Converting Text into FIBO-aligned Semantic Triples
Outline

• Introduction: Text to triples
• FIBO-driven Use Cases:
  – Contract processing
  – Transaction information extraction
• How It Works
• Applications
Introduction: Text to Triples
Text Mining Framework

TEXT
- Emails
- Articles
- Reports
- Contracts
- Web
- Social

Foreign languages
Different domains

KNOWLEDGE
- Concepts
- Semantic Relations
- Events
- Sentiments
- Opinions
- Ontologies

APPLICATIONS
- Text Analytics
- Question Answering
- Intelligent Search
- Decision Support
- Info Extraction
- Doc Classification

Drive your applications
Text Mining Framework

**TEXT**
- Emails
- Articles
- Reports
- Contracts
- Web
- Social

**FOREIGN LANGUAGES**
- Different domains

**KNOWLEDGE**
- Concepts
- Semantic Relations
- Events
- Sentiments
- Opinions
- Ontologies

**APPLICATIONS**
- FIBO-aligned
- Question Answering
- Intelligent Search
- Decision Support
- Info Extraction
- Doc Classification

**Drive your applications**
FIBO Use Cases
FIBO Triples: Motivation

• Federation: knowledge from text together with RDBMS
• Visualization & Analytics
• Natural language querying:
  — What are the agreements with not-for-profit companies that has effective date in March 2017?
  — What are the assets of ACME Inc.?
• Ontology extension/population based on proprietary or regulatory documents
Use Case: Contract Processing

LICENSING AGREEMENT

This agreement is made between Lymba Corporation (the "Licensor") and ACME Inc. (the "Licensee"), collectively referred to as the "Parties".

Named Entity: Organization

Named Entity: FIBO Concept
Use Case: Contract Processing

More items to extract:

• Parties and their attributes
• Time restrictions
• Money values and their roles (penalty, fee, etc)
• Jurisdiction
• Governing laws
Sonoco (NYSE: SON) today announced it has completed the sale of its rigid plastics blow molding operations to Amcor, a global leader of packaging products, for $280 million.
How It Works
Customizable Semantic Knowledge Extraction

- Preprocessing of documents: pdf, tables, structure recognition
- Entities (over 86 standard types): persons, organizations, locations, dates/times, etc.
- Semantic relations (26 basic types): Agent, theme, instrument, location, etc.
- Custom relations and entities
- RDF/TriX representation of knowledge
Layered Extraction

- Named Entity extraction
- Basic Relations
- Iterative application of extraction and inference rules to extract more knowledge:
  \[ \text{NE}(X, \text{XType}) \land \text{ISA}(X, C) \land \text{NE}(C, \text{FIBO}) \land \text{IsCompatible}(\text{XType}, C) \]
  \[ \rightarrow \text{IsInstanceOf}(X, C) \]

\[ X = \text{Lymba Corporation} \]
\[ C = \text{Licensor} \]
The Process

- Decide on data model: FIBO fragment+
- Data/document sampling
- Collect gold annotations for testing and training
- Annotator agreement
- Model refinement
- Iterative training/testing of the solution
Applications
International Economics Is Different

Nations are not like regions or families. They are sovereign, meaning that no central court can enforce its will on them with a global police force. Being sovereign, nations can put all sorts of barriers between their residents and the outside world. A region or family must deal with the political reality that others within the same nation can outvote it and can therefore coerce it or tax it. A family or region has to compromise with others who have political voice. A nation feels less pressure to compromise and often ignores the interests of foreigners. A nation uses policy tools that are seldom available to a region and never available to a family.
Applications in Financial Domain

- Contract processing: information extraction, metadata enrichment & semantic search
- Compliance requirements understanding: who is responsible for what
- Risk assessment
- Customer support
- Report generation
- Money laundering detection