BIG DATA: CHALLENGE OR OPPORTUNITY FOR THE FinTech INDUSTRY?

Abraham Bernstein
• Volume
• Velocity
• Variety
• Veracity
Intelligent System

Data Integration

Action Selection

Action

Monthly values for the AMO index, 1856-2013
Data Integration
Data Space
## Data Semantics

<table>
<thead>
<tr>
<th>Good</th>
<th>Price</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apples</td>
<td>4.75</td>
<td>3</td>
</tr>
<tr>
<td>Oranges</td>
<td>4.25</td>
<td>4</td>
</tr>
</tbody>
</table>

...
Data Semantics

Financial Industry Business Ontology™

The Financial Industry Business Ontology (FIBO) is a business conceptual ontology developed by the members of the EDM Council. FIBO provides a description of the structure and contractual obligations of financial instruments, legal entities and financial processes. FIBO is used for harmonization of data across repositories as a common language (i.e. Rosetta stone) for risk analysis and business process automation. FIBO is expressed in the triplestore language of the Web (RDF/OWL) for machine readable inference processing and UML for people readable analysis.

Access to JIRA/Wiki and GitHub requires a username and password. The links that follow will not be accessible until you are logged in: 1) register with GitHub, 2) send your GitHub username to Dennis Wisnosky for permission to access FIBO Data Repositories, 3) log into GitHub before clicking on the Wiki links below.

- Development Process: The Council manages FIBO via GitHub (for content management), JIRA/Wiki (for project management) and Jenkins (for testing).
- FIBO-Vocabulary: Terms and definitions for financial instruments and processes
- FIBO URI: The Uniform Resource Identifier for FIBO is maintained in GitHub.
Data Semantics

\[ P(A|E) = \frac{P(E|A)P(A)}{P(E)} \]
Graphical Models
MRF - CRF - PSL

Brown and Lowe, 2007
Graphical Models
MRF - CRF - PSL
Data Integration on the Web
Intelligent System

Action Selection

\[ P(A|E) = \frac{P(E|A)P(A)}{P(E)} \]
Decomposing Fraud Patterns

Participants can be labeled as:
- Splitters
- Aggregators
- Forwarders
Traditional Processing

- Volume
- Velocity
- Variety
- Veracity
Flow Processing
Matching Financial Transactions With the Subset Sum Algorithm
Elicitation Process
Action: Fraud Patterns in Bitcoin Transactions

Dataset Size: 50M Transactions, Matching Duration: 1 week

- Volume
- Velocity
- Variety
- Veracity
Intelligent System

Action Selection

\[
P(A|E) = \frac{P(E|A)P(A)}{P(E)}
\]
ROBOT FINANCE?
Intelligent System

- Volume
- Velocity
- Variety
- Veracity

Data Integration

Action Selection

Action

Monthly values for the AMO index, 1856-2013