

# Ontological Representations for Supporting Learning in Business Communities

**Bojan Cestnik**

Temida d.o.o. & Jozef Stefan Institute, Ljubljana, Slovenia  
bojan.cestnik@temida.si

**Tanja Urbančič**

University of Nova Gorica, Slovenia  
tanja.urbancic@ung.si

**Ingrid Petrič**

University of Nova Gorica, Slovenia  
tanja.urbancic@ung.si

**Presented by**

**Stojan Košti**

Temida d.o.o., Ljubljana, Slovenia  
stojan.kosti@temida.si

# Contents

---

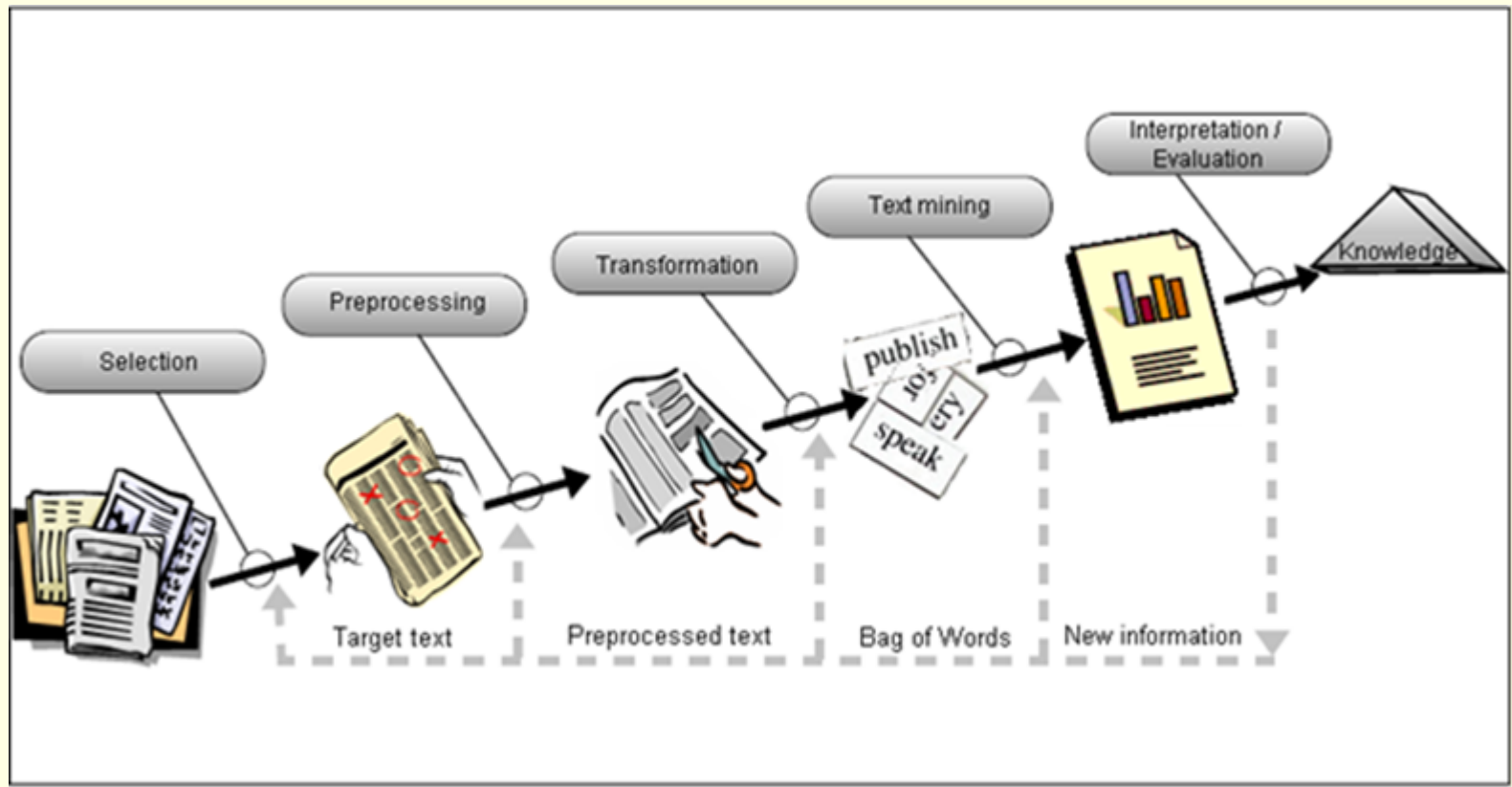
- Introduction
- Background
- Ontology construction case
- Evaluation of results
- Conclusions

# Introduction

- Learning tasks in business communities
  - Adequately recognize partners' products and skills
  - Explore heterogeneous data sources
  - Modern techniques (data mining, text mining, etc.)
- Knowledge acquisition from various sources
- Business knowledge representation: topic ontology

# Extracting knowledge from text

- Text mining steps:



# Introduction to ontologies

- Ontologies
  - Organize scientific information
  - Common vocabulary of concepts
  - Contribute to common understanding of problem domains
- Formal representation of a shared conceptualization of a particular domain
- Construction of ontologies
  - Manual (e.g. Protégé)
  - Semi-automatic (e.g. OntoGen)

# NHSS *Protégé* ontology

**CLASS BROWSER**

For Project: ●

Class Hierarchy

- :THING
- :SYSTEM-CLASS
- ▼ ● NHSS GENERAL TERMS
  - CONTROLLING
  - PENALTIES
  - TEMPORAL TERMS AND CONDITIONS
- ▼ ● EARMARKED SAVINGS
  - CATEGORIES OF SAVERS
  - INCLUDED BUSINESS BANKS
- ▼ ● TERMS OF SAVING
  - TYPE OF SAVING
  - DURATION
  - INTEREST RATES
  - PREMATURE TERMINATION
  - PREMIUM ACCURALS DETERMINATION
  - HOUSING LOANS
- ▼ ● REPORTING
  - **CONTRACTS**
  - PAYMENTS
  - PREMIUM ACCURALS
  - LOANS

Superclasses

- REPORTING

**CLASS EDITOR**

For Class: ● CONTRACTS (instance of :STANDARD-CLASS)

Name: CONTRACTS

Documentation: BANK ACCOUNT, CONTRACT DURATION, CONTRACT DATE, MONTHLY INSTALLMENT, NAME, ADDRESS

Role: Concrete ●

Template Slots

Name	Cardinality	Type
ADDRESS	single	String
BANK ACCOUNT	single	String
CONTRACT DATE	single	String
CONTRACT DURATION	single	String
MONTHLY INSTALLMENT	single	String
NAME	single	String

# NHSS *OntoGen* ontology

The screenshot displays the OntoGen software interface, which is used for creating and visualizing ontologies. The interface is divided into several panels:

- File:** A menu bar at the top left.
- Concepts:** A tree view on the left side showing the hierarchy of the ontology. The root node is expanded, showing sub-nodes such as "General terms", "Earmarked savings", "Terms of saving", "Premium accruals", "Housing loans", and "Reporting".
- Concept properties:** A panel at the bottom left for editing the properties of the selected concept. It includes fields for "Id" (0), "Name" (root), "Keywords" (slovenije, posojila, pogodb, republike, republike\_slovenije, na, varcevanja, zakona, tega\_zakona, za), "SVM Keywords", "All documents" (45), "Unused documents" (0), and "Avg. similarity".
- Ontology details:** A panel at the top right showing the current visualization settings, including "Ontology visualization", "Concept's documents", and "Concept Visualization". It also includes sliders for "Concept font size" (12) and "Relation font size" (0).
- Ontology visualization:** A large central area displaying a hierarchical diagram of the ontology. The root node is highlighted in red. The diagram shows the following structure:
  - root
    - Procedural remarks
    - Requirements for accruals
    - Premium accruals
      - Contract termination
    - Housing loans
    - Reporting
    - Terms of saving
      - Interest rates
      - Saving duration
      - Saving types
      - Eligible savers
    - Earmarked savings
      - Participating banks
      - Saver categories
    - General terms
      - Terms and conditions
      - Controlling mechanisms
      - Temporal terms

# Ontology learning from text

- Constituting parts of ontology
  - Terms *tender, apartment, buyer*
  - Synonyms *apartment, flat, dwelling*
  - Concepts *apartment: = {1-room, 2-room, ...}*
  - Taxonomy *is\_a(buyer, person)*
  - Relations *buy(buyer, apartment)*
  - Rules  $\forall x, y : (has\_a\_child(x, y) \rightarrow family(x))$
- Methods
  - Statistical
  - Linguistic
  - Hybrid

# Areas of text processing

- **Information retrieval** – Search in DB
- **Natural language processing** – Computational linguistics
- **Text mining** – Data analysis
- **Semantic web** – Knowledge representation and Reasoning

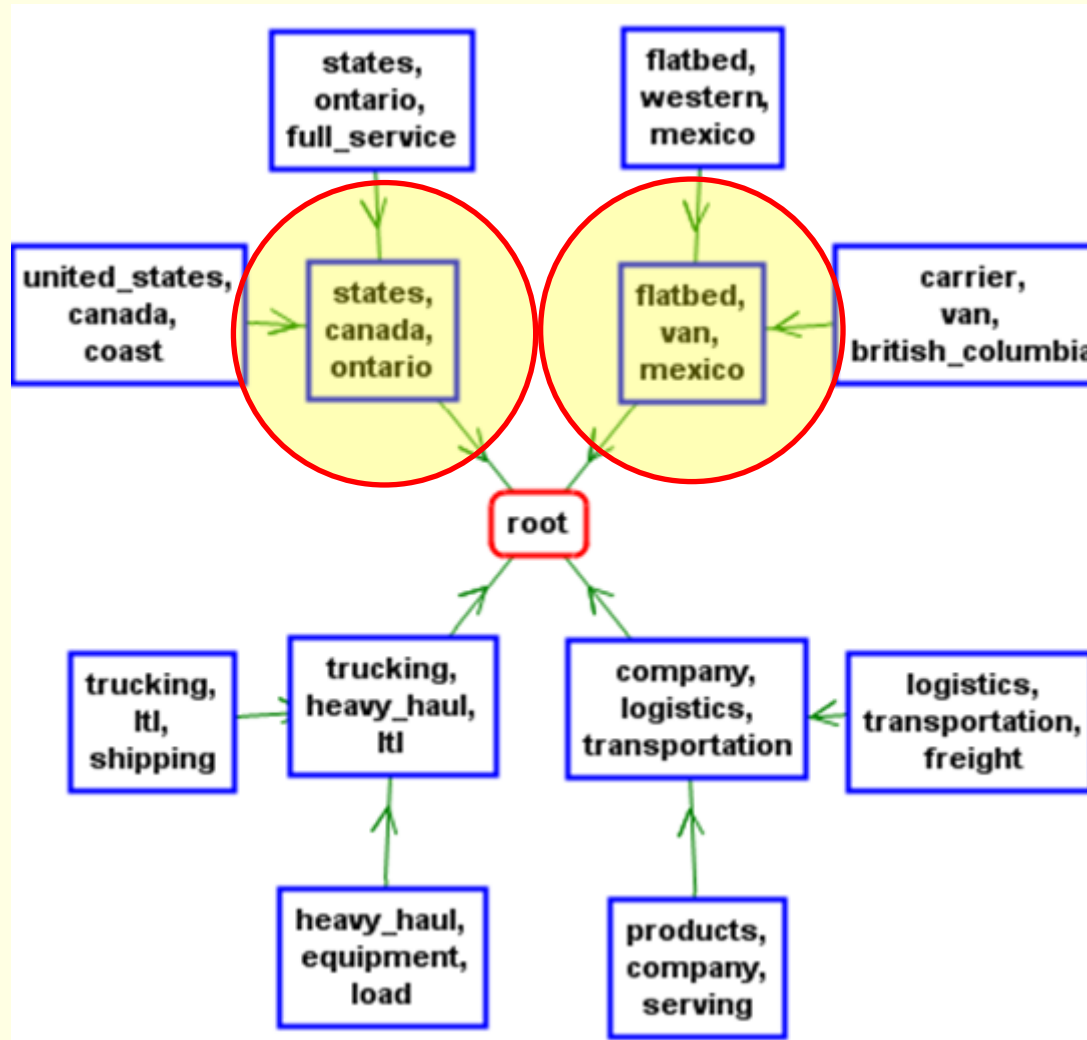
# Elements of text processing

- Word
- Sentence
- Paragraph
- Document
- Collection of documents

# Ontology construction case

- Input document formats: XML, HTML, doc, pdf
- Preprocessing
  - Removing graphics, markup and meta tags
  - Lemmatization
  - Stop-words removal
- Tool: OntoGen
- Case: North American transportation domain
- Input documents
  - Professional descriptions of NA transport companies
  - From Open Directory Project DMOZ

# Topic ontology: cluster descriptions



# Document map



company  
name  
keyword  
cluster

# Conclusion I

- Ontologies for efficient learning in business communities
- OntoGen: constructing ontology for a given problem domain from description documents
  - Obtain overview of a given problem domain
  - Requires substantially less time to construct
  - Can be useful for exploiting existing but overlooked knowledge

# Conclusion II

- Constructed ontology quickly reveals top level concepts
  - Birds-eye view over a given problem domain
  - Result: easier and more effective business concepts identification and management