What’s New in Semantic Enrichment?

Marjorie M.K. Hlava
President
Access Innovations, Inc.
Mhlava@accessinn.com
LOTS! A few examples

- Smart Submit
- Recommender
- Linking data
- Search
- Auto Legacy tagging
- Automating Medical Coding
- Author and Affiliation Disambiguation
- Visualization based on Semantics
- And MUCH More
- Author or Editor does the tagging
- Tags automatically suggested by
  - Data Harmony MAI
- From Controlled Vocabulary
  - Thesaurus
  - Taxonomy
  - Ontology
  - Knowledge Organization System
  - Using Data Harmony Thesaurus Master
Smart Submit - Let Editors Choose
The author fills in the data to the document template, attaching images and graphs as necessary.

An API calls Data Harmony and generates a list of indexing terms based on the content.
A popup list of contributor role options appears for the author to choose from:

- Study conception
- Methodology
- Formal analysis
- Computation
- Investigation
- Resources
- Data Curation
- Publication
- Supervision
- Project administration
- Funding acquisition

Application of statistical, mathematical or other formal techniques to analyze study data.
Recommender – “more like these” based on the semantic fingerprint
Alcohol, Folate, Methionine, and Risk of Incident Breast Cancer in the American Cancer Society Cancer Prevention Study II Nutrition Cohort

Heather Spencer Feigelson¹, Carolyn R. Jonas, Andreas S. Robertson, Marjorie L. McCullough, Michael J. Thun and Eugenia E. Calle

Department of Epidemiology and Surveillance Research, American Cancer Society, National Home Office, Atlanta, Georgia 30329-4251

Recent studies suggest that the increased risk of breast cancer associated with alcohol consumption may be reduced by adequate folate intake. We examined this question among...
Search Integration

- Leverage the controlled vocabulary
- Basis for faceted search
- Let customers contribute
Predation by Bears Drives Senescence in Natural Populations of Salmon

Stephanie M. Carlson, Ray Hiborn, Andrew P. Hendry, Thomas P. Quinn

Published: Dec 12, 2007 • DOI: 10.1371/journal.pone.0001286

Abstract

Classic evolutionary theory predicts that populations experiencing higher rates of environmentally caused (“extrinsic”) mortality should senesce more rapidly, but this theory usually neglects plausible relationships between an individual’s senescent condition and its susceptibility to extrinsic mortality. We tested for the evolutionary importance of this condition dependence by comparing senescence rates among natural populations of sockeye salmon (Oncorhynchus nerka) subject to varying degrees of predation by brown bears (Ursus arctos). We related senescence rates in six populations to (1) the overall rate of extrinsic mortality, and (2) the degree of condition dependence in this mortality. Senescence rates were determined by modeling the mortality of individually-tagged breeding salmon at each site. The overall rate of extrinsic mortality was estimated as the long-term average of the annual percentage of salmon killed by bears. The degree of condition dependence was estimated as the extent to which bears killed salmon that exhibited varying degrees of senescence. We found that the degree of condition dependence in extrinsic mortality was very important in driving senescence: populations where bears selectively killed fish showing advanced senescence were those that senesced least rapidly. The overall rate of extrinsic mortality also contributed to among-population variation in senescence—but to a lesser extent. Condition-dependent susceptibility to extrinsic mortality should be incorporated more often into theoretical models and should be explicitly tested in natural populations.

Figures
Predation by Bears Drives Senescence in Natural Populations of Salmon

Stephanie M. Carlson, Ray Hilborn, Andrew P. Hendry, Thomas P. Quinn

Published: Dec 12, 2007  DOI: 10.1371/journal.pone.0001286

Abstract
Auto-completion using the taxonomy

Guide the user

Navigate the full taxonomy “tree”

BROWSE
Automatic Tagging

- Legacy collections
- Fast processing
- 92% accuracy
- Full text
- Inline
- Metadata enrichment
oil prices rose on Thursday as hopes that the White House would move quickly on an economic stimulus package outweighed flagging demand and rising inventories in the world’s top consumers. U.S. crude settled 12 cents higher at $43.67 a barrel, after falling as low as $40.41 earlier. London Brent settled at $45.39 a barrel, up 37 cents. Earlier in the day, crude prices had dropped after a U.S. government report showed that crude oil, gasoline and distillate fuels rose last week as demand for fuels weakened again. The U.S. stock market pared early losses after a spokesman said President Obama’s administration is committed to moving as quickly as possible on an economic stimulus plan. "crude is still resilient, despite the big build you have seen in the EIA data. I have a feeling that there will soon be a rebound in the stock market and that will spill over to the energy markets," said Mark Waggoner, president of Excel Futures in Huntington Beach, California. "In the last week, I’ve seen investor confidence improving, with the new Obama administration now installed ... I feel that if this confidence continues, it will spill into more (consumer) buying and that will improve demand for gasoline and other energy products," he added. oil fell in early trade after U.S. energy Information Administration data showed that crude oil inventories jumped 6.1 million barrels last week, well above expectations for a build of 1.4 million barrels.
MAI Batch Processing

Files Processed Report

Customize by file batch

Suggested indexing in Relevance order
Automating Medical Coding

- Semantic techniques
- ICD-9 and ICD-10
  - Diagnosis Codes
- American Medical Association CPT
  - Procedure Codes
- American Medical Association HCPCS
  - Medical Supplies
Semantic Enrichment and Concept Extraction Provides Disruptive technology to the Medical Coding Industry
Lists contextually appropriate Codes
Highlights text related to code
Author Disambiguation

Semantic Fingerprint,
Affiliation Disambiguation

Remaining authors: 1

<table>
<thead>
<tr>
<th>Author #</th>
<th>Surname</th>
<th>Given-name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1044556</td>
<td>O'Brien</td>
<td>Paul G.</td>
</tr>
</tbody>
</table>

Search authors
- Given-name
- Surname
- AUTHOR#
- Keywords
- Institution names:

Add as a New Author  Update  Delete
All data up-posted to the top level
This is a radial graph of “plosthes”. The number of records for which each index term occurs is reflected by circle sizes.
Database Plus Search Workflow

Raw Full text data feeds
Printed source materials
Data Crawls on data sources

Source data

Clean and enhance data

XIS Creation
XIS repository
Taxonomy terms
MAI Concept Extractor
MAI Rule Base
Taxonomy Thesaurus Master
Add metadata

Load to Search
Search for ecommerce
Search Harmony Display Search

Search data

SQL for ecommerce
Authors at a Place
Member Profile Tagging

User pastes or uploads CV

Button to auto-extract taxonomy attributes

A Critical Problem for Scholarly and Professional Publishers is identifying authors and managing author relationships. Publishing increasingly becomes an automated process of identifying and contact potential reviewers. The many overlapping networks of author names often appear in multiple variations, and even when they have the same or similar names, they are often used by researchers, and greater exposure for their work is desirable. Various efforts are underway to standardize author names globally, including the ORCID (Open Researcher and Contributor Identification) and emerging standards like ORCID (Open Researcher and Contributor Identification).

Using xPeerent’s IT Marketplace Tagging, the CV is described and used by the system to meet customer requirements.

SOFTWARE: Auto-suggest

- Business Applications
- Collaborative Apps
- Business Process Management Apps
- Conferencing
- Videoconferencing Apps
- Web-conferencing Apps
- Integrated Collaborative Environments
- Knowledge Management Applications
- Learning Mgmt Apps
- Messaging
- Digital Signage Solutions
- Email and Group Calendar
- Enterprise Instant Messaging
- Unified Messaging Apps
- Team Collaboration Apps
- Document Management Apps
- Project Mgmt Apps
- Social Networking Apps
- Wiki Apps
- Scientific Apps
- Data Visualization Applications
- Informatics Applications
- Lab Information Mgmt Systems
- Mathematical Analysis Applications
- Molecular or Multi-scale Modeling and Simulation
- Statistical Analysis Applications
- Engineering Apps
- Product Quality Management
- Customer Relationship Mgmt Apps
- Marketing Information Systems
- Event Management
- Info Access, Analysis and Delivery
- Text Analysis and Discovery Tools
- Content Apps
- Authoring and Publishing Software
- Data Mgmt and App Development
- Quality and Life-cycle Tools
- Automated Software Quality (ASQ)
- Info and Data Mgmt Software
- Data Integration and Access Tools
- Data Warehouses and Databases
- Database Development and Mgmt Tools
- End-user Demos
- Object-oriented Demos
- Post-relational Demos
- Recommender
- The OUP mouse over
- PLOS implementation of taxonomy in search
- Author disambiguation using taxo for semantic fingerprint /author profiling
- Affiliation disambiguation - affiliation fingerprint.
- other author disambig shots
- Affiliation disambiguation - affiliation fingerprint.
- other author disambig shots
- Inline tagging - several options
- MAI Batch GUI
- XIS using Lucene search
- Ontology Master mock up
- Media Sleuth with and without search within a search options
- Media Sleuth with two sets of parens - one for term and one from branch hits
- Access Integrity Medical Claims Compliance - how it works slide?
- Author Submit / Smart Submit
Using taxonomies in applications

- Improve search
- Subject browsing
- Mobile intelligence
- Targeted resources based on subject or user role
- Link to society resources
- Author submission module
- Author authority database
- Expert reviewer identification
- Member profiles
- Data visualization
- More like this

- In “indexing” or categorizing, as subject metadata
- In content management systems
- In SharePoint
- In mash-ups
- In social networking sites
- In author tagging
- In filtering data – e.g., spam filters and RSS feeds
- In web crawlers

Social media - community
More Innovations!

- Link topic to article to author to event
- Make visual links within domain
- Enable authors to submit and *categorize* conference submissions
- Create author authority database linking to co-authors, topics, locations, etc.
- Create expert reviewer database
- Create member profiles with alternate names, publications, tagged by topic
- Visualize data and domain distribution
- Display interest connections in social network
- Deliver accurate targeted information through mobile applications
Access Innovations - Semantic Enrichment Services Provider

Creator of Data Harmony Tools

More than 600 taxonomies created

More than 2000 Engagements

Financed by Sweat, Persistence, and Good Cash Flow management

Accurate, On time, Under budget!
Marjorie Hlava
- mhlava@accessinn.com
- +1-505-998-0800

Jay Ven Eman
- J_ven_eman@accessinn.com