A Hybrid Data and Text System for Downgrading Sensitive Documents

Mikhail J. Atallah, Victor Raskin
with
Christian F. Hempelmann,
Dina H. Mohamed

The Problem

• Extensively required for security purposes
• Types of suppressing sensitive information:
  • weak declassification
  • strong declassification
  • downgrading, sanitizing
  • on the fly downgrading, sanitizing, surveillance
• different types of representation: text / data
State of the Art

• Hundreds of declassification rules
• Human implementation of declassification is slow and very costly.
• Downgrading rules are often not even codified.
• No rules for a declassified copy
• So far no work to filter e-mail on the fly

Ontological-Semantic Approach

• Based on computer understanding of the information
• Takes full advantage of the new technologies in computational semantics
• Uses 3 major resources:
  • 1. lexicon (words of an natural language explained in terms of an ontological concept)
  • 2. ontology (a tangled hierarchy of concepts)
  • 3. text-meaning representation (TMR): (composes sentential meaning out of ontological concepts)
Defining Illegal Information

• Task of the customer
• Illegal information must be translated into ontological concepts
• e.g. allow mentions of nuclear submarines, not their specific deployment, reactor capacity, or mode of refueling

Minimal Analysis

• Analyzer spots nuclear submarine, or even submarine
• Analyzer evokes the appropriate lexical entry
• Lexical entry produces the corresponding ontological concept
Minimal Analysis – contd.

submarine
(isa warship)
(theme-of build, commission, decommission, deploy, destroy, attack)
(instrument-of attack, support, transport, threaten)
(manned-by naval crew)
(propel-mode surface, sub-surface)
(engine-type nuclear-engine)
(range $N < x < M$)
(speed $K < y < L$)
(current-location body-of-water and/or geographic point and/or coordinates and/or relative, time-range)
(prior-location …)
(next-location …)
(current-mission $Z$)

Minimal Analysis – contd.

• General principles:
  • Most of the slot fillers of concepts are ontological concepts as well
  • Analyzer is interested only in selected slots (e.g. LOCATION), or specific subslots (e.g. CURRENT LOCATION)
  • If the slot is empty, the text will be passed down unmodified
  • Depth of meaning penetration is limited to 3 levels, with a possible enhancement to 5
Problems

• Literal meaning:
  • Generality of the type of illegal information
  • Determining the exact parts to be deleted or replaced
• Non-literal meaning: Inferences, Implicatures, Presuppositions, etc.

Sophisticated Analysis

• For problematic sections:
  • Full analysis
    but locally restricted to a rigidly circumscribed amount of text
  • User intervention
    for marked section, i.e. step down from automatic approach
Downgrading the Text

• General Strategies:
  • Suppress any illegal information
  • Keep as much of the original wording as possible when rewording is necessary
  • Keep the changed wording as consistent as possible throughout
  • Make the necessary modifications to make up for any damage done to the text

Downgrading the Text - contd.

• General types of downgrading:
  1. **Deletion**: of one or more lexical items, one or more sentences, one or more paragraphs, etc.
  2. **Replacement**: using a more generic synonym, (taking one or more steps upward in the ontology)
  3. **Addition**: rare, mostly of articles; mainly to compensate for the damage done to the structure of the text by deletions.
Degree of Automation

- **Manual:**
  - **Defining Illegal Information:** determining what is classified and how it should be treated.
  - (Optional) **Post-Editing:** double-checking that the automatic NLP module below did all right.

- **Automatic:**
  - **NLP Implementation:** all the modifications of the sensitive data as per human definitions above.

---

### Degree of Automation - Graph

- **Automatic NLP Module**
  - Static Resources
  - Ontology
  - Lexicon
  - PSI

- **Human Security Expert Module**
  - SIL
  - PEM

**Abbreviations:**
- PSI: Permitted Substitution Inventory
- MPAG: Meaning-preserving Analyzer-Generator
- SIL: Sensitive Information List
- PEM: Post-editing Module (optional)
Example – Original Text

“I have a dream”
Fivescore years ago, a great American, in whose symbolic shadow we stand today, signed the Emancipation Proclamation. This momentous decree came as a great beacon light of hope to millions of Negro slaves who had been seared in the flames of withering injustice. It came as a joyous daybreak to end the long night of their captivity. But one hundred years later, the Negro still is not free; one hundred years later, the life of the Negro is still sadly crippled by the manacles of segregation and the chains of discrimination; one hundred years later, the Negro lives on a lonely island of poverty in the midst of a vast ocean of material prosperity; one hundred years later, the Negro is still languished in the corners of American society and finds himself in exile in his own land.

Example – Downgraded Version

“I have a dream”
In the past, many great Americans, in whose symbolic shadows we stand today, signed important documents. These momentous decrees came as a great beacon light of hope to millions of oppressed citizens who had been seared in the flames of withering injustice. They came as a joyous daybreak to end the long night of their captivity. But today, the oppressed citizen still is not free; today, the life of the oppressed citizen is still sadly crippled by the manacles of segregation and the chains of discrimination; today, the oppressed citizen lives on a lonely island of poverty in the midst of a vast ocean of material prosperity; today, the oppressed citizen is still languished in the corners of American society and finds himself in exile in his own land.