IMAGINING EMERGENT METADATA, REALIZING THE EMERGENT WEB

Jason Bengtson, MLIS, AHIP
Jason Bengtson, MLIS, AHIP
Emerging Technologies/ R&D Librarian
University of New Mexico Health Sciences
Library & Informatics Center
Tel: (505) 272-0645
Email: JBengtson@salud.unm.edu
WHAT IS EMERGENCE?

- Independent Agents Interacting
- Agents following relatively simple rules
- New levels of complexity EMERGE in the system
- Organizes spontaneously, without intelligent control
EXAMPLES OF EMERGENCE

Complex behaviours by creatures of relatively low intelligence working in concert (such as social insects)

Flickr image created by Ian Armstrong
EXAMPLES OF EMERGENCE (CONTINUED)

Genetic Expression (genotype becoming phenotype)
HOW DO WE KNOW IT’S EMERGENCE?

- No organizing intelligence can be identified
- Outcomes can only be predicted through SIMULATION
- Simulations cannot be compressed
- Outcomes exert constraints over the individual agents
WHY IS THIS IMPORTANT?

- Emergent phenomena cause systems to
  - SELF-ORGANIZE
  - Organize from the bottom up
  - Generate outcomes that cannot be predicted through ordinary means
  - Operate outside of traditional reductionism
WHAT COULD THIS MEAN FOR INFORMATION SYSTEMS?

- Information systems that:
  - SELF-ORGANIZE
  - Operate outside the limitations of human design
  - “Self repair” shortcomings in original configuration
  - Respond to change DYNAMICALLY
  - Respond to change without human intervention
WHY METADATA?

- Metadata, in one form or another:
  - Lies at the heart of all modern information systems
  - Allows for interoperability
  - Allows for searchability
  - Forms the basis of MACHINE SEMANTIC systems
WHAT ABOUT METADATA NEEDS TO CHANGE

- Metadata units need to interact with each other
- Metadata units need room to be more machine-centric
- Metadata units need to be ontology-agile
HOW CAN WE MAKE THESE CHANGES?

- Use tools like bots, browsers, and evaluators to cross-pollinate Metadata units
- Recreate Metadata units as code with encapsulated Metadata tags
- Create room in Metadata schemas for tags reflecting non-human organizing principles
- Allow Metadata units to dynamically reference external ontologies
  - Actively (as software)
  - By proxy (cross-pollination)
WHY HASN’T THIS BEEN DONE ALREADY?

- Digital technology defies the physical
  - Items can exist anywhere on a network
  - Items just need to be linked digitally
  - Extensive, offsite resources can be easily referenced
  - Moving from resource to resource happens at near light-speed

- Artificial processing and evaluation are now more sophisticated
- Digital resources need less direct human oversight
- Processes can be automated easily
CROSS-POLLINATION

- Simpler
- Uses less computing power
- Could be implemented using multiple mechanisms
- More realistic in the short term
- Would probably rely heavily on user navigation
Browsing tools would act as catalysts

As users moved from one item to another:

- Browsers could make changes to level 2 and 3 tags
- Evaluate tags for retention
- Metadata agents would interact by proxy
METADATA AS SOFTWARE

- Allows for direct interactions
- User navigation less important
- More potential for novel connections/channels to Emerge
METADATA AS SOFTWARE

- Metadata would exist as information within small programs
- These programs could interact
- These programs could run on a shared network or the internet
DEFINING OUR TERMS

- Metadata files are AGENTS
- AGENTS contain data in FIELDS
- Each discrete piece of data in a FIELD is a TAG
METADATA TAGS SET IN LEVELS

- **Level 1**: regular, base level
- **Level 2**: identical to level 1 but generated through navigation
- **Level 3**: tags employing machine-semantic metadata
LEVEL 1

- Will often be assigned by a human cataloger
- Designed to be machine readable and human semantic
- Not meant to be edited by mechanical agency
LEVEL 2

- Uses the same tag content as Level 1 tags
- Meant to be assigned/edited by mechanical rather than human agency
LEVEL 3

- Tag content meant to be machine semantic only
- Might be human readable/semantic but doesn’t have to be
- Meant to be assigned/edited by mechanistic agency
AGENTS CAN BE FILE GESTALTS

- With digital technology files can be split up
- Level 1 tags could be local
- Level 2 and 3 tags could be linked from a remote server
AGENTS FOR THE WEB

- Should probably be stored in offsite indexes/networks
- Similar to keyword indexes used by search engines now
- This would keep even Level 1 tags from being modified unscrupulously
FOCUSING ON SUBJECT TAGS

- At this time I have chosen to focus on subject heading tags
- Some navigational tags are also part of Level 3
- Probably the easiest place to identify useful effects
WALKING THROUGH CROSS-POLLINATION

- Takes place in three cycles
- Should probably be catalyzed by a browsing agent
- On the web process would need to be anonymized and transparent
LEVEL 2 EXCHANGES-DYNAMISM

- Designed to augment the Level 1 tags
- May rectify shortcomings in original cataloging
- May help respond dynamically to change
- All Level 2 and 3 tags must keep a counter value
LEVEL 2 HANDSHAKE CYCLE

1. Handshake Cycle

For exchange to take place, there must be a certain threshold of matching tags and time on agent.
LEVEL 2 INTERACTION CYCLE

2. Interaction Cycle

With compatibility established tags are exchanged. Set as Level 2 with a Counter value 1
The last part of the cycle evaluates Level 2 tags by counter value for retention
EVALUATION CYCLE

- Uses two types of filter:
  - Survival Of The Fittest (SOTF)
  - Strength Of Weak Ties (SOWT)
THE WHOLE LEVEL 2 SEQUENCE

1. Handshake Cycle

**Agent One**
- Neoplasms
- Neoplasms by Histologic Type
- Leukemia
- Adolescent, Hospitalized
- Neoplasms by Site
- Abdominal Neoplasms

**Agent Two**
- Neoplasms
- Neoplasms by Histologic Type
- Lymphoma
- Adolescent, Hospitalized
- Neoplasms by Site
- Abdominal Neoplasms
LEVEL 3 EXCHANGES-ATTEMPTING REAL EMERGENT BEHAVIOR

- Using tag types distinct from the Level 1 and 2 tags
- Designed to generate/cultivate unique information channels
- Could be exploited by search and aggregation tools in a variety of ways
SOME POSSIBLE LEVEL THREE TAG TYPES

Ontology: Synonyms for tags from level 1 and 2 drawn from external ontology libraries.

Folksonomy: Synonyms for tags from level 1 and 2 drawn from external folksonomy libraries.

CrossLink: Links to other Agents that have been the subject of a successful exchange.

References: References from the Agent's Item and from any Agents that have been the subject of a successful exchange.

Search: The search terms present in the browsing mechanism at the time of a successful exchange.

Identity: Presents as string. Three separate tags generated. Top three Ontology tags + top three Folksonomy tags as determined by Counter values and random number if too many Counter values are equal.

Path: Presents as string. Three separate tags generated. Top three Reference tags + top three Crosslink tags as determined by Counter values and random number if too many Counter values are equal.

Route: Presents as string. Three separate tags generated. Top three Search tags + top three Identity tags as determined by Counter values and random number if too many Counter values are equal.
For exchange to take place, there must be a certain threshold of matching tags and time on agent.
With compatibility established tags are exchanged. Set as Level 3 with a Counter value 1
LEVEL 3 EVALUATION CYCLE

The last part of the cycle evaluates Level 3 tags by counter value for retention.
1. Handshake Cycle
NEXT STEPS

► Determine if the idea has utility through simulation
► Use an emergence simulator like NetLogo (if possible)
► If these tools are inadequate, create a proprietary simulation methodology
WHAT ARE WE LOOKING FOR

- Useful adaptations to navigation or other environmental changes
- The appearance of novel channels or networks in the form of linking and navigational pathways
QUESTIONS?