Improving Play Concept Development Using Semantic Technologies*

Anthony Gary[1], Arun Majumdar[2], William Full[3], and John Sowa[2]

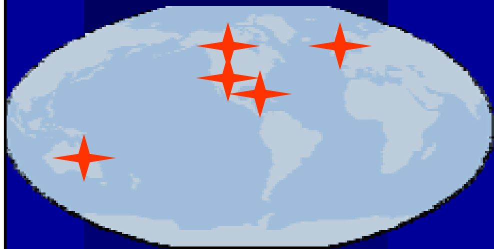
Search and Discovery Article #40377 (2009)
Posted January 20, 2009

*Adapted from oral presentation at AAPG Annual Convention, San Antonio, TX, April 20-23, 2008

[1]Energy & Geoscience Institute, University of Utah, Salt Lake City, UT. (tgary@egi.utah.edu)
[2]VivoMind Intelligence, Inc., Rockville, MD.
[3]Tramontane, Inc., Salt Lake City, UT.

Abstract

The critical ingredients for developing a successful play concept are data, experienced geoscientist(s) and time. In today's E&P environment the latter two ingredients are in short supply. It generally requires ten or more years of post-graduate experience to become a seasoned explorationist. Our objective is to develop a system that will enable geoscientists or teams at various skill levels to close the gap in time and capabilities for play concept development. The system functions can locate and retrieve relevant data, and intelligently process these to form quantified hypotheses regarding potential plays. We overcome two significant obstacles: firstly, since much of the data relevant to a play is fragmented and isolated among different databases, a method is developed to associate them semantically; and secondly, processes for analyzing structured and unstructured data are coupled to avoid missing important links between concepts. Structured data in databases can be processed by computers while unstructured data, like published reports and articles, have required a human. The geoscientist must read reports, fuse results from the structured data, and then incorporate years of experience to form a play concept hypothesis. The system we are developing streamlines this process, augments efficiency for the geoscientist and produces a quantitative evaluation of the play concepts. The outputs use evidential measures from background knowledge for reasoning, analog identification and association to produce quantitative results. The technology our system utilizes has been developed in other domains with similar processes and data as hydrocarbon exploration. Initial results applied to the E&P domain illustrate that semantically meaningful context can link concepts across multiple data sources, both structured and unstructured, to form a new hypothesis.



EGI

Energy & Geoscience Institute
at the University of Utah

Improving Play Concept Development Using Semantic Technologies

Anthony Gary, Arun Majumdar,
William Full & John Sowa

Presentation Outline

1. Statement of problem.
2. Overview of the semantic technology – analogy engine.
3. Proposed process for implementing the technology.
4. Simple tests of geologic application.
5. Overview of proof-of-concept study.

E&P: High Risk Business of Conjectures & Hypotheses

EGI

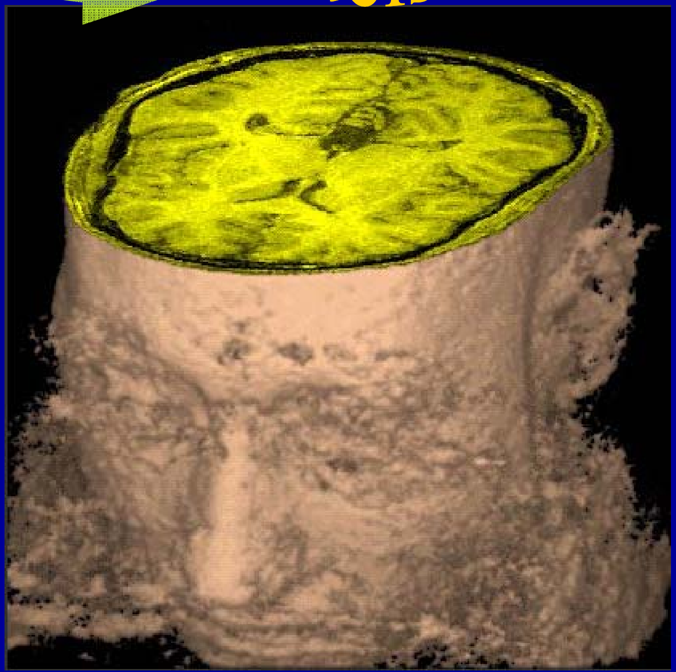
- Determine new prospects
- Discover new play concepts
- Develop infill drilling strategy
- Determine appropriate system analogs
- Determine the ‘best’ bid amount

The E&P Knowledge Creation Dilemma

New concepts & hypotheses result from the confluence of past & novel experiences.

Experienced E&P geoscientists are a limited resource.

The Industry needs strategies to increase the productivity of E&P knowledge creation.



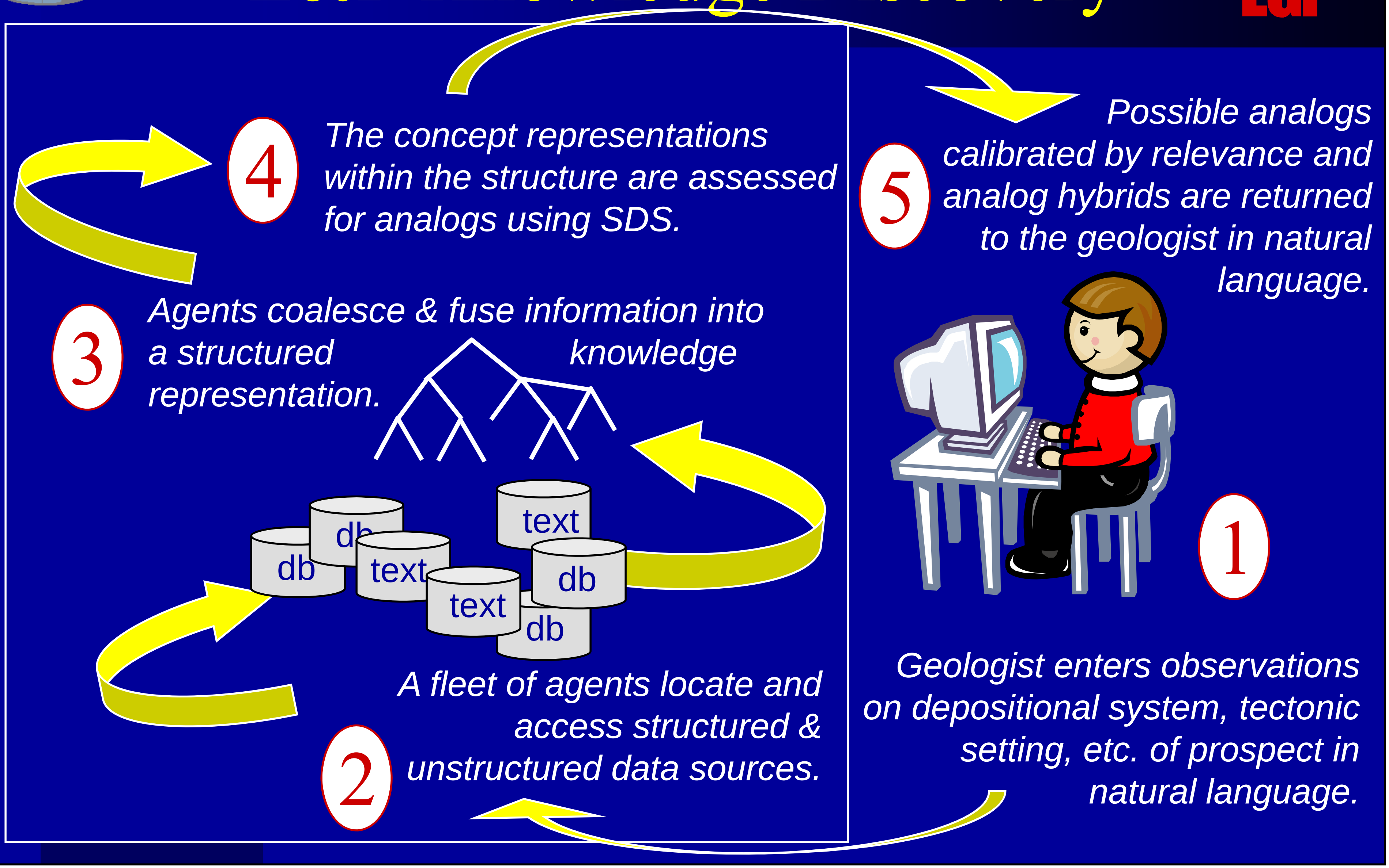
Michal Abbott, www.mabot.com

Barriers to Knowledge Creation EGI

- Data are highly fragmented, residing in different databases (*e.g.*, vendor, discipline, M&A, etc.).
- Related data are often isolated by discipline (*e.g.*, different vocabularies, scales, etc.).
- E&P data are heterogeneous – structured and unstructured (*e.g.*, attributes in databases, text in reports & articles).
- Reports and articles cannot be sufficiently analyzed by algorithms, and therefore require human processing.
- SQL limits data accessibility & types of queries.

Proposed Process for E&P Knowledge Discovery

EGI



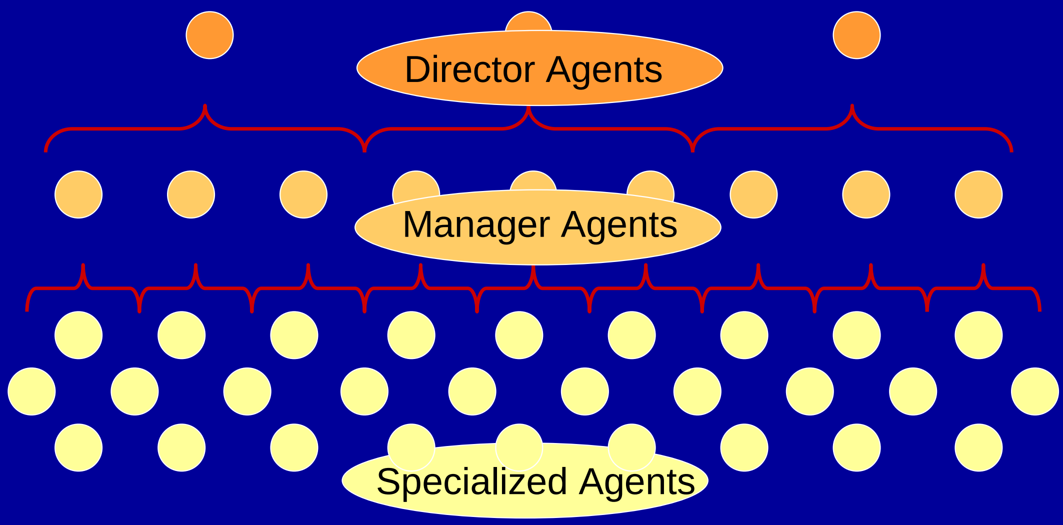
Multi-Agent System

EGI

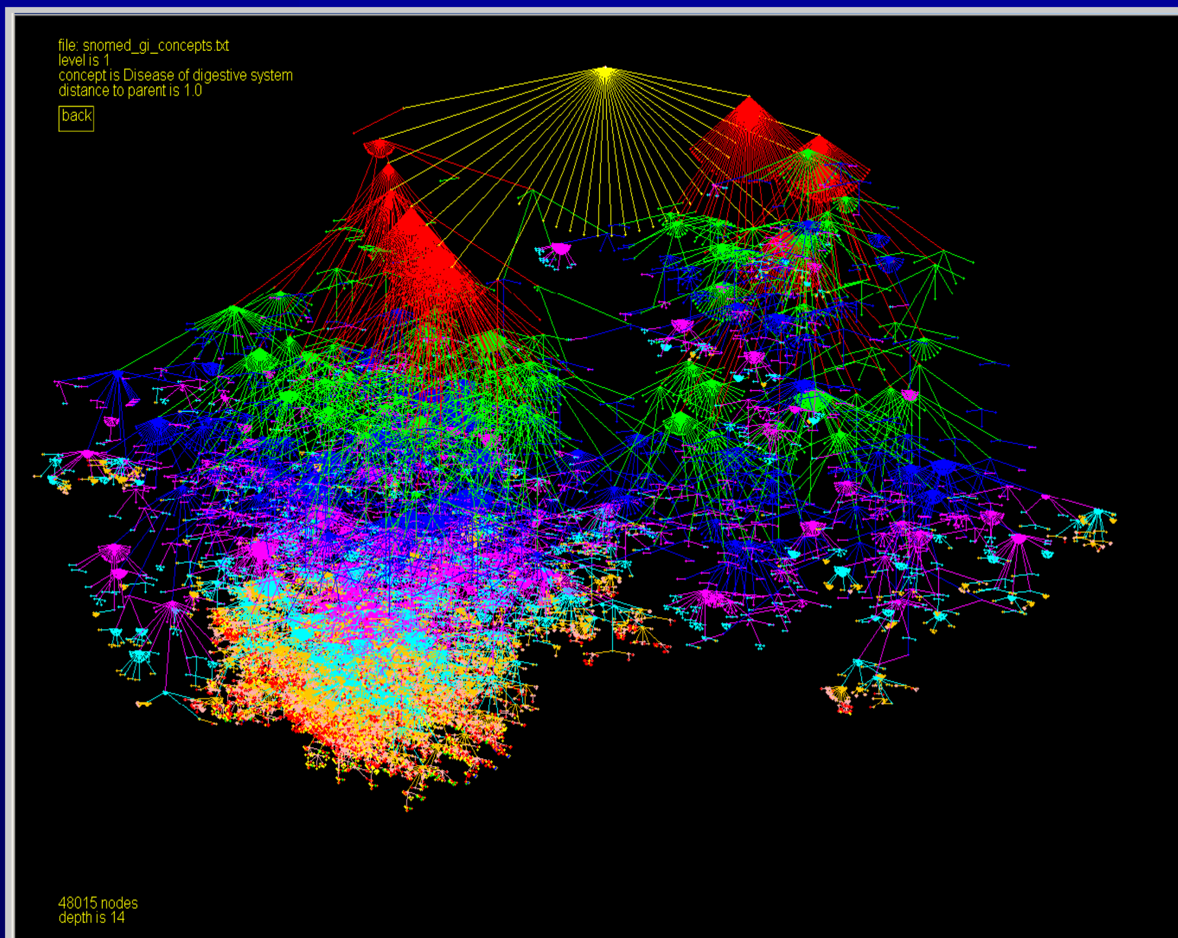
Agents are (semi)autonomous, assist in locating and processing information for the requester.

Agents perform specific tasks (e.g., Oracle agents, Wiki agents, petrophysical data agents) generally simple ones, such as parse text within a document or resolve synonymies.

A collection of agents with different specialties interact to perform tasks for the requester (*i.e.*, society of agents).



Example of a Conceptual Graph
(Step 3)



Conceptual graph for gastrointestinal disease hierarchy .

Has 48,066 concepts and a maximum hierarchical depth of 14 levels.

Each color represents a specific depth in the hierarchy.

*R&D with The Mayo Clinic, (To Appear) - **Peter L. Elkin, MD1, Arun Majumdar, B.Sc,, Steven H. Brown, MD, MS 2, 3 S. Trent Rosenbloom, MD MPH2, John Sowa4, PhD, Dietlind Wahner-Roedler, MD1, Brent A. Bauer, MD1, Zami Temesgen, MD1, Luka Bajzer1 [Copyrighted and proprietary to the authors]**

Processing Semantic Information

(Step 4)

EGI

Similarities between conjectures, hypotheses and analogs within CGs are determined by Semantic Distance Signatures (SDS).

A SDS is a value that represents both the topological structure and the ontological features of a conceptual graph (as a semantic value in a semantic distance field).

Simple metric measures, such as Jaccard or cosine-theta, can be used to determine concept similarities from a field of SDS. Can be biased to select either reasoning context or analogs.

Overview of Semantic Technologies

EGI

Diversity of data filters at the low level: some filters just sense numbers and relations, others text, others pictures…

Diversity of methods of reasoning – induction, deduction, abduction and combinations thereof.

Diversity of strategies of hypothesis formation – from evidence, from observation, from facts, from invention.

Varied methods of learning – by example, by rote (*i.e.*, schema or script), or by human interaction (via controlled language).

Simulative reasoning (aka., counterfactual or scenario based "What-If" reasoning).

A Simple Test: Forming a Concept Across Two Articles

EGI

"Natural fractures in the Barnett Shale and their importance for hydraulic fracture treatments" (Gale et al., 2007)

"Oil and gas geochemistry and petroleum systems of the Fort Worth Basin" (Hill, et al., 2007):

CGs were produced using common ontology & formal name word list.

A simple conjecture was input ("high oil and gas in reservoir").

Only a few general agents were used.

Did not utilize a domain-specific ontology, rules or heuristics.

The Resulting Hypothesis

EGI

"...the Barnett Shale consists of mudstones interspersed with carbonate layers, and the Austin Chalk comprises chalk-marl couplets..." (Gale et al., 2007)

"Fine grained, layered chalk or mudstone with high depth and heat flow, and with low fracture, implies high oil and gas reservoirs."

"...gas generation and consequent gas production are largely controlled by high heat flow..." (Hill, et al., 2007)

Multiple Evidence for Inference of Low Fracture

EGI

"We see no evidence of widespread open natural microfractures..."

"Three natural fractures were observed in the mudstone, each being less than 0.05 mm..."

"Fine grained, layered chalk or mudstone with high depth heat flow, and with low fracture, implies high oil and gas reservoirs."

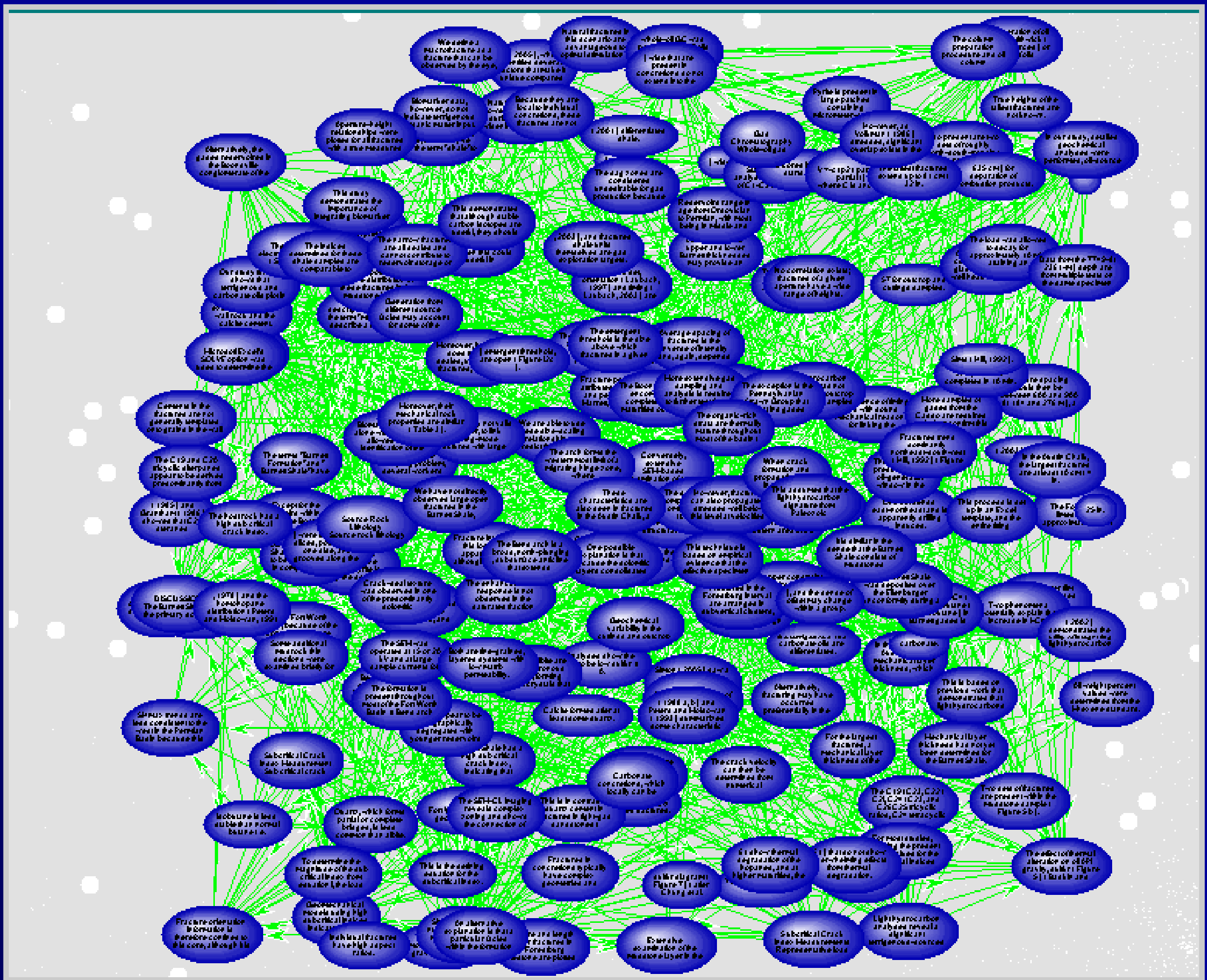
"Fracture intensity in this location is apparent low..."

"Most natural fractures in this core are found in the Forestburg Interval..."

(Gale et al., 2007)

Graphical Representation of Semantic Relationships

EGI



System has a diverse set of tools for manipulating, analyzing and visualizing semantic data.

Advantages of Semantic Technologies

EGI

- Locates relevant data across different databases.
- Recognizes comparable data across different sources (*e.g*., corporate, legacy & vendor).
- Processes unstructured data (*e.g*., articles, reports).
- Integrates structured (*e.g*., databases) and unstructured (*e.g*., articles, reports) data.
- Links data from different disciplines that share a concept.
- Quantifies analog similarities.
- Uses natural language, not query languages (*e.g*., SQL).

Semantic Technology Research Project Underway

EGI

Objective:

Develop a natural language system that can locate analog data from different sources, both structured and unstructured, and integrate them into a conceptually-relevant structure that can be analyzed for analogs and possible 'hybrid' analogs.

Proof-of-Concept:

Deepwater turbidites; sources: approx. 50 oil & gas publications, and Cossey reservoir & field databases.

Proof-of-Concept: Field and Outcrop Analogs

EGI

Outcrop and field analogs are critical for "ahead of the bit" predictions from prospect evaluation to infill drilling.

The problem is finding the 'right' analog. Comprehensive evaluation of all possible analogs at multiple-scales (*e.g.*, tectonic setting, depositional system, facies architecture, etc.) is not economically possible.

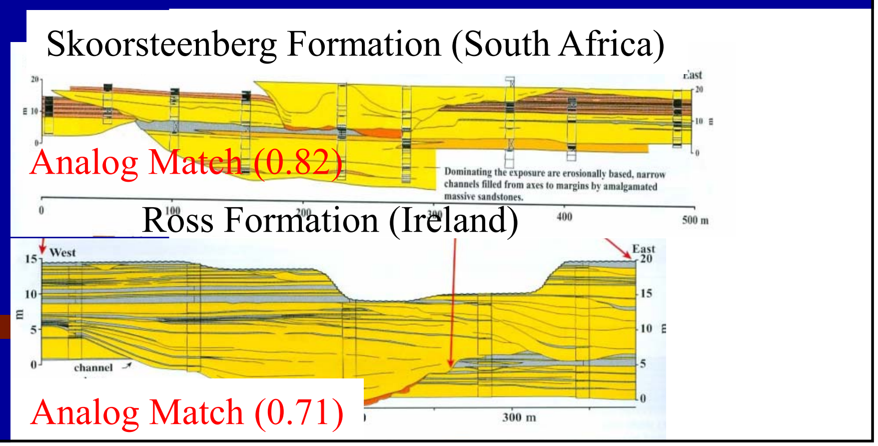
Bias to the familiar analogs or those getting the most press.

The Bottom Line

1. Reduce the likelihood of overlooking the “best” analog.
2. Reduce the likelihood of selecting the “wrong” analog.

How?

- Capability of software to capture & semantically process text.
- Integrating both quantitative & qualitative data/information.
- Quantify the similarity between prospect and analog(s).
- Possible to produce “hybrid” analogs.
- More robust and lower cost than re-engineering & migrating data.
- Specialized agents post process analog parameters (*e.g.*, connectivity) for easy evaluation.
- System capabilities grow as agents, data & information are added.



References

Gale, J.F.W., R.R. Reed, and J. Holder, 2007, Natural fractures in the Barnett Shale and their importance for hydraulic fracture treatments: AAPG Bulletin, v. 91/4, p. 603-622.

Hill, R.J., D.M. Jarvie, J. Zumberge, M. Henry, and R.M. Pollastro, 2007, Oil and gas geochemistry and petroleum systems of the Fort Worth Basin: AAPG Bulletin, v. 91/4, p. 445-473.

Hill, R.J., E. Zhang, B.J. Katz, and Y. Tang, 2007, Modeling of gas generation from the Barnett Shale, Fort Worth Basin, Texas: AAPG Bulletin, v. 91/4, p. 501-521.

EGI

Thank you.