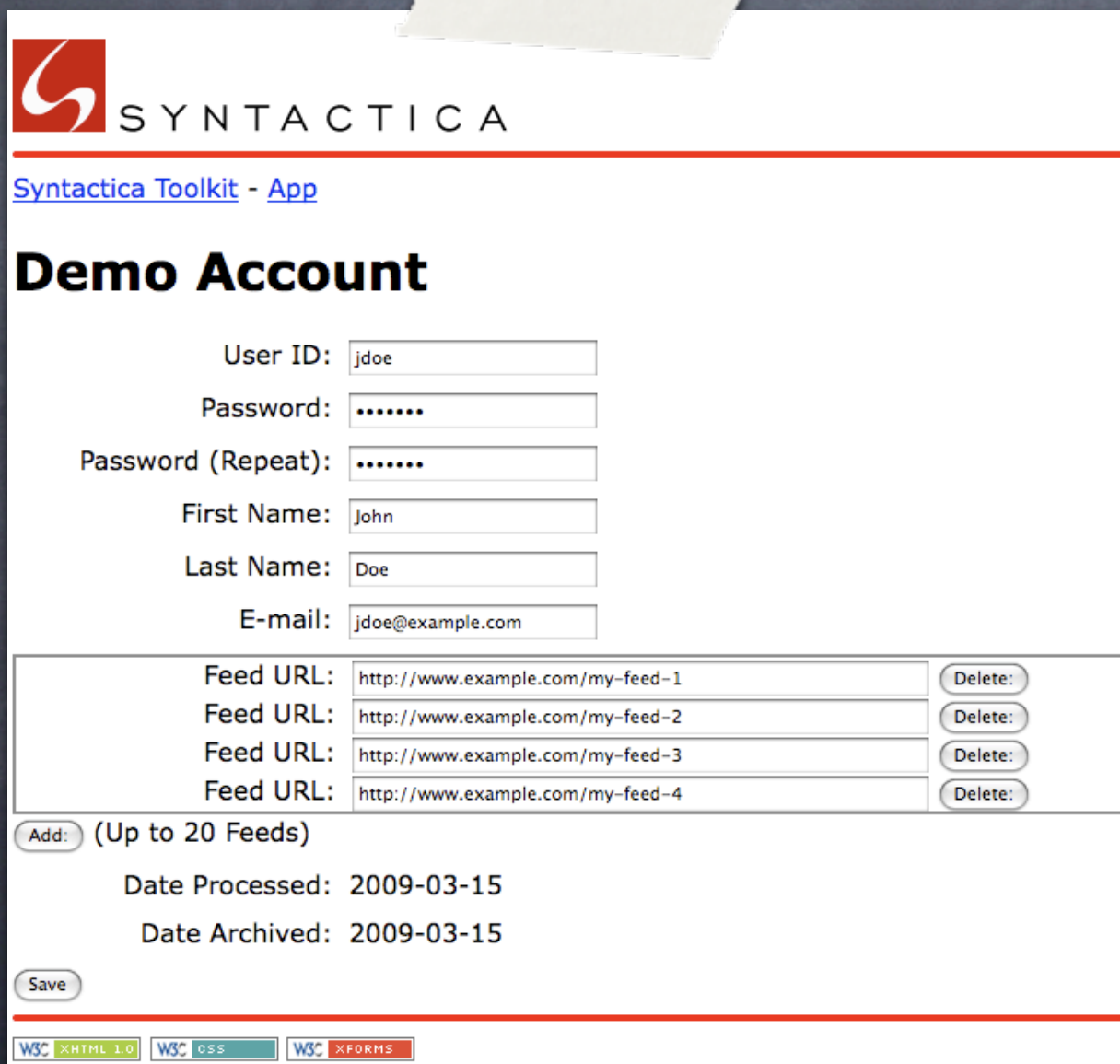


Rome Feed Processing

Technologies

• UIMA	
• Syntactica	
Weighted	
Corellation	
• Rome	
• Quartz	
• eXist	
• HtmlCleaner	

RSS/Atom Demo



The screenshot shows a web form for creating a demo account. At the top is the Syntactica logo and the text 'SYNTACTICA'. Below it is a link 'Syntactica Toolkit - App'. The main heading is 'Demo Account'. The form contains several input fields: 'User ID' (filled with 'jdoe'), 'Password' (filled with dots), 'Password (Repeat)' (filled with dots), 'First Name' (filled with 'John'), 'Last Name' (filled with 'Doe'), and 'E-mail' (filled with 'jdoe@example.com'). Below these is a table with four rows, each containing a 'Feed URL' and a 'Delete' button. The feed URLs are 'http://www.example.com/my-feed-1', 'http://www.example.com/my-feed-2', 'http://www.example.com/my-feed-3', and 'http://www.example.com/my-feed-4'. Below the table is an 'Add:' button followed by '(Up to 20 Feeds)'. Below that are two labels: 'Date Processed: 2009-03-15' and 'Date Archived: 2009-03-15'. At the bottom left is a 'Save' button. At the very bottom are three small icons: 'W3C XHTML 1.0', 'W3C CSS', and 'W3C XFORMS'.

SYNTACTICA

[Syntactica Toolkit - App](#)

Demo Account

User ID:

Password:

Password (Repeat):

First Name:

Last Name:

E-mail:

Feed URL:	<input type="text" value="http://www.example.com/my-feed-1"/>	Delete:
Feed URL:	<input type="text" value="http://www.example.com/my-feed-2"/>	Delete:
Feed URL:	<input type="text" value="http://www.example.com/my-feed-3"/>	Delete:
Feed URL:	<input type="text" value="http://www.example.com/my-feed-4"/>	Delete:

Add: (Up to 20 Feeds)

Date Processed: 2009-03-15

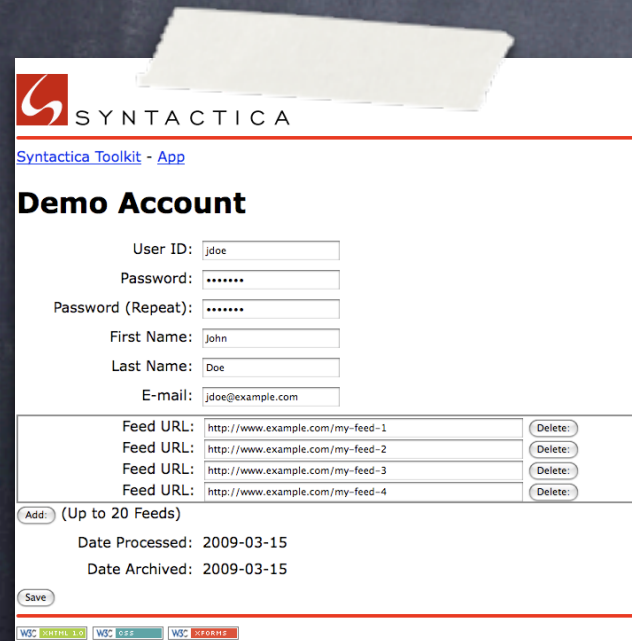
Date Archived: 2009-03-15

Save

W3C XHTML 1.0 W3C CSS W3C XFORMS

- Create the demo account
- place each feed URL into the processing queue
- Once all feed URLs are processed, then email user with link to the processed data

Queue Processing



SYNTACTICA
Syntactica Toolkit - App


Demo Account

User ID:
Password:
Password (Repeat):
First Name:
Last Name:
E-mail:

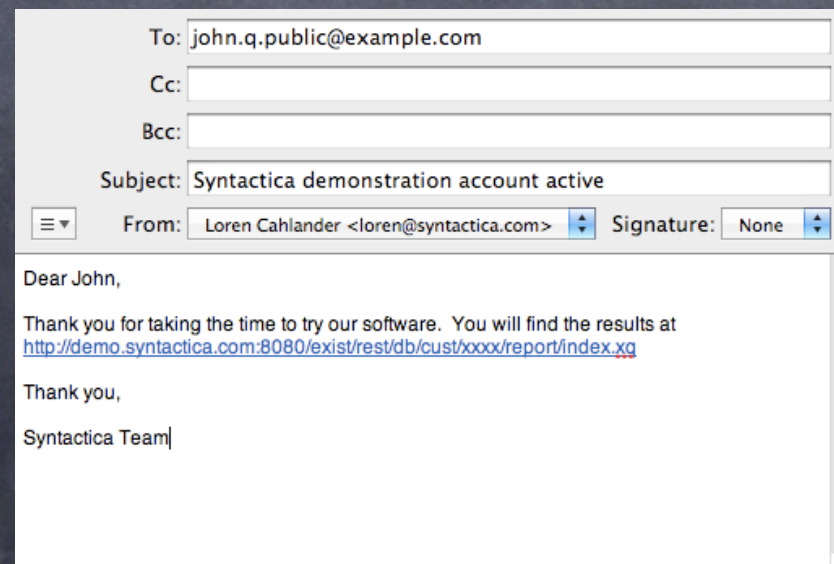
Feed URL:	<input type="text" value="http://www.example.com/my-feed-1"/>	<input type="button" value="Delete"/>
Feed URL:	<input type="text" value="http://www.example.com/my-feed-2"/>	<input type="button" value="Delete"/>
Feed URL:	<input type="text" value="http://www.example.com/my-feed-3"/>	<input type="button" value="Delete"/>
Feed URL:	<input type="text" value="http://www.example.com/my-feed-4"/>	<input type="button" value="Delete"/>

(Up to 20 Feeds)

Date Processed: 2009-03-15
Date Archived: 2009-03-15



RSS/Atom
RSS/Atom
RSS/Atom
RSS/Atom
RSS/Atom






To: john.q.public@example.com
Cc:
Bcc:
Subject: Syntactica demonstration account active
From: Loren Cahlander <loren@syntactica.com> Signature: None

Dear John,

Thank you for taking the time to try our software. You will find the results at <http://demo.syntactica.com:8080/exist/rest/db/cust/xxxx/report/index.xq>

Thank you,
Syntactica Team



To:	john.q.public@example.com		
Cc:			
Bcc:			
Subject:	Syntactica demonstration account active		
	From:	Loren Cahlander <loren@syntactica.com> 	Signature: None 

Dear John,

Thank you for taking the time to try our software. You will find the results at <http://demo.syntactica.com:8080/exist/rest/db/cust/xxxx/report/index.xg>

Thank you,

Syntactica Team|

Feed Processing

- Feed process array of size N threads
- Array size can be changed through an mbean
- When the feed process is invoked, the array is scanned for empty or completed slots
- If an available slot is found then the first RSS in the queue is added as a new FeedProcessor thread

FeedProcessor

- Use Rome processing code to get the feed header and the list of entries
- Each entry is processed through the threaded Entry process.
- Once all entries are processed, then the FeedProcessor thread is terminated.

Entry Processing

- Entry process array of size N threads
- Array size can be changed through an mbean
- The list of entries is processed by scanning the array for empty or completed slots
- If an available slot is found then the entry is added as a new EntryProcessor thread

EntryProcessor

- Add SyndFeed data to root
- Add SyndEntry data to root
- Get filtered text from entry server and get UIMA annotations
- Get abstract from Syntactica server
- Get weighted abstract from Syntactica server
- Get weighted subject from Syntactica server

SyndFeed

SyndEntry

Sofas

Text

Abstract

WAbstract

WSubject

EntryProcessor

