#### NRC-CNRC

## NRC Publications Archive Archives des publications du CNRC

#### Federated Searching in an Age of Web Services

Newton, Glen

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. / La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

#### Publisher's version / Version de l'éditeur:

Access 2005 [Proceedings], 2005

NRC Publications Archive Record / Notice des Archives des publications du CNRC : <a href="https://nrc-publications.canada.ca/eng/view/object/?id=d9c5b6fb-2485-450e-827b-f707b0318936">https://nrc-publications.canada.ca/eng/view/object/?id=d9c5b6fb-2485-450e-827b-f707b0318936</a> <a href="https://publications-cnrc.canada.ca/fra/voir/objet/?id=d9c5b6fb-2485-450e-827b-f707b0318936">https://publications-cnrc.canada.ca/fra/voir/objet/?id=d9c5b6fb-2485-450e-827b-f707b0318936</a>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at <a href="https://nrc-publications.canada.ca/eng/copyright">https://nrc-publications.canada.ca/eng/copyright</a>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site <a href="https://publications-cnrc.canada.ca/fra/droits">https://publications-cnrc.canada.ca/fra/droits</a>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

**Vous avez des questions?** Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.







# Federated Search in an Age of Web Services

**Glen Newton** CISTI Research

National Research Council Canada

Michael Fortin, Concordia University

Access 2005

October 17-19 2005







### **Outline**

- Original Problem
- Initial solution
- Expermentation
- Implications
- Conclusion



## **Two Important Points**

This is a talk on federated search



## **Two Important Points**

- This is a talk on federated search
- This is NOT a talk on federated search



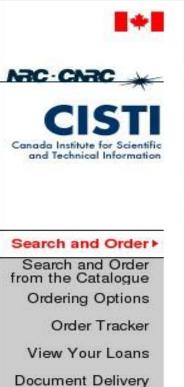
# Original Problem Part 1

- Canadian government has standard look & feel for web pages, known as "Common Look and Feel" (CLF)
- Includes accessibility, design, navigations, official languages, metadata, etc.
- All web sites must comply (limited exceptions)
- Sometimes difficult to make compliant



# Original Problem Part 1

- CISTI has web interface to Catalogue (Innovative)
- Not CLF compliant



Change Your Client

NRC Research.

Profile

Information Specialists Health Libraries

Media >

About CISTI >

National Research Conseil national Council Canada de recherches Canada



Français	Contact Us	Help	Search	Canada Site
CISTI	RP Home	DocDel	CISTI Source	NRC Site

GO to:	Select a search page	<u> </u>
	51	

#### **Special Notice**

On the Monday of the Canadian Thanksgiving long weekend, October 10, 2005, CISTI's Direct Supply service orders will be filled as usual. Client services and Urgent services will not be available. Thank you for your understanding.

#### Search and Order from the Catalogue

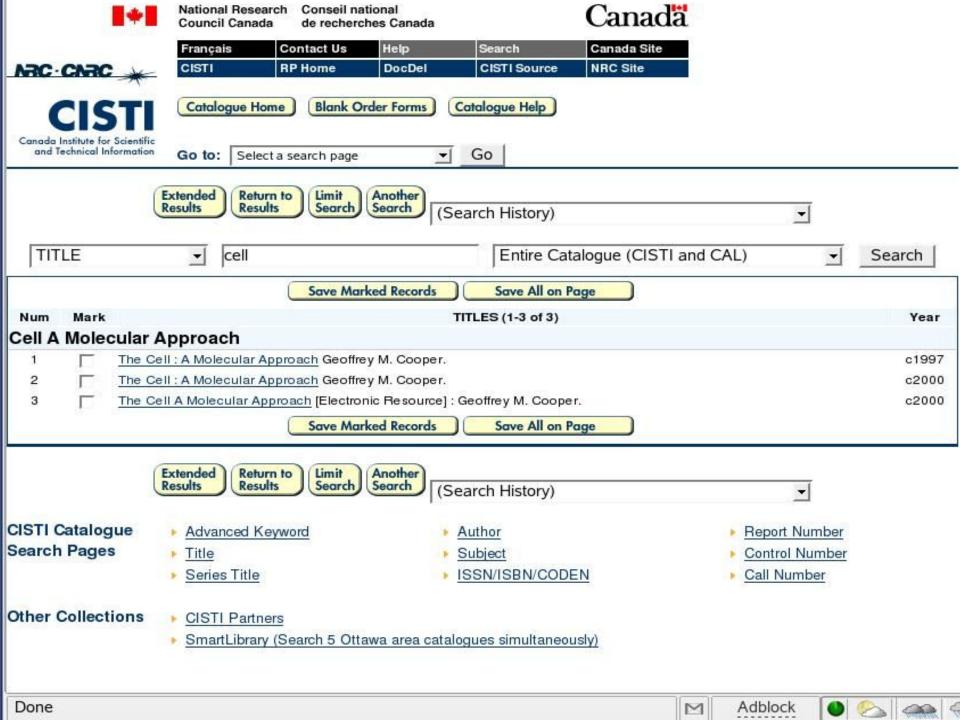
The Catalogue allows you to search the <u>CISTI collection</u> and that of its partner, the Canadian Agriculture Library (<u>CAL</u>). To order an item, access the full record display of that item and click on "Order this Item." New users must register before ordering.

Search type:	Title begins with	_	
Search terms:			
Catalogue subset:	Entire Catalogue (CISTI	and CAL)	<u>•</u>













# Original Problem Part 2

- NRC has adopted a commercial content management system (CMS) for managing web sites (Interwoven TeamSite)
- Proprietary, but relatively open, using XML & other standards
- Web sites & applications to migrate to CMS
- CLF templates
- While the CMS understands HTML, XML, JSP, ASP, PHP, servlets, etc. did not understand the catalogue web interface
- How to migrate the Web view of the catalogue to CLF and into the CMS



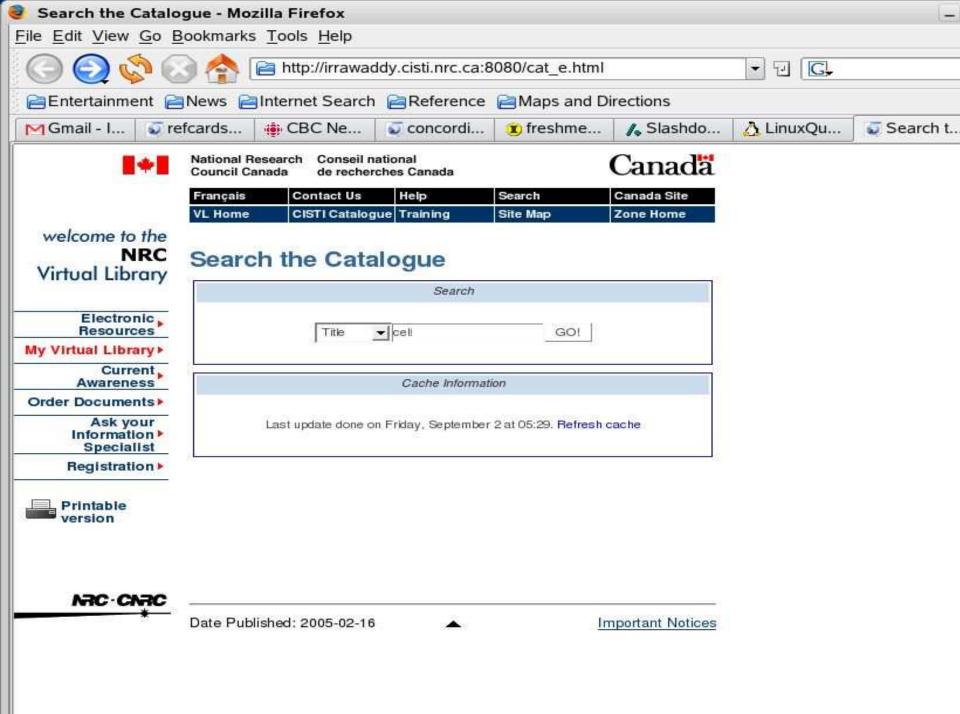
#### **Initial Solution**

- "Bridging" Web application which the CMS would understand, and which could talk to catalogue
- PHP 5
- Used Z39.50 to query catalogue
- Could be embedded into CMS
- Due to Z39.50 performance & connection dropping issues, used YAZ Z39.50 accelerating & caching Z39.50 proxy
- Implemented using Linux, Apache, PHP 5



# Initial Solution (Problems)

- Z39.50 problems: server could not handle sustained, but not high level, of queries
- Connections would be dropped, new connection attempts would time-out
- Due to Z39.50 performance & connection dropping issues, used YAZ accelerating & caching Z39.50 proxy









# Initial Solution Completed

- Successfully showed that a view of the Catalogue could be moved to both the CMS and to CLF
- Altered ("cleaned up?") web GUI
- Limitations:
  - Could not limit to specific collections (only electronic, etc)
  - Could not build patron web GUI using Z39.50



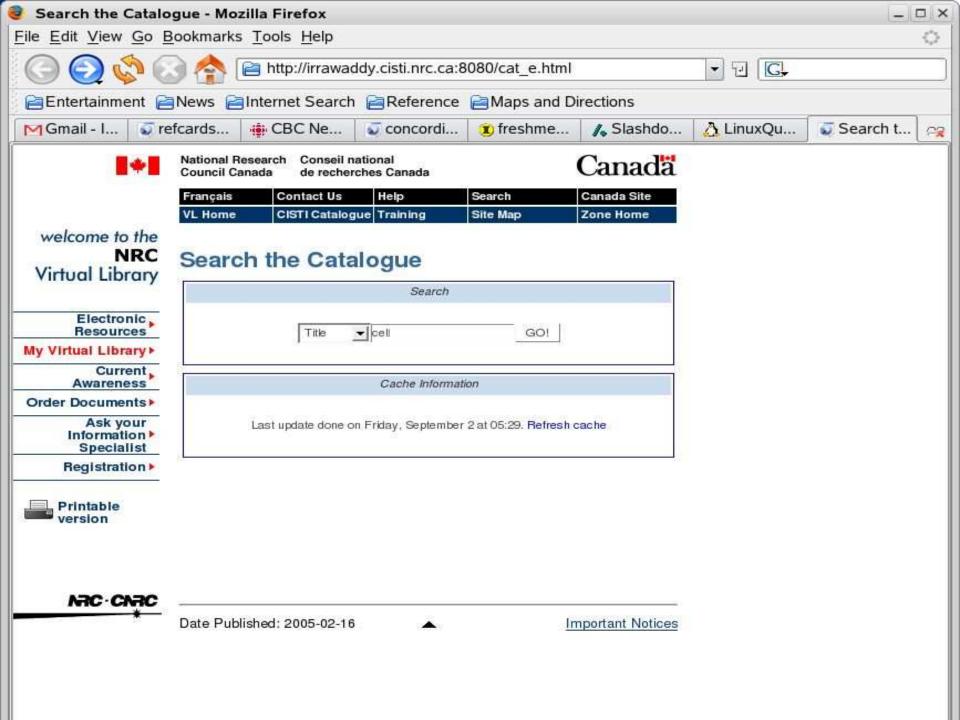
### What next?

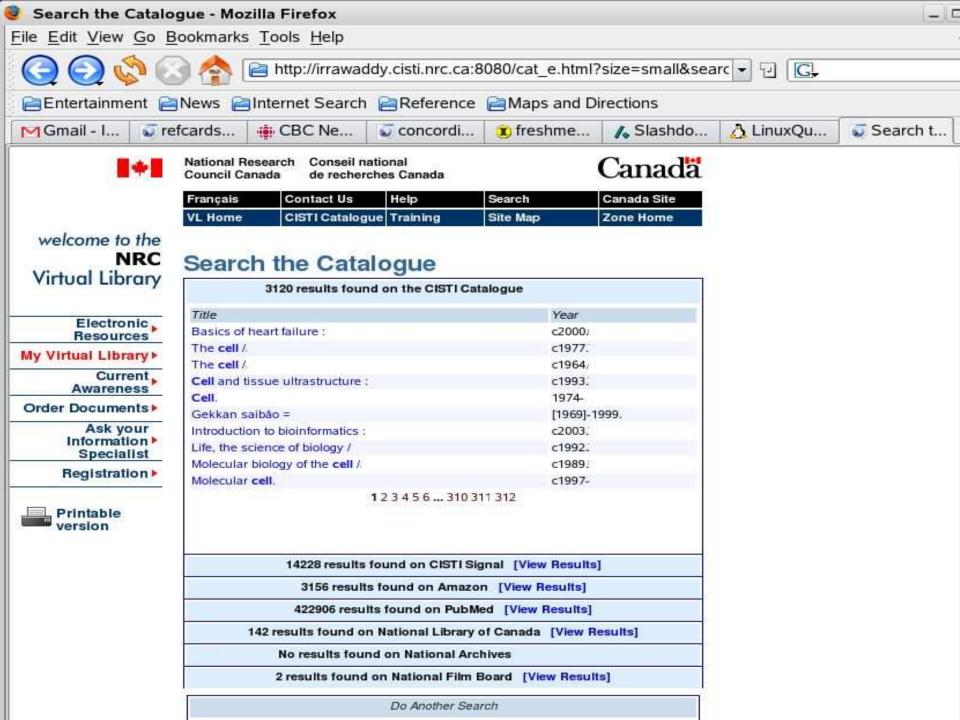
- Having the application built using PHP, this allowed for easy extensions
- AJAX (Asynchronous Javascript and XML) ?
- Other Z39.50 databases?
- Other non-Z39.50 databases?



### **Experimentation**

- Having the application built using PHP, allowed for easy extensions
- First extension was to move to AJAX (Asynchronous Javascript and XML)
- Search of other Z39.50 databases
- Search of PubMed using HTTP REST
- Search of Amazon using Web Services
- Use Web Services to get cover art from Amazon
- Search of our own metadata database (MySQL)



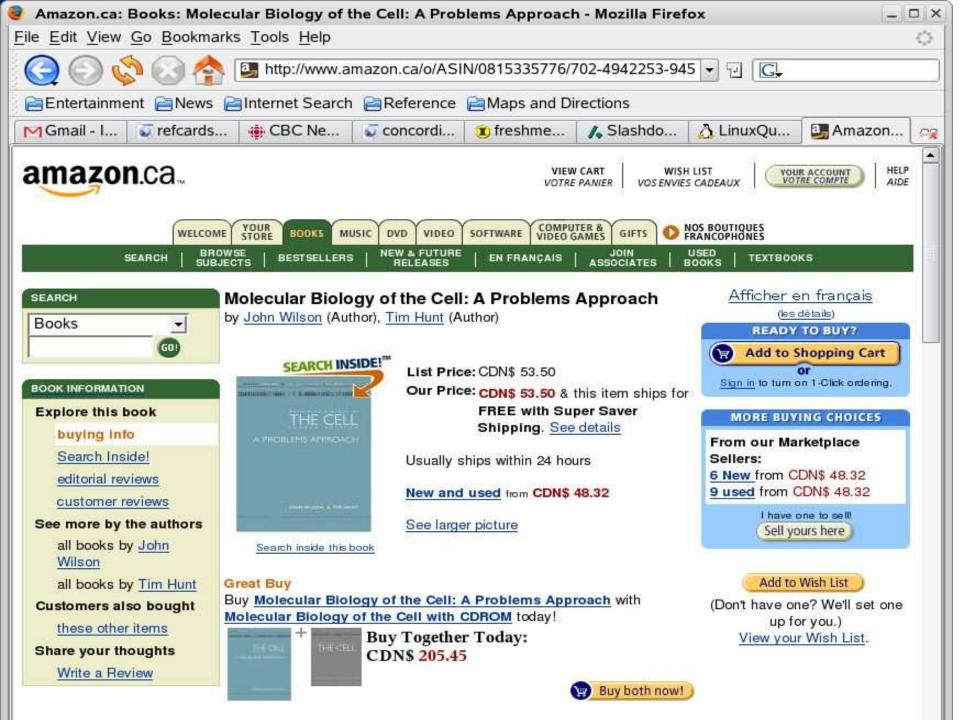




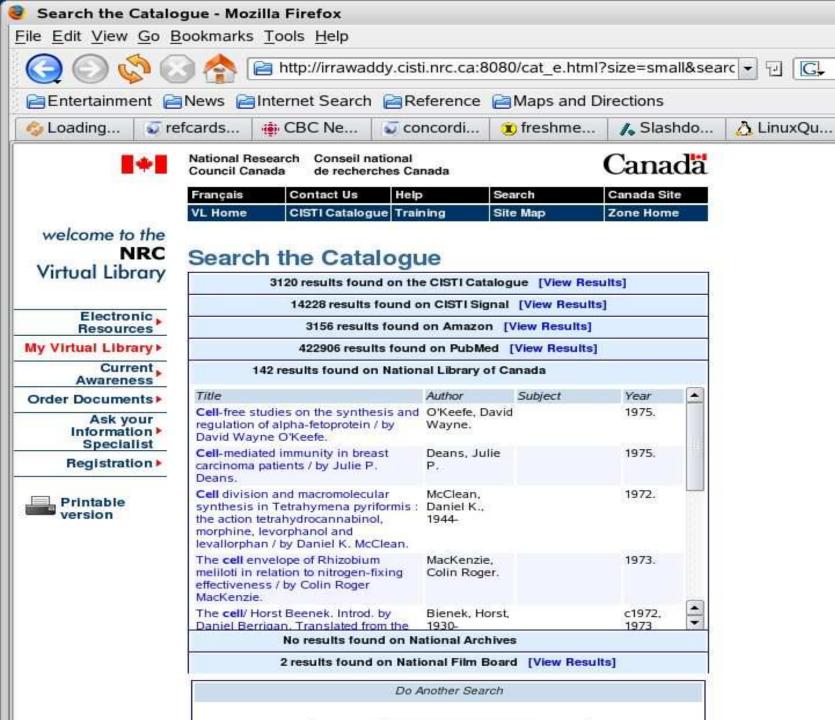




Search







Search t...



# Creating good user interfaces for search results

- 1) One source containing one kind of metadata: very easy
- 2) 5-8 sources containing one or a small number of kinds of metadata: **fairly easy** (we did this)
- 3) Many sources containing many different kinds of metadata: heavy lifting involved

### **AJAX**

- Asynchronous Javascript And XML
- Important step in Web-based applications becoming more desktop GUI-like, with finer level of interaction granularity (no more "full-in-form-and-submit" entire page update cycle)
- Google maps best known AJAX-driven application
- While initially rather complicated to implement applications using AJAX, toolkits are becoming mature for most programming languages making implementation very easy and flexible



### **Amazon Web Services**

- Very rich Web Services API allowing searching by various metadata (ISSN, title, author, etc)
- Ability to request cover art for books, CDs, etc. via API
- Additional Amazon information also available, like reviews, ratings, wish lists, etc.
- If you become an Amazon partner, any users who make purchases at Amazon, you get a cut from the purchase



### Amazon Web Services: *Caveats*

- Make sure you read the Amazon Web Services Licensing Agreement
- Peace Palace Library, The Netherlands



## **Implications**

- Explosion of Web Services by many disparate players beyond the initial few (Google API, Google Scholar API, Amazon, etc)
- M2M becomes basis of applications
- Services Oriented view of the universe
- Z39.50 history of library world puts it in good standing
- (Yet still more!) legal & licensing issues with which to deal
- Rich applications with better & more responsive GUI presenting users with information from a number of

#### Conclusion

- Able to build prototype embedding Web view of catalogue into CMS & compliant to CLF
- Open standards helped(Z39.50 to communicate to Catalogue;
   XML & others with CMS; Web Services)
- Easily implemented using Open Source (PHP, Apache, Linux)
- Issues also resolved using Open Source & open standards (YAZ Z39.50 proxy)
- Implications of proliferation of Web Services and AJAX signify major change in the Web



### **Thank You**

**Glen Newton** CISTI Research

glen newton at nrc-cnrc.gc.ca

Michael Fortin, Concordia University

NRC CISTI – Canada's National Science Library and Publisher: http://cisti-icist.nrc-cnrc.gc.ca/

CISTI Lab: http://lab.cisti-icist.nrc-cnrc.gc.ca/





#### References

- AJAX: http://en.wikipedia.org/wiki/AJAX
- Amazon Web Services: http://www.amazon.com/gp/aws/landing.html
- Canadian Government Common-Look-and-Feel standards and guidelines: http://www.tbs-sct.gc.ca/clf-nsi/index e.asp
- Innovative Interfaces: http://www.iii.com/
- Interwoven Teamsite CMS: http://www.interwoven.com/
- PHP: http://www.php.net/
- PubMed scripting page: http://eutils.ncbi.nlm.nih.gov/entrez/query/static/eutils\_help.html
- YAZ Z39.50 proxy: http://www.indexdata.dk/yazproxy/



## Presentation produced using an Open Source office suite, Open Office

http://www.openoffice.org/







## **Acknowledgements**

- Danielle Langlois, Hans Vulker, NRC Web Office
- Greg Kresko, CISTI Research