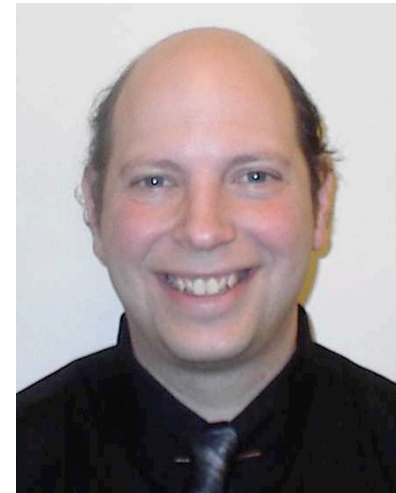


Some Ideas on Research Directions in: KR and Info Integration for Financial Services

Benjamin Grosf*

July 21-22, 2010



Working material for the National Science Foundation Workshop**

Sr. Research Program Manager, Vulcan Inc.; and
Principal Consultant, Benjamin Grosf & Associates, LLC
<http://www.mit.edu/~bgrosf/>

** Organized by University of Maryland, Robert H. Smith School of Business;
held at Waterview Conference Ctr., Arlington, VA, USA

Outline

- **Bigger Ideas**
- **Smaller Ideas**
- **Appendix: Semantic Web Primer**

Use Semantic Web and other KR Technology

- **Represent and Integrate**

- Financial data
- Instrument descriptions
- Regulations and laws
- Business/government policies
- Economic statistics and financial
- Other relevant aspects of business/government processes

- **New building blocks beyond conventional data mgmt.**

- Semantic Web: rules, query; naming, vocabulary, ontologies, schemas
 - Recent progress esp. in rules, ontologies, querying
 - Higher-order defaults. Monitoring event flows. Parallelization. Tools. Standards.
- Probabilistic and Strategic AI
 - Machine learning, data mining, statistics
 - Game theory, “mechanism design”, utilities, decision theory

Converge XBRL etc. with main Semantic Web KR

- “Etc.” here means roughly-similar-flavor financial/economic data
- **Converge XBRL with main Semantic Web**
 - Focus on Rules, leverage recent web rule standards
 - W3C RIF, W3C OWL RL, OMG SBVR
- **Develop (more) financial vocabulary and rules**
- *Background on XBRL (“eXtensible Business Reporting Language”)*
 - *Standardized web format for financial reporting data – “the main game in town”*
 - *Usage already required by SEC and many other countries’ regulatory/tax agencies*
 - *Primarily for public companies. Used also internally for compliance and CFO function.*
 - *Data is in XML. Includes vocabulary and rules of accounting definitions.*
 - *Grew up in parallel with Semantic Web standards and technology*
 - *Developed by accountants more than computer scientists*

E-contracts and E-law

- *Background*
 - *Advanced semantic rules well represent most logical aspects of contracts and regulations*
 - *Technically: Involves conflict handling and exceptions (prioritized defaults)*
 - *Technically: Involves meta-knowledge, e.g., about provenance (higher-order)*
- **Represent logical content of contracts and legal provisions, in:**
 - Financial instruments and transactions
 - Regulations and other laws
 - Finer-grain representation of investment strategies, instruments, and vehicles
- **Derivatives and structured finance**
- **Analyze and aggregate**

Open-Source Model of Financial System (“Finux”)

- **Collaboratively developed**
- **Simulation and what-if analysis**
- **Agent-based cognition to transcend “math-based” blindnesses**
 - Business/government processes and contracts/laws
 - Gaming strategy and herd behavior
 - Stochastics and lags
 - Use machine learning to estimate parameters, lags
- **Combine “in the small” with “in the large” risk management**

New Govt. Funding Org. for Financial KR/IT R&D

- **Potential org. models: more NIH and DARPA than NSF**
 - Relationship to Office of Financial Research?
- **Finance is primarily a cyber-industry**
- **In the arms race, the public and the regulators are technologically way behind**
- **Budget scale that's justifiable**
 - ~\$25M in yr 1 could have considerable impact initially
 - Accelerate development of uncontroversial information models
 - Analyze KR requirements and attack soft spots; proof of concepts
 - Grow fast, e.g. ~\$50M yr 2, \$100M yr 3, \$150M yr 4, ...
 - Long-term: grow to, e.g., ~1 basis point on US financial assets
 - thus ~\$1B/yr. **Cheap insurance against another \$multi-trillion hit.**

KR Challenges Needing Applied Research

- **Combine numerical reasoning, more closely**
 - Equalities and equations; Inequalities, “constraints”; Integrals
 - Money, utility; Time (and dates)
 - Probabilities, statistics
- **KR context mappings (reformulations)**
- **Representing contracts, regulations, policies, legal aspects**
- **Bring spreadsheets into the semantic KR world**
- **Combine probabilistic reasoning, more closely**
- **Combine processes descriptions**
- **Map English to and from KR, for knowledge entry and explanation**

Smaller Ideas

Public debate wiki about financial public policies

- **Prime topic: Regulatory reform**
- **Tool opportunities:**
- **Semantic wiki software as infrastructure**
 - e.g., Semantic MediaWiki+ with plugins for office productivity, semantic web
- **Argumentation systems**
 - e.g., cf. MIT Ctr. for Collective Intelligence, plus default rule systems

Other Ideas

- **Track closed/merged/acquired co.'s/funds**
- **Apply game-theoretic dynamics and incentives**
 - Analyze market decomposition. Automated mechanism design.
- **Analyze co./fund control, pay practices not just ownership**
- **Expose "invisible" leverage**
 - Systemic, as well as per-deal
 - undercapitalized insurers (AIG), single-movers (sovereign dollars)
- **Expose opacity**
- **Expose securitization-based liquidity amplifications, flows**
- **Company "living wills"**

Appendix:

Semantic Web Primer

Heart of the SW Revolution

- Pre-Semantic Web (1.0/2.0) breakthrough:
 - Radically improves sharing of human-readable info
- Semantic Web (3.0+) breakthrough:
 - Dramatically improves sharing of machine-readable info
 - Info structure based on Knowledge Representation
 - Logical principles that sanction what inferences should/should-not be drawn from what's explicitly communicated
 - Rules + Facts = Structured Knowledge

Today's SW Standards & Core Tech

- Phase I: basic database schemas (RDF, OWL-DL)
 - Filled industry vacuum, enabling shareability
- Phase II: database queries and simple rules (SPARQL, OWL-RL, RIF)
 - Leverages core of legacy database and business rules technology
 - RIF (Rule Interchange Format) has the most general framework

What's Next for the Core of SW

- SILK – Rules that extend RIF
 - Defeasible: permit exceptions, handle conflicts
 - Cope with knowledge quality and context
 - Reactive: take actions, based on event flows
 - Activate knowledge
 - Higher-order: knowledge about knowledge
 - Knowledge modularity, dynamism, lifecycle
- Raise the KR abstraction level underlying structured data/knowledge management
 - Most significant since relational databases and business rule systems invented in the 1980's

What's Next for the Core of SW, cont'd

- Tools for rules will take a while to mature:
 - Engines for inferencing+action
 - User interfaces for creating and editing rules
- *Longer-Term*: Deep Probabilistic and Statistical knowledge representation
 - Shareable data mining and inductive learning
 - Natural language processing

Thank You

Disclaimer: The preceding slides represent the views of the author only.
All brands, logos and products are trademarks or registered trademarks of their respective companies.