

SEMANTIC SEARCH IN E-DISCOVERY: AN INTERDISCIPLINARY APPROACH

DESI V Workshop

UvA: David Graus, Zhaochun Ren, Maarten de Rijke

HvA: David van Dijk, Hans Henseler

Leiden University & Fox-IT: Nina van der Knaap

ICAIL 2013 Conference, Rome, 14 June 2013

PROBLEM STATEMENT

- Explosion of information
- Beyond predictive coding
- From unstructured to structured information
- Text mining is still far from perfect
- We need more precision and we need better context
- Increase precision by focusing on specific problems
 - Finding privileged information
 - Finding facts related to violation of competition laws
 - Find facts related to bribery

FOUR SUB PROJECTS

- Subproject 1 :
 - Criminology/Law
- Subprojects 2-4 :
 - Information Extraction/Information Retrieval

SUBPROJECT 1: SCOPE AND REQUIREMENT ANALYSIS OF SEARCH PROCESS

- Multidisciplinary literature study
 - IT
 - Criminology
 - Law
- Interviews with professionals and practitioners

SUBPROJECTS 2-4: INFORMATION EXTRACTION & RETRIEVAL

- Semantic Search: Applying **structured knowledge** as a complement to **text retrieval**
- Using techniques from Information Extraction and Information Retrieval

Scope and Requirements Analysis



Investigators

 Search Interface |

Exploration and contextualization based on **entities**

Exploration and contextualization based on **relations**

Exploration and contextualization based on **topics**

Semantic Search

Entity Extraction

Relation Extraction

Topic Extraction

Semantic Analysis

THANK YOU

j.henseler@hva.nl

www.hva.nl