonstrated a method of creating documents that web-enabled them from the outset. One of the challenges when adopting this approach in the context of Open Access is to make access to content both seamless and bi-directional within desktop content creation tools. WebDAV [2] is an XML-based standard for saving content over Internet protocols that goes back to 1999 and is supported in most operating systems. It is one of the few, if only, widely available conduits between desktop storage and web systems that does not depend on a web browser.

At the University of Windsor, one of the proposed collections for archiving and access is centered around photos taken by the official campus photographer. The photo collection can grow at a rate of 40 to 50 objects per day, and there is a high level of incidental metadata required for this type of content. A picture may be significant, for example, because it captures a particular individual at a specific function in a certain setting with a select group of dignitaries. A typical Institutional Repository (IR) setup, like that in DSpace [3], would require assigning Dublin Core [4] descriptors to the photos. Going through the “fill out the web form” and “attach and upload” process that is typically supported by IR systems is, at best, cumbersome for this type of content.

A WebDAV folder, on the other hand, allows the photos to be dragged and dropped into arbitrary folder layouts. Since every object has a URL in this scenario, metadata can be attached at the folder and/or object level as appropriate. The enabling technology for this sequence is Cocoon [5], an XML publishing framework from the Apache Foundation. Cocoon uses a system of “pipelines” to let content flow through transformations powered by XSLT [6] stylesheets and other tools. The end result is that any content delivered by WebDAV can be reworked and repurposed as appropriate.

We have developed a tool for assigning metadata to objects (see below) that describes objects and collections on the web. We have also targeted some specific desktop tools, like OpenOffice [7], an office suite that brings strong advantages to creating web-savvy documents. OpenOffice uses a publicly recognized XML standard for storing content and is currently adapting the OASIS OpenDocument [8] standard as its native syntax. Thanks to an ingenious file format structure, OpenOffice can work with WebDAV to blur the distinction between the desktop and the web, and allow collaborations to reach deep into the content creation process.



## E-Resource Admin

Namespace	Predicate	Value	Select
DC	Title	Agenda: User Group Me	<input type="checkbox"/>
DC	Resource Identifier	http://www.dspace.org/c	<input type="checkbox"/>
LOCAL	Link Creator	art	<input type="checkbox"/>
SUBJ	General	assigned by default	<input type="checkbox"/>
DC			
LOCAL			
SUBJ			

[Go to Admin List](#)

**Information Text**

Arial 7 (36 pt) Heading 1 **B** *I* U ~~S~~ \*x\* x\* ? i

**This** is an example of ~~good interesting~~ **rich text**.

Path: [body](#)

Related to this is work using the FOAF (Friend Of A Friend) [9] standard, an XML syntax for describing individuals and contacts on the web. In addition to identification, there have been some intriguing uses of FOAF to describe subject expertise, for example:

```
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:trust="http://www.perceive.net/schemas/20020725/trust#">

<foaf:Person>
<foaf:nick>pixel</foaf:nick>
<trust:trustsHighly>
<foaf:Person rdf:about="#John">
<rdfs:subject rdf:resource="http://dmoz.org/Computers/Programming/Languages/Perl/" >
</foaf:Person>
</trust:trustsHighly>
</foaf:Person>
</rdf:RDF>
```

In this example, FOAF is used to identify someone with expertise in the programming language Perl. A low-barrier method for describing expertise and a mechanism for informal trust networks could greatly enhance how content is created and shared on the web. In 1945, Vannevar Bush proposed a “Memex”[10], a tool for tracking and facilitating research, it seems appropriate, on the 60<sup>th</sup> anniversary of Bush’s famous proposal, that we break down the barriers between the desktop and the web, and foster new kinds of connections and collaborations based on the availability of a global network.

## References

- [1] <http://www.xml.com/pub/a/2003/07/09/udell.html>
- [2] <http://www.webdav.org>
- [3] <http://www.dspace.org>
- [4] <http://dublincore.org>
- [5] <http://cocoon.apache.org>
- [6] A good introduction to XSLT can be found in Wikipedia:  
<http://en.wikipedia.org/wiki/XSLT>
- [7] <http://www.openoffice.org>
- [8] <http://www.oasis-open.org/specs/index.php#opendocumentv1.0>
- [9] <http://www.foaf-project.org>
- [10] Bush’s article appeared in the July, 1945 issue of *The Atlantic Monthly*, one of several copies can be found here: <http://www.ps.uni-sb.de/~duchier/pub/vbush/vbush-all.shtml>