

Supporting Collaboration by Large Scale Email Analysis



Michal Laclavík, Martin Šeleng, Ladislav Hluchý

laclavik.ui@savba.sk, <http://ikt.ui.sav.sk/>

Institute of Informatics, Slovak Academy of Science, Dúbravská cesta 9, 845 07 Bratislava, Slovakia

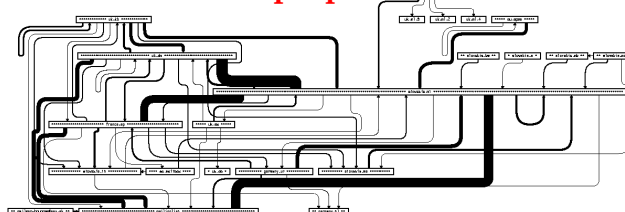
Motivation

To exploit information and knowledge included in email communication

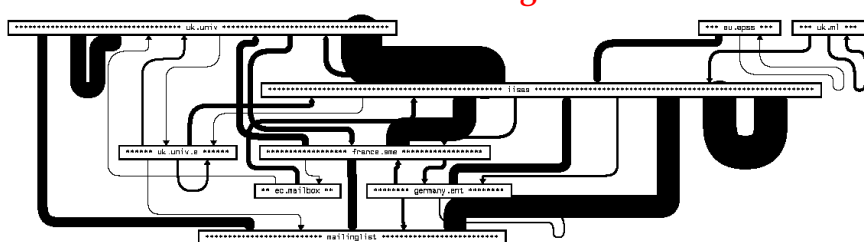
Approach

- Social Network Extraction
- Objects detection – adding of semantics
Persons, Organizations, Locations, Contact data

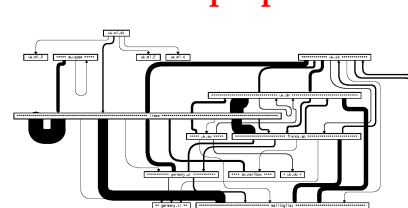
network of people



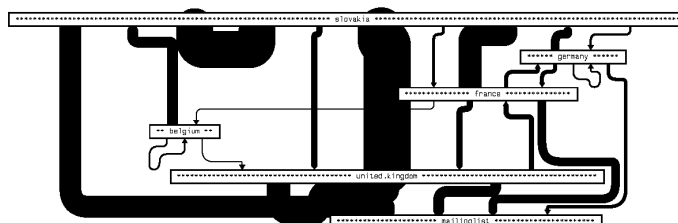
network of organizations



network of people and IISAS



network of countries

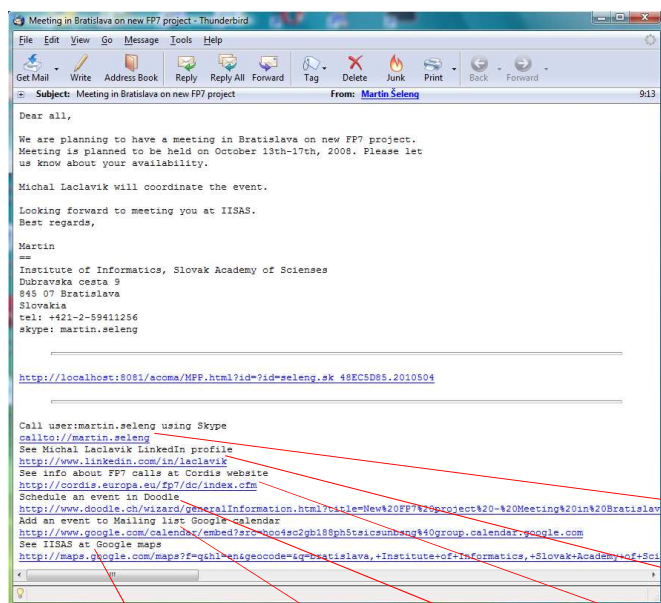


Used Technology

- Regular Expressions
- RDF/OWL
- Java Mail
- MapReduce

Used Tools

- Jena
- Ontea, Acoma
- aiSee
- Hadoop



Features

- Social network of communicating people
- Including level of interaction (how many email's were send and received)
- Social network graph transformed to different graphs based on email addresses, people or organizations
- Graph can be enriched with other objects mentioned in communications depending on ontology model: e.g. organizations, enterprises, people or geographical locations
- Large scale email archive processing on MapReduce
- Collaboration through adding hints into the email message such as – see location on the map; manage event in calendar or doodle; use contact information

