Technological Assembly of Search Challenges, 1974 - 2007

Lynda Moulton, *The Gilbane Group* at

Boston KM Forum, April 5, 2007

Topic promised: *Major innovations from early electronic search to today OR progress made with an assessment of what we are still lacking*

TOPIC Expanded

- Why search?
- The millennium challenge (pre-electronic)
- The 20th century challenge
- The current state-of-the challenge

WHY SEARCH?

- *To Find*

not to be

- Annoyed
- Frustrated
- Confused
- Overloaded
- Overwhelmed
- Lost

WHY SEARCH?

More than just find

- To Discover
- To Stimulate Thinking
- To Learn
- To Get our Bearings - place in the world, our locale, our organizations, our communities, our families
- To Confirm Facts
- To Be Entertained or Navigate Life
- To Become More Than What We Are Without Searching - Increase our Knowledge

THROUGH THE MILLENNIA

- Physical navigation
- Physical Finding Devices: sextants, maps, sounding devices, telescopes, concordances, edge-notched cards, indexes
- Dependencies on:
  - massive data capture and gathering
  - human data interpretation
  - intellectual organizing devices
- Results:
  - back of book indexes
  - atlases
  - dictionaries/directories/catalogs

THROUGH THE MILLENNIA

- How did these get built?
  - Researchers captured and recorded data
  - Subject matter experts interpreted the data
  - Analysts and SMEs collaborated to transform data into assemble or aggregated information resources
  - Indexers/librarians created indexes based on authors' terminology or controlled vocabulary terminology to make information resources accessible in appropriate finding lists
  - Technique: Indexing or "Original Cataloging"

THROUGH THE MILLENNIA

- But who knew?
  - Experts
  - Researchers
  - Librarians
  - Scholars
  - "The educated"
  - People with a need to find:
    - a phone number
    - a pair of jeans
    - a machinery part
    - a car dealer

20th Century - From Print to Electronic

For What did we Build Electronic Indexes?

- To facilitate search - speed, ease of use for non-search specialists
- To provide contextual frameworks for the content
- To define the scope of a content repository

Some Search Jargon

- Structured search
- Free text search
- Structured with authority control
- Federated search
- Semantic search *
- Embedded search
- Indexing
- Metadata
- Tagging
- Spidering/Crawling
- Categorization (auto) *
- Keyword
- Phrase
- Date
- Numeric
- Controlled vocabulary
- Thesaurus
- Taxonomy
- Ontology *

Search Types Access Preparation Data Concepts Terminology Controls

Use Google: *define:xxx* to get definitions or clarity around these terms

What and why taxonomies?

- Control and Uniformity
- Quality Assurance
- Confidence Building
- Navigation
- Definition
- Cross-references
- Terminology relationships

Examples of Electronic Indexes

Government: fistgov.gov, PubMed, Herbs at a Glance , DOE,

Defense Technical Information Center (DTIC)

Libraries: Structured search – Specified Fields; Directory of http://www.publiclibraries.com/

Academic: MIT, Harvard Business School

Directories: Thomas Register, Encyclopedia of Associations

E-commerce: Computer Equipment, Automobiles, Clothing, Florists,

Publishers: Oxford Univ. Press Information Week

Specialized Indexes: Chemical Abstracts, MITRE, Raytheon, Lincoln Laboratory, Air Products, DuPont, Johnson & Johnson

21st Century - Search Challenges

Why we search has not changed BUT

Fundamentals have been transformed

Physical Location > Virtual Location

Browsing Collection Proper > Keyword Searching

Locating by Call Numbers > Locating by RFID

Print Directories > Topical Navigation

Abstracting Essential Content > Business Analytics through Algorithmic Extraction from Databases

<u>What</u> we find and <u>How</u> gives a richer <u>Result</u>

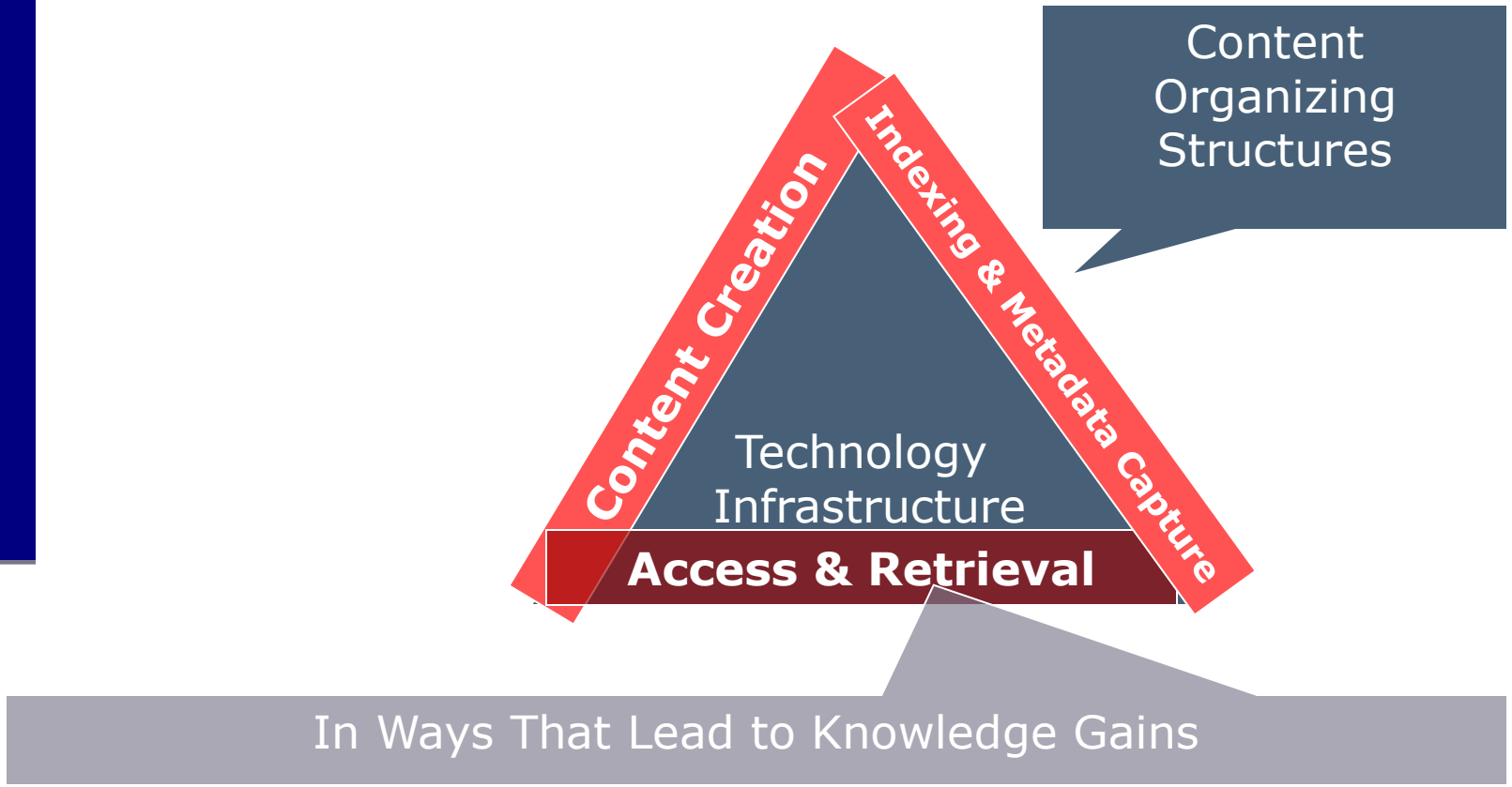
21st Century - Search Challenges

- From a Library of Congress with 6487 volumes purchased from Jefferson in 1815 ($23, 940)
- 34,000 volumes arriving/day - 7K permanent (1992 100 million items in LC)
- Supplemented by NLM, National Agriculture Library, plus all the government documents
- One of our speakers will describe a repository of 7,000,000 items for a firm of under 1000
- The challenge in the enterprise is to find what is:
  - Unique
  - Critical to R&D, business operations, worker efficiency
  - Required by regulation
  - Legally significant
  - **What you want and need!**

Changing the Search - Knowledge Connections

- Change leadership attitudes about the human/computing relationship
- Change ways in which organizations do work
- Change ways in which academia prepares the work force for communicating their learnings and methods of knowledge discovery
- Change the technology vendor client relationship to a collaborative venture
- Change how search tools are sold (licensed), bought and implemented.

Search - Content - Knowledge Connection



THE END

Contact: lmoulton@lwmtechnology.com

For additional articles related to the topics presented please see:
http://www.kmforum.org/content/kmf-04052007.pdf