

A minimum effort distributed library for KMi

(HackLab Project #3)

**Tom Heath and Mark Gaved
Knowledge Media Institute
The Open University, UK**

ESWC Workshop on Scripting for the Semantic Web

Problem scenario: Tom couldn't find a book...



“Does anyone have a copy of *‘Readings in Nonmonotonic Reasoning’*?”

Question: how can we find out who has a particular book on their shelf in KMi?

Answer: we need a system!

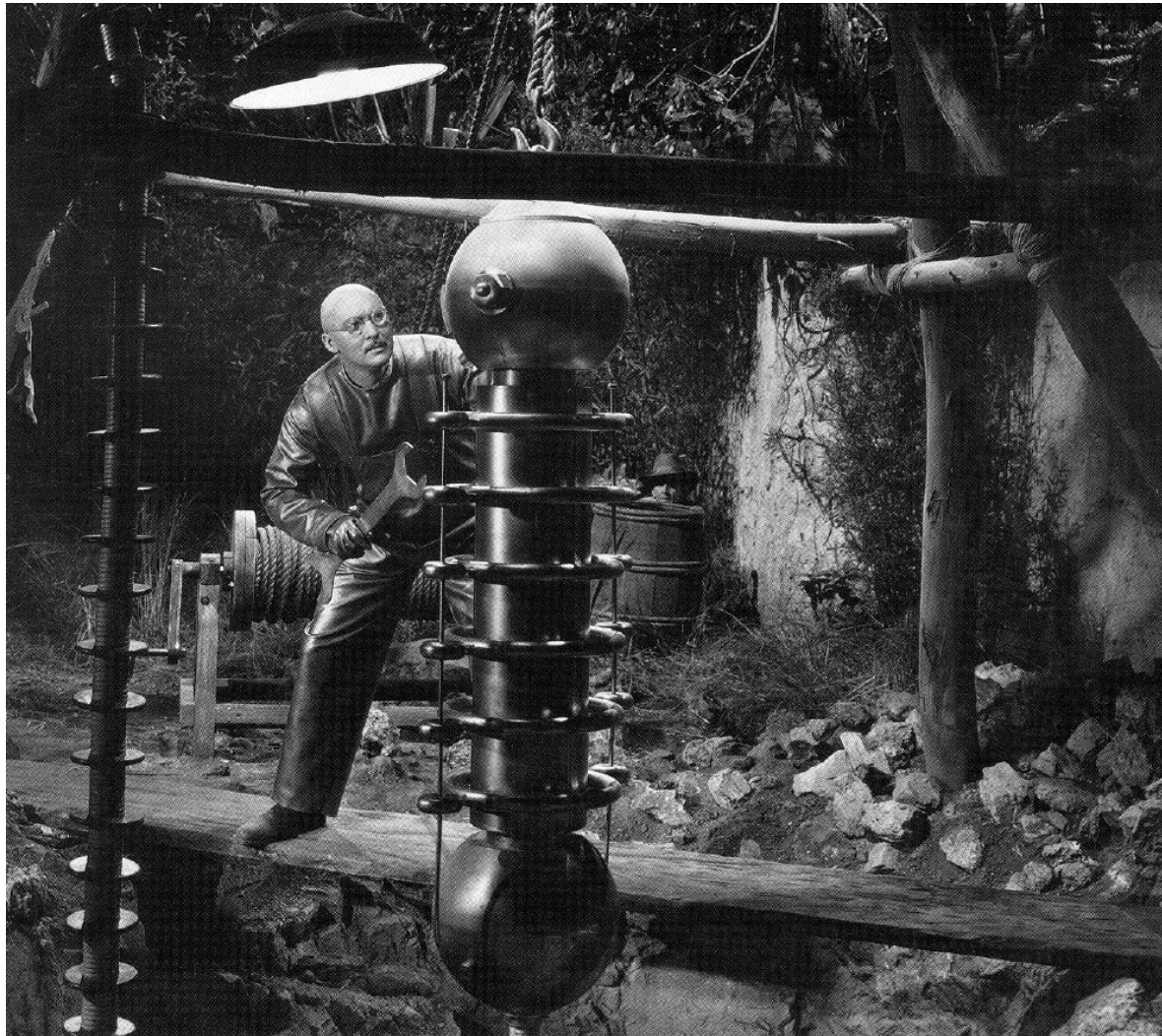
John's Champagne challenge

"So Tom, if you manage to set up a running system (technology + people) which captures over 50% of what is on the shelves in KMi and maintains this for at least 3 months

I'll buy you a bottle of

champagne (or equivalent other alcohol)"

Tom and Mark go to work on the challenge



- **Issues to address**

- Acquisition
- Maintenance
- Access/Use

- **Approaches we rejected**

- A central database
- Endnote libraries on a network share

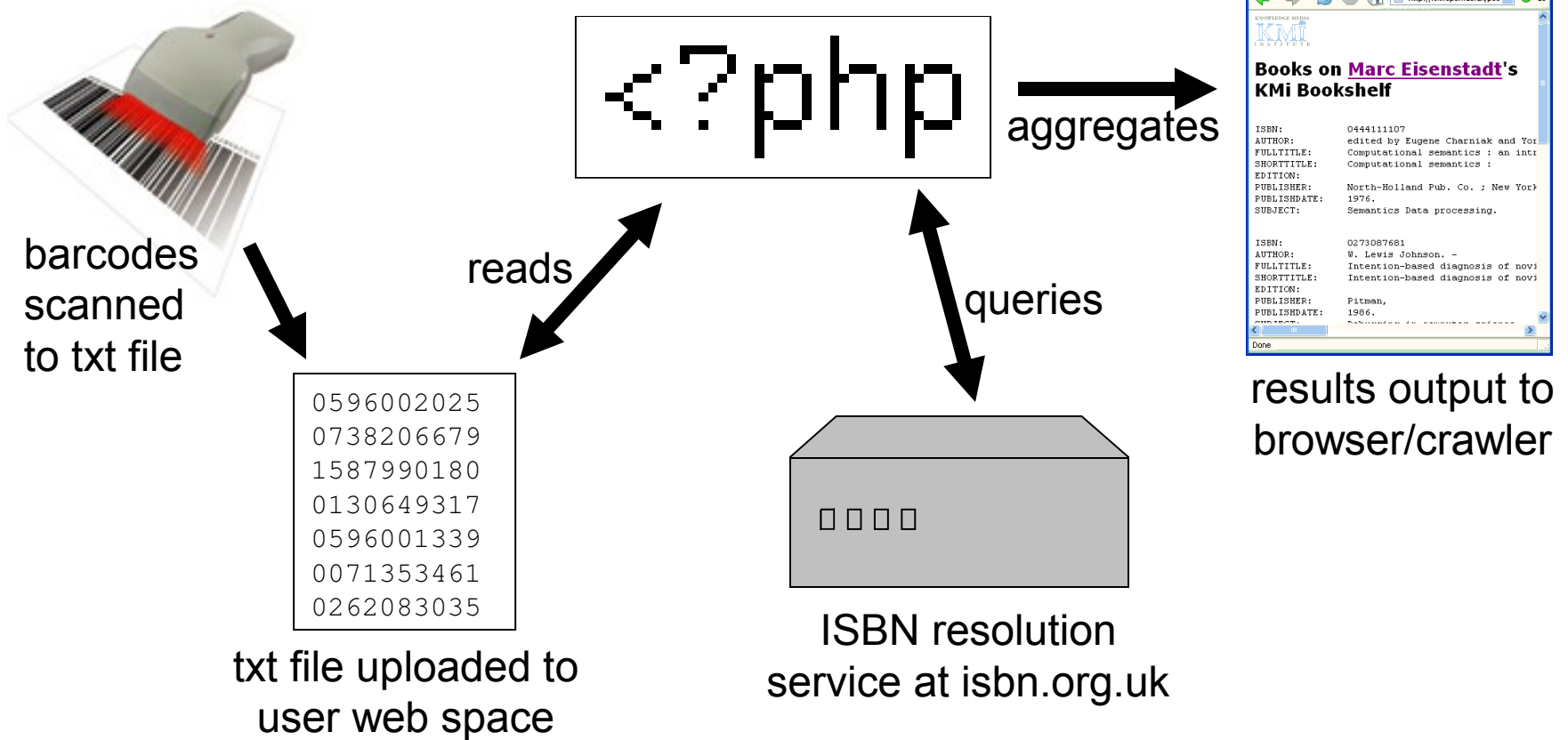


Our approach

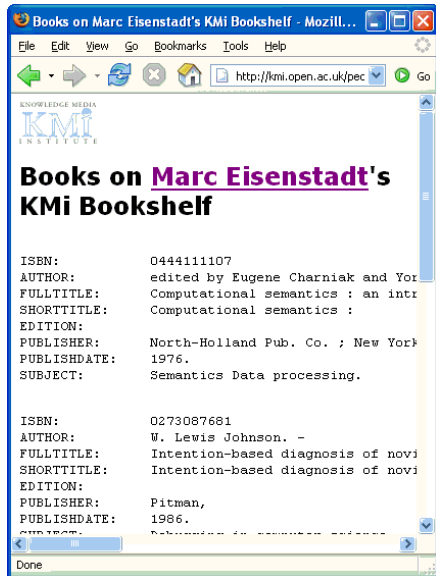
- Principles
 - minimum effort
 - maximum devolution, maximum reuse
- Solution
 - ISBNs, 3rd party web services, some PHP glue, existing search tools

Architecture: Acquisition and Maintenance

php script on kmi.open.ac.uk

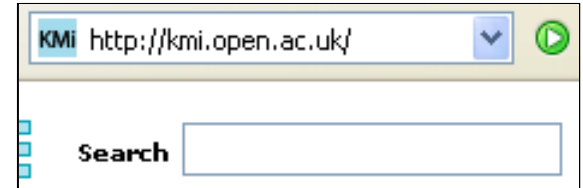
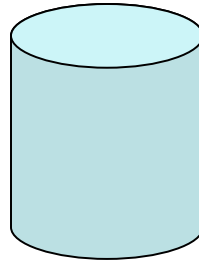


Architecture: Access/Use

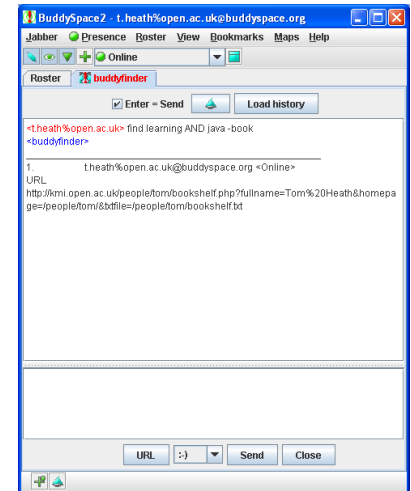


“bookshelf”
pages crawled

search index



KMi search engine



“Buddyfinder” Jabber bot

Quick demo

Some more semantics please...

- Next steps
 - Transform XML output to RDF/XML on the fly
 - Make person-to-book link machine readable
 - Some technical bits (caching, error handling..)
- Result
 - Nice big SW data set to play with!
 - Integrate with the KMi Semantic Portal
- Other applications
 - placename-to-lat/long resolution, email-to-mbox_sha1sum...

Thankyou!

- Acknowledgements
 - Marc Eisenstadt and John Domingue (for feedback)
 - Chris Denham (for Buddyfinder)
 - the OU Librarians (for the barcode scanner)
- **Questions?**
- Contact details
 - {t.heath,m.b.gaved}@open.ac.uk
 - <http://kmi.open.ac.uk>