Domain-Specific Search and the Encyclopedic Internet Vision

Network and Complex Systems Group
Indiana University

Anthony F. Beavers, Ph.D.
Philosophy / Cognitive Science
The University of Evansville

Funding by the Digital Humanities Initiative,
National Endowment for the Humanities
What is Noesis?

- An “intentional” act of consciousness that is correlated with an object
  - The act is termed *noesis* (adj. noetic)
  - The object is termed *noema* (pl. noemata, adj. noematic)
  - The correlation is termed *noetic-noematic* to indicate the unitary structure of act and object
Intentionality & Hypertextuality

- Hyperlinks are also intentional in this sense
  - A link in order to be a link must point from a source to a target
  - Thus the hyperlink forms an intentional or bi-polar relation with one pole at its source and the other pole at its target
Two Points of View on Noesis

- From the noetic point of view:
  - A limited area search engine dedicated to the discipline of philosophy

- From the noematic point of view:
  - A library of reliable philosophical research designed according to the affordances of the Internet
Some Noetic Affordances

- Gateway or portal to off-site information
- Server-side centralization for information analysis, organization, visualization, and quality control
- Ability to impact a range of noematic affordances
Some Noematic Affordances

- Organic domain representation (regarding both topical scope and the number of documents)
- Hypertextual Relatability
- Decentralization (regarding both data management and data storage)
Definitions

- Search Engine: The mechanism that allows users to find documents

- Search Request: The query string that prompts the search engine to respond

- Search Return Set: The set of documents returned for a particular search request

- Search Space: The set of documents that can potentially be returned by an unspecified search request
Parameters

- May vary according to discipline and task

- Search and Browse Modes
  - Random Discovery, Bookshelf Browsing, Exploratory Discovery, Programmatic Research Discovery, Particular Resource Location

- Quality Control
  - Explicit Peer Review – Our Traditional Method
  - Emergent Quality Control – Based on Page Rank, Citation Rank, Search Space Design, etc.
Some Interesting Problems

- Emergent organization and quality control
- Organic search space design for context-sensitive searching; automatic subset creation
  - Interactive visual interface construction to navigate large information stores
  - Ethical prioritization of information to respect the topical needs of the user and the politics of the profession
Some Solutions

- Limited Area Searching
- Dynamic Classification with an Organic Formal Taxonomy
- Document Comparison and Classification with a Recursive Artificial Network
Limited Area Searching

- Targets quality control issue and other issues connected to relevance
- Two-pronged approach divides task into 1) where to search and 2) what to search for
- Three models
  - Noesis 2.0 and earlier – 1998
  - Noesis 4.0 – 2006
The Argos/Hippias Model

- First ‘peer-reviewed’ search engine online
- Used a set of ‘associate sites’ to determine search space
- Searched the associate sites and everything to which they linked

(more)
The Argos/Hippias Model

- Handed editorial control of content over to the editors of the associate sites
- Provided backbone for EAWC context-sensitive searching
The Noesis 2.0 Model

- Based on Plato SE prototype
- Database driven and hand-catalogued
- Editorial control managed by a team of content editors who manually checked author credentials
- Topic tree formation developed by professional editors (dismal failure with important lessons)
The Noesis 4.0 Model

- Search space based on mapping the profession of philosophy online
- Regions include associations, departments, faculty webspace, online journals and reference works

(more)
The Noesis 4.0 Model

✓ “By indexing regions, in effect, directories and subdirectories, rather than their contents, Noesis passes editorial control of its search space over to the individuals who, in managing their own web resources, add to, edit, and delete from the content searchable by Noesis.... The result is that the shape and texture of Noesis's search space is determined organically by credentialed scholars whose actions directly determine content.”
Dynamic Classification

- The Indiana Philosophy Ontology Project (InPhO)
- Targets the organizational issue by providing an emergent topic tree based on an analysis of the Stanford Encyclopedia of Philosophy (SEP)
- Artificial intelligence-based with feedback from human users
Organic Search Space Design / Automatic Subset Creation

- Based on a comparison of semantic features using a recursive artificial network
- Will provide targeted context-sensitive hypertextuality
- Will enable exploratory discovery and “bookshelf browsing” by potentially linking every document in Noesis’ search space to every other relevant document
Affordances of Inphormed Noesis

- Organic and emergent representation of the profession of philosophy
  - Unbiased judgment of professional import
  - Ability to track changes in the profession

- Topical catalog that lets documents “speak for themselves”
  - Emergent paradigm based in a bottom-up cataloging strategy
  - Overcomes the discovery problem even in one’s own area of research; fosters *intradisciplinarity*
Some Limitations

- Bypasses philosophical resources that are coming from other professions
- Insofar as InPhO is based in the SEP, editorial bias may influence emergent representation
- Model does not currently make allowances for long-term resource archiving
Partial Bibliography


* * *

Any views, findings, conclusions, or recommendations expressed in this presentation do not necessarily reflect those of the National Endowment for the Humanities.