



XML Document Editing With Lemon8

Leading Edge Developments

PKP
PUBLIC
KNOWLEDGE
PROJECT

MJ Suhonos
Public Knowledge Project
First International PKP Scholarly Publishing Conference
Vancouver, Canada, July 12, 2007

Overview

- Background / Problem Domain
- Technology / Rationale
- Application Design
- Current Limitations
- Future Directions

What's this all about?

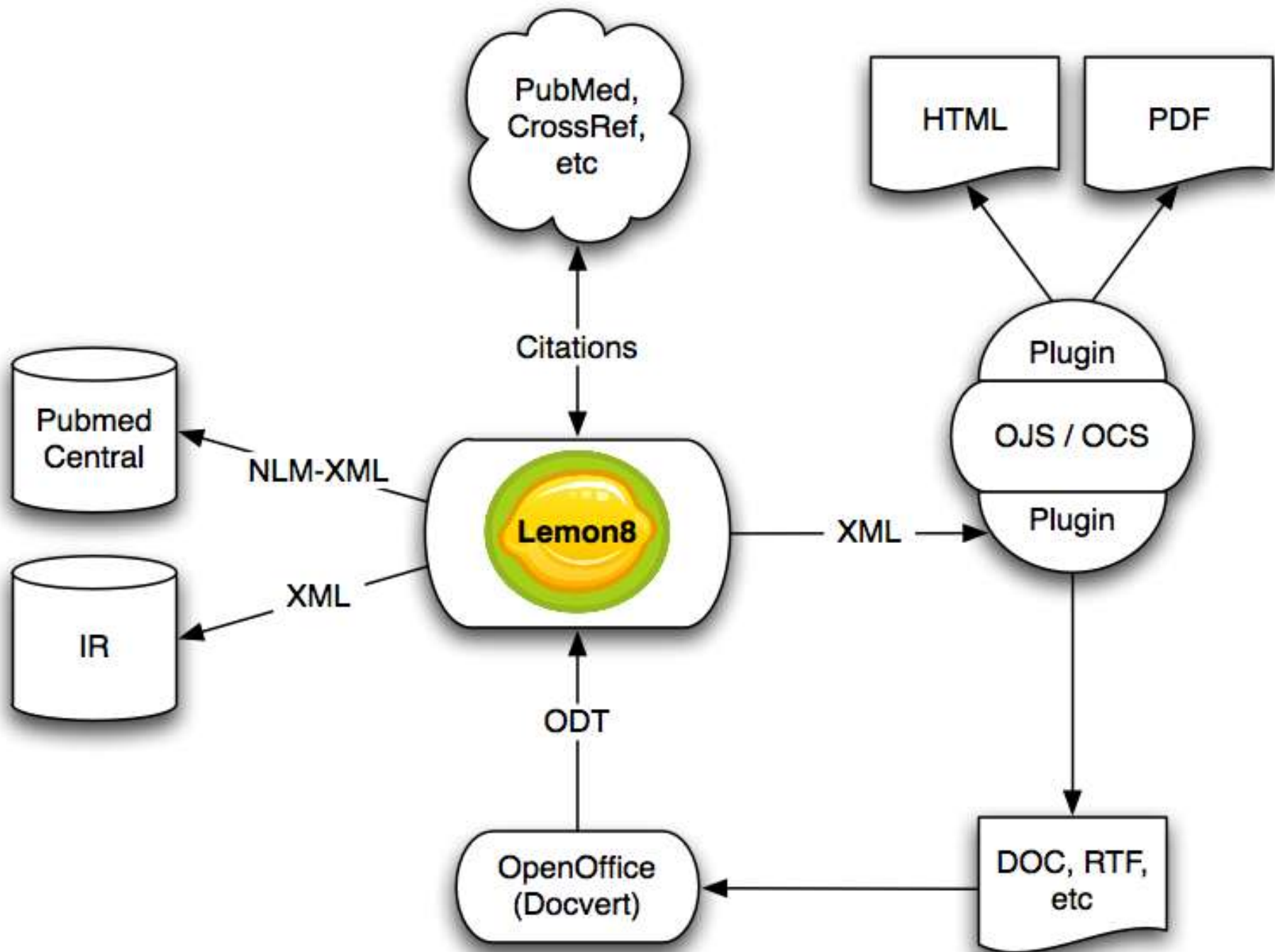
- Move to an XML-based layout system
 - Render XHTML and PDF dynamically
 - Support digital archiving (eg. Pubmed Central)
 - Add semantic markup
- Benefits for non-techies
 - Reduced effort for layout, editing, galleys
 - Free alternative to commercial applications
 - eg. eXtyles (~\$20K + \$1500 / journal) for FM

A great idea, but...

- Leverage existing FOSS tools
 - OpenOffice.org + Docvert
 - LAMP + CakePHP (PHP5 for DOM/XSL)
- “Best guess” parsing with human correction
- Help with hard things (eg. citation checking)
- Intuitive UI: “5 steps to an XML document”
- Schemaless metadata (NLM-JP for now)

“XML Never Tasted So Good”







1. Metadata



2. Components



3. Citations



4. Preview/Save

Section 1 (92 words)	<input type="text" value="Introduction"/>	
Section 2 (80 words)	<input type="text" value="Methods"/>	
Section 3 (84 words)	<input type="text" value="Subheading"/>	
Section 4 (79 words)	<input type="text" value="Subheading"/>	
Section 5 (281 words)	<input type="text" value="Results"/>	
Section 6 (79 words)	<input type="text" value="Subheading"/>	
Section 7 (0 words)	<input type="text" value="Subheading"/>	
Section 8 (79 words)	<input type="text" value="Sub-subheading"/>	
Section 9 (79 words)	<input type="text" value="Discussion"/>	
Section 10 (79 words)	<input type="text" value="Conclusions"/>	



1. Metadata



2. Components



3. Citations



4. Preview/Save

11 citations (82% edited, 64% looked up) [Lookup All Citations](#)

1. Worldwide PDA & Smartphone Forecasts Report: 1998-2008. eTForecasts 2003.

Edited

2. Martin S. MD's computer, PDA use on the upswing. CMAJ 2002;167(7):794.

Looked up

3. Fischer S, Stewart TE, Mehta S, Wax R, Lapinsky SE. Handheld computing in medicine. J Am Med Inform Assoc 2003;10(2):139-149. [Medline]

Looked up

4. Koop A, Moesges R. The use of handheld computers in clinical trials. Control Clin Trials 2002;23(5):469-480.

Looked up

5. Lapinsky S, Weshler J, Varkul M, Mehta S, Stewart T. Evaluation of handheld computers in critical care. In: Bruch HP, Kockerling F, Bouchard R, Schug-Pab C, editors. New Aspects of High Technology in Medicine. Bologna, Italy: Monduzzi Editore; 2000:279-284.

Edited

6. Benbasat I, Zmud RW. Empirical research in information systems: the practice of relevance. MIS Quarterly 1999;23(1):3-16.

Looked up

7. Baumgart DC. Personal digital assistants in health care: experienced clinicians in the palm of your hand? Lancet 2005; 366(9492):1210-1222.

Looked up

8. Gautheir, B (Circum Network Inc). Assessing survey research: a principled approach (work-in-progress report). In: Proceedings of the Professional Marketing Research Society Annual Conference; April 2001; Ottawa, Ontario, Canada:1-28.

Not Edited



[Return to Document](#)

Reference:

[Fill out form](#)

Baumgart DC. Personal digital assistants in health care: experienced clinicians in the palm of your hand? Lancet 2005;366(9492):1210-1222.

Copy and paste or type in the reference as it may appear in the References section of a paper. For example, Shah, S. (2006). A new kind of bibliography. Journal of Bibliography, 4(3), 1-24.

[Lookup citation](#)

Attempt to locate the citation in a number of online databases, including: PubMed, CrossRef, and Amazon.com. If found, the form will automatically be filled with the correct values.

Author(s):
(required)

Baumgart Daniel C;

Required unless there are editors or an institution, the author(s) of the work.

Year:
(required)

2005

The year of the work. If unknown, write n.d.. If forthcoming, try to predict the year of publication. Otherwise, write forthcoming.

Title:
(required)

Personal digital assistants in health care: experienced clinicians in the palm of your hand?

The title of the work.



1. Metadata



2. Components



3. Citations



4. Preview/Save

[View PDF Preview](#)

[Export in NLM-XML](#)

. NaNN ;

Introduction

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Cras sed leo nec ligula adipiscing lacinia. Duis turpis. Phasellus id ligula pulvinar erat tempus ultrices. Suspendisse eleifend, sem at hendrerit scelerisque, enim ante nonummy urna, ac consequat ligula sem in augue. Fusce venenatis. Fusce pellentesque vestibulum mauris. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus [1]. Ut vitae ante quis tellus eleifend hendrerit. Sed a mi facilisis lectus placerat eleifend. Praesent feugiat libero eget nisl nonummy commodo. Aliquam nec neque. Donec magna nisi, luctus eu, pharetra vitae, commodo in, leo [2].

Methods

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Cras sed leo nec ligula adipiscing lacinia. Duis turpis. Phasellus id ligula pulvinar erat tempus ultrices. Suspendisse eleifend, sem at hendrerit scelerisque, enim ante nonummy urna, ac consequat ligula sem in augue. Fusce venenatis. Fusce pellentesque vestibulum mauris. Cum sociis natoque penatibus et

Current Limitations

- Substantial installation requirements
- Document parser could be more robust
- Limited by OpenOffice (eg. EndNote, WMF)
- Citation handling only for Vancouver style
- Only exports to NLM DTD

Future Directions

- Actively developing with Open Medicine
- Planned enhancements:
 - Integration with OJS, OCS via plugins
 - Liberal addition of AJAX
 - More citation lookup services
 - Additional XML schemas for export

Thank You!

“98.5% of DNA is considered to be junk DNA with no known purpose.”

“Maybe it's XML tags.”

