

# eGranary Metadata Issues

Lessons “learned” from our offline library and our many users...

# eG Catalog Remains the Most Popular

- According to subscriber logs, 90% of “first hits” are to items listed in the catalog.
- Thousands of years of organizing things, we still conceptualize in “chunks.”
- Like shopping in a store, the clerks and the customers are always trying to figure each other out
- Sometimes it’s best to put a product in multiple locations



[What's New](#)



[Check Email](#)

# YAHOO!



[Personalize](#)



[Help](#)

[Yahoo! Auctions](#)  
bid & sell for free

**FAMILY GUY** Win 6 days in Hawaii!  
Go to [fox.com](#) **FOX**

[Park Your Domain Free](#)

Search

[advanced search](#)

[Yahoo! Mail](#) - Get your **free** e-mail account today!

[Shopping](#) - [Yellow Pages](#) - [People Search](#) - [Maps](#) - [Travel Agent](#) - [Classifieds](#) - [Personals](#) - [Games](#) - [Chat Email](#) - [Calendar](#) - [Pager](#) - [My Yahoo!](#) - [Today's News](#) - [Sports](#) - [Weather](#) - [TV](#) - [Stock Quotes](#) - [more...](#)

## [Arts & Humanities](#)

[Literature](#), [Photography](#)...

## [Business & Economy](#)

[Companies](#), [Finance](#), [Jobs](#)...

## [Computers & Internet](#)

[Internet](#), [WWW](#), [Software](#), [Games](#)...

## [Education](#)

[Universities](#), [K-12](#), [College Entrance](#)...

## [Entertainment](#)

[Cool Links](#), [Movies](#), [Humor](#), [Music](#)...

## [Government](#)

[Military](#), [Politics](#), [Law](#), [Taxes](#)...

## [Health](#)

[Medicine](#), [Diseases](#), [Drugs](#), [Fitness](#)...

## [News & Media](#)

[Full Coverage](#), [Newspapers](#), [TV](#)...

## [Recreation & Sports](#)

[Sports](#), [Travel](#), [Autos](#), [Outdoors](#)...

## [Reference](#)

[Libraries](#), [Dictionaries](#), [Quotations](#)...

## [Regional](#)

[Countries](#), [Regions](#), [US States](#)...

## [Science](#)

[Biology](#), [Astronomy](#), [Engineering](#)...

## [Social Science](#)

[Archaeology](#), [Economics](#), [Languages](#)...

## [Society & Culture](#)

[People](#), [Environment](#), [Religion](#)...

## In the News

- [King Hussein of Jordan dies](#)
- [Online: Lewinsky video testimony](#)
- [NASA comet mission](#)
- [NBA season opens](#)
- [Weekend's top movies](#)

[more...](#)

## Inside Yahoo!

- [Y! Personals](#) - find a Valentine
- [Shop](#) for your Valentine
- [Y! Clubs](#) - create your own

[more...](#)

# Only Just Begun to Catalog

- We already catalog
  - Web Sites
  - Web Pages
  - Sections
  - PDFs
  - Videos
  - Audio

# Drilling Down

- A Web site hosts several journals
- Each journal has a home page
- Each home page lists the issues
- Each issue has many papers
- Each paper has its own sections
- Each section has footnotes...
- Metadata... where does it stop?

# Drilling Deeper

- Users may be seeking images, charts, quotes, etc.
- How do these items get cataloged?
- No standard URL
- No creator-provided anchors
  - `<a> Bob </a>`
  - `http://cliff.com/goodstuff.htm#Bob`
- WC3 Working Group Selectors and States
  - Allows others to identify segments of a document and create links
  - Basically a browser redirect

# Atomized Cataloging

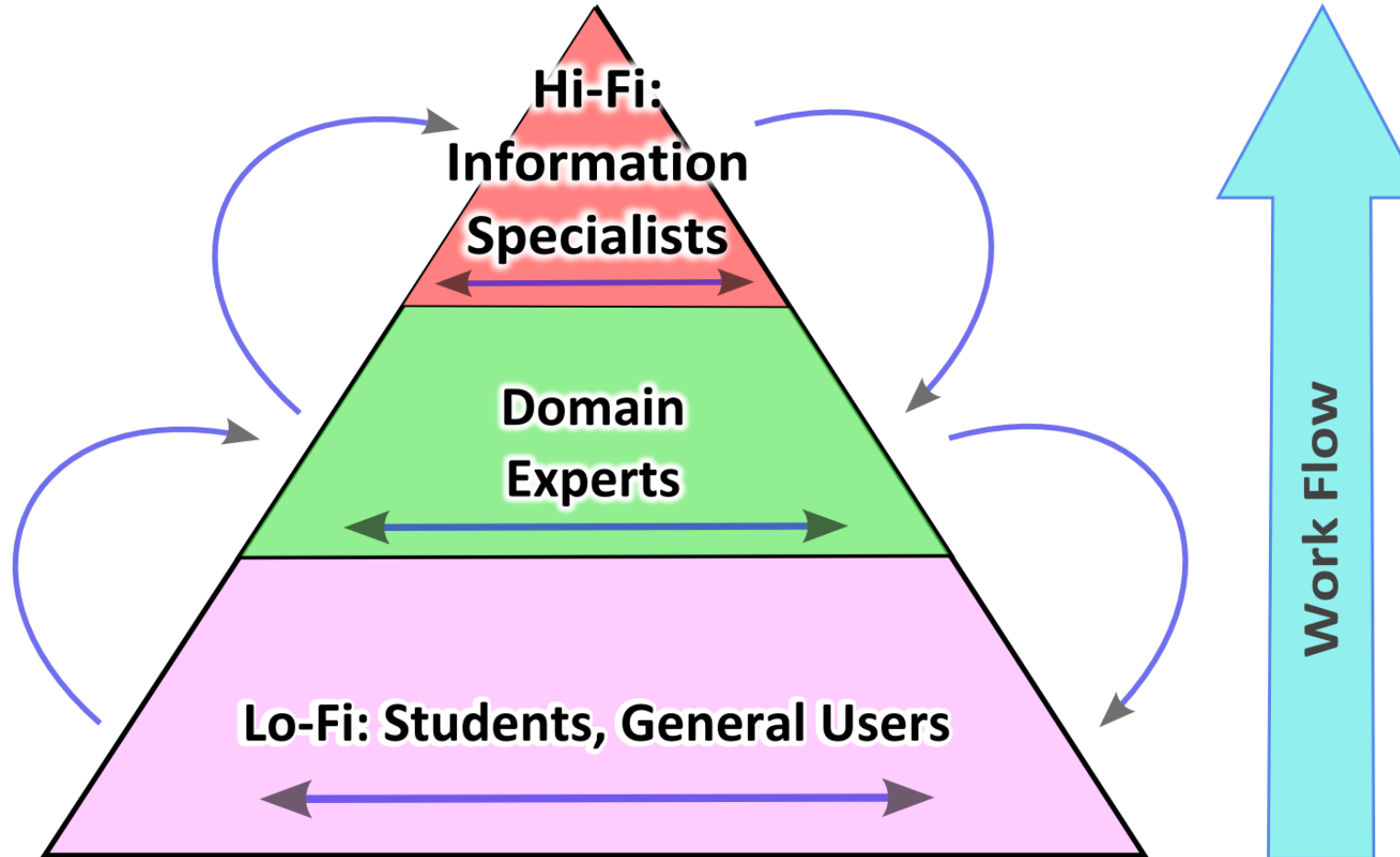
- With Selectors and States, documents can be broken down into smaller elements, with links to:
  - A particular phrase or paragraph
  - A photo or chart, with requested zoom
  - A specific time in an audio file
  - A single frame in a video
  - A extract from an animation
- `http://cliff.com/goodstuff.htm #selector (type=TextPositionSelector,start=412,end=795)`

# Pre-Screen by Machine

- Using computers to identify metadata
- Determining types (audio, video, presentations, etc.)
- Reading document headers
- File size and date
- Acquire usage data (Google, Bing, etc.)
- Analyze the document for common properties

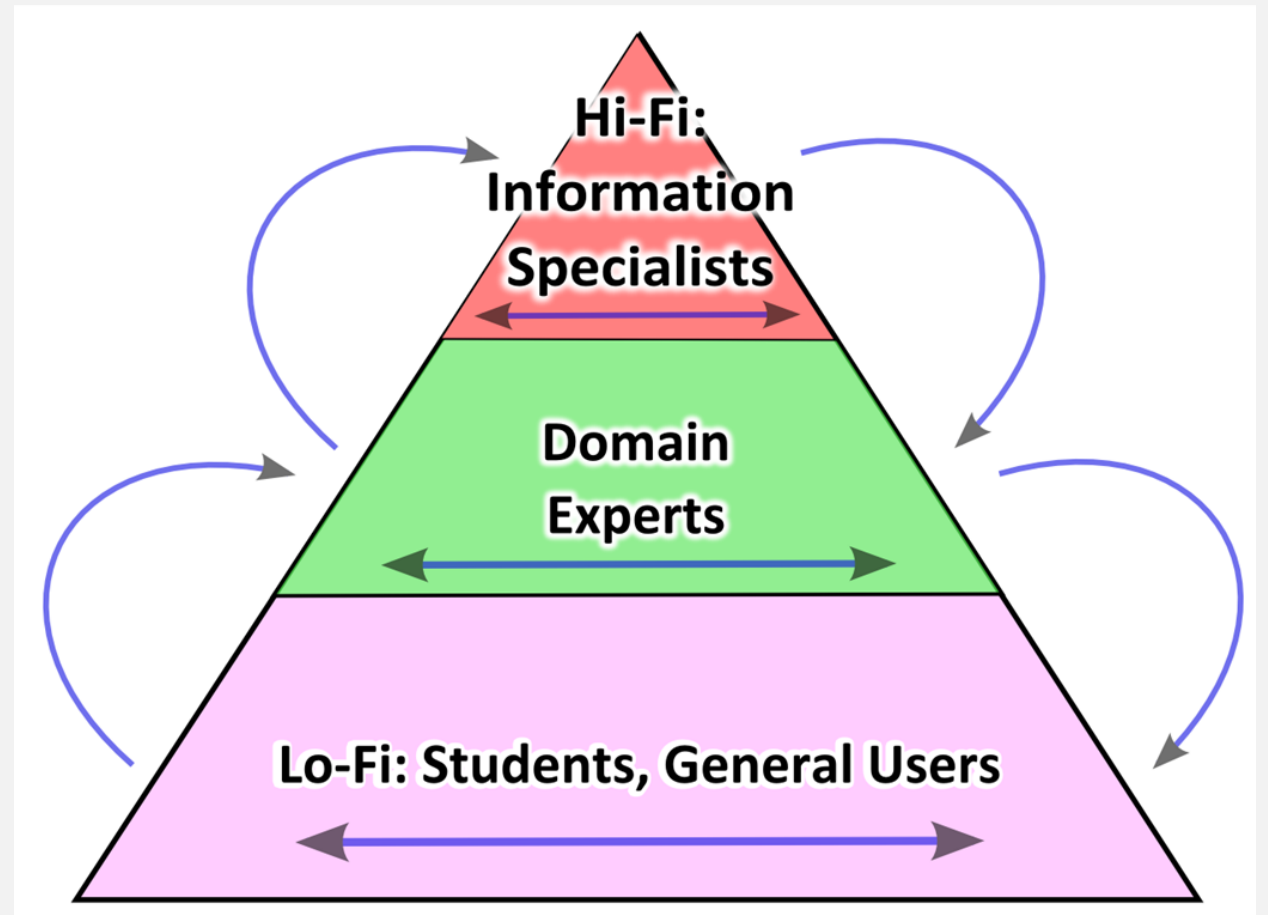


# Crowd Cataloging



# Metadating the Masses

- Level of expertise
- New records created by cataloger
- Approvals at each level
- Feedback by cataloger and reviewer
- Discern patterns of errors
- Do A/B testing
- Accommodations for other applications



# eGranary Metadata Issues

Lessons “learned” from our offline library and our many users...