CERIAS

the center for education and research in information assurance and security

Extension of NLP Techniques for Privacy Management

Olga Krachina, Victor Raskin, CERIAS, Purdue University

Motivation:

❖ deployment of Privacy Policies (PP) occurs on the level of a formal language, but PP are written primarily in natural language

❖ level of natural language expressiveness necessary to accomplish a certain task or goal is unknown a priori

❖ inherent ambiguity associated with natural language creates obstacles for formal languages that have syntax-based structure and, as a rule, are domain-specific

Framework of Choice:

Ontological Semantics

A domain-independent NLP framework characterized by expressive semantics, comprehensive information representation modules (TMR), modularity, easily extendable to accommodate and process domain-specific knowledge (for detailed description of the framework see [2]).

Applications in the Field of Privacy:

❖ natural language to formal language translation

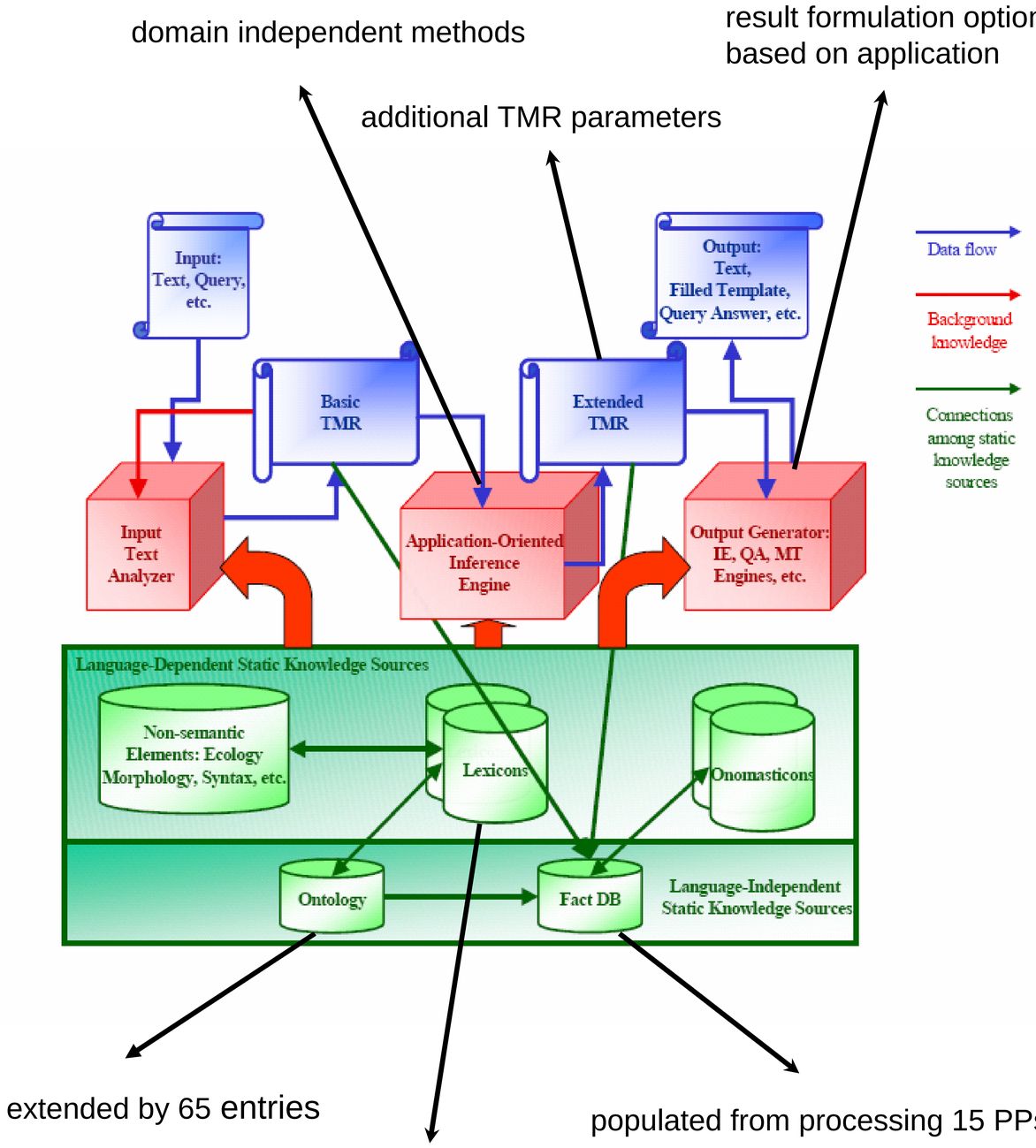
❖ tasks pertaining to inference process:

- policy compliance verification, i.e. testing alignment of a particular policy and regulation for that specific category [1]
- question-and-answering session, i.e. user-policy interaction

Future NLP Goals in the Domain:

abstracting a standard set of queries that are necessary to determine compliance, which will require development of a ranking system of the queries

Overview of basic components of Ontological Semantics: arrows indicate modifications to the framework necessary to accommodate domain of Privacy Policies.



Adopted from Nirenburg and Raskin, 2004

References:

1. Krachina O., Raskin V., Triezenberg K. 2007. Reconciling Privacy Policies and Regulations: Ontological Semantics Perspective, *to appear in* Proceedings of HCII 2007.

2. Nirenburg S., Raskin V., 2004. Ontological Semantics, Cambridge, MA: MIT Press.

PURDUE UNIVERSITY

CERIAS

Discovery Park e-Enterprise Center