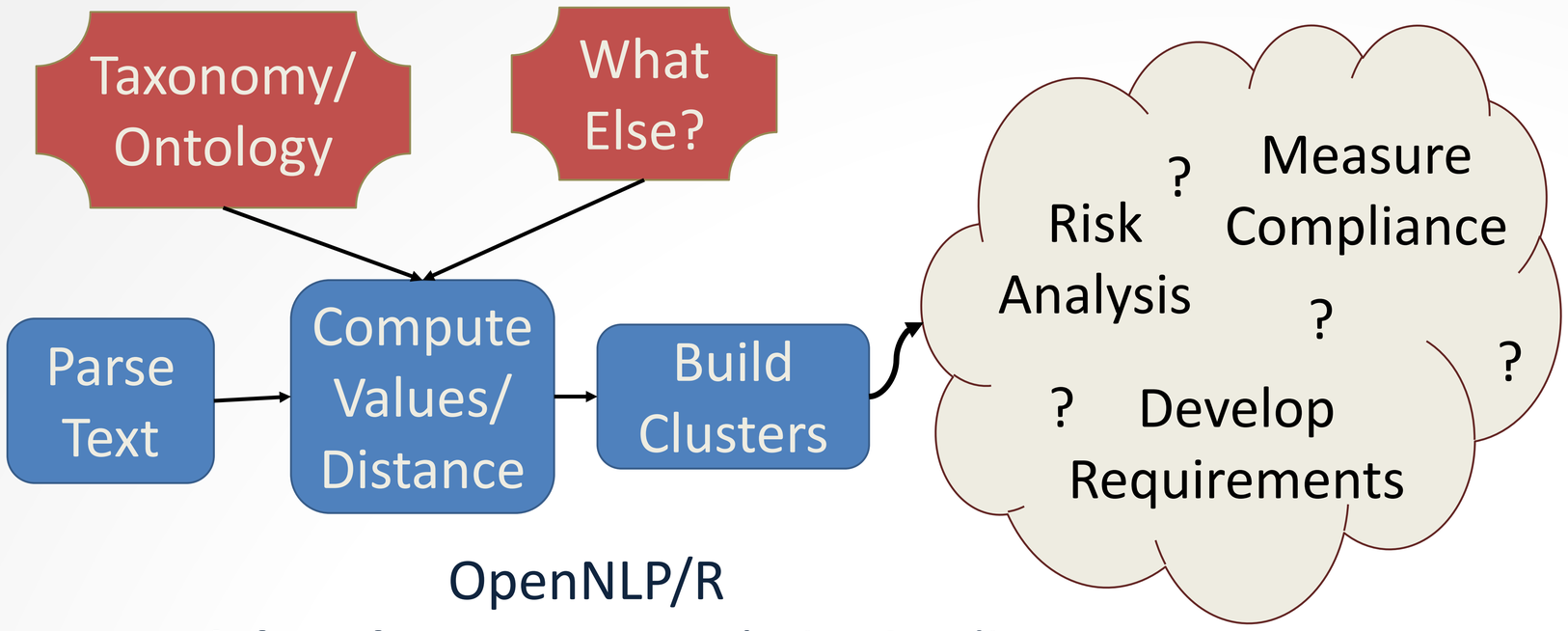
Semantic Similarity in Security Regulations

Sarah Elder Hui Guo Munindar Singh Laurie Williams

{seelder, hguo5, mpsingh, lawilli3 } @ncsu.edu

Identifying Semantically Similar Statements



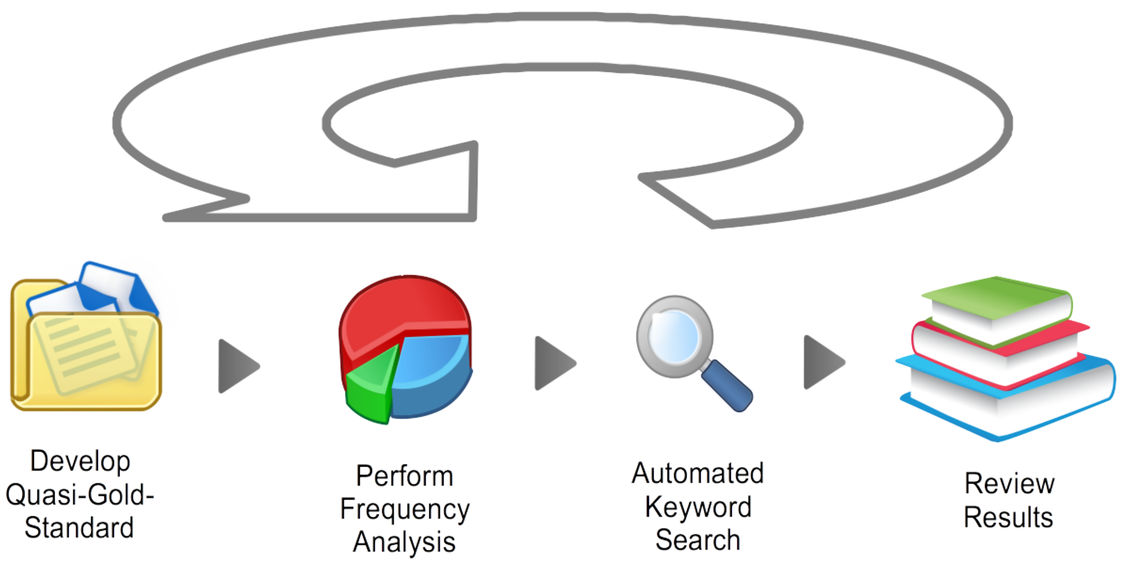
Potential Tools: OpenNLP/R
Watson Analytics Studio
MALLET
WEKA

Key Terms	
Semantic Similarity	similarity in meaning
Regulation	legal documents, organizational policies, etc.
Ontology/ Taxonomy	Ways to organize concepts, their attributes, and relations

Systematic Literature Review

How has semantic similarity been calculated in the domain of software security?

How have ontologies and taxonomies been applied in the security domain for the calculation of semantic similarity?



Methodology

- Based on Kitchenham and Brereton
- Quasi-Gold-Standard is a core set of documents in the domain. Will be based on sources cited in other SLRs
- Automated Keyword Search may be supplemented by Manual Search (Shakeri et. al)

Preliminary Findings

- Ontologies and taxonomies are often evaluated based on whether they contain all relevant concepts
- Ontologies/taxonomies often used in combination with each other
- Additional information content sometimes used

B. Kitchenham and P. Brereton, "A systematic review of systematic review process research in software engineering," *Information and software technology* 55, no. 12 (2013): 2049-2075.

Z. Shakeri Hossein Abad, G. Ruhe and M. Noaeen, "Requirements Engineering Visualization: A Systematic Literature Review," in *IEEE 24th International Requirements Engineering Conference*, Beijing, 2016.

HoTSoS *Symposium and Bootcamp*
HOT TOPICS *in the* **SCIENCE OF SECURITY**
APRIL 4-5, 2017 | HANOVER, MARYLAND