ticTOCron: An Automatic Solution for Propagating Quality Metadata to Scholarly TOC RSS Feed Metadata

Santy Chumbe
Heriot Watt University
S.Chumbe@hw.ac.uk
This talk is about:

- Getting good use out of TOC RSS feeds
- Enabling the reuse of TOC RSS feeds

TOC: Table of Content
RSS: Really Simple Syndication
Presentation Plan

- The “Context”
- “Problems”
- “Our Approach”
- How it works
- Use Cases
- Questions
“Problem” 1: Scientific Publishing

5.5 million researchers worldwide

1.4 million papers written annually

24,000 journals from more than 2,000 publishers

3.5% + journals/year
3% + researchers/year
“Problem” 2: Technology
“Problem” 3: Usage (quantity)

16,000 RSS TOC Feeds
“Problem” 3: Usage (admin)

- Need of different “RSS Readers”
- Maintaining your own Feeds
- It can take up time to browse and search RSS feeds from different sources

It used to be the “Lost and Found” Department
“Problem” 4: Format Inconsistencies

- RSS 0.90
- RSS 0.92
- RSS 1.0
- RSS 1.1
- RSS 2.0
- Atom

- Title
- Authors
- Abstract
- Journal
- Vol, Issue
- Date
- Pages
- Publisher
- URL
- DOI
- News
- Jobs
“Our Approach” 1: Define a “Best Practice” schema

- Title
- Authors
- Abstract
- Journal
- Vol, Issue
- Date
- Pages
- Publisher
- URL
- DOI
- News
- Jobs
“Our Approach” 2:
Implement the “Best Practice”

From a diversity of feeds to a normalized feed
How it works 1:
Based on Berson’s Methodology and Metadata Quality Propagation Techniques

TOC RSS Feeds
(13,000)

WWW

Harvesting
Cleansing
Auditing
Normalizing

"Know-how" human input

"ticTOCs Best Practices" metadata set
How it works 2: The Big Picture

ticTOCron toolkit system context
Use Case 1: Alerting IR Managers

- Facility to alert IR managers when a submitted paper from their IR has been published in a scholarly journal.
Use Case 2: Localised TOC RSS for Libraries

- service or facility to generate a localised version of ticTOCs for an Institutional Library
Use Case 3:
Subject rr Community based TOCs

- vortal

Subject based portal
Thank you!

Any questions?