Roundtable Discussions: Descriptions - Tuesday, October 18

Roundtable 1: Data visualizations - dashboards for text analytics

The results of text mining can yield precise answers using sophisticated techniques. These can be viewed in tables, but there are many occasions when users need to look at the whole set of results to analyze distribution, identify trends or spot anomalies. This discussion will look at existing use cases, discuss what kind of visualizations are important and look at what kind of reporting dashboards are needed to facilitate this.

Roundtable 2: Mining patient data: challenges and opportunities

Interest in mining unstructured patient data has grown rapidly in recent years to support improved analytics, disease understanding and patient care. This group will discuss the challenges faced by healthcare organizations to mine pathology data, socioeconomic factors and disease insights for example. It will look at data input from EHRs, dealing with format variation and using I2E to explore data sets

Roundtable 3: Big data analytics

Big data refers to the volume, variety, velocity and veracity of information that has to be dealt with across an Enterprise. It can range from terabytes to petabytes of data in size and rate at which it has to be processed can be overwhelming. This roundtable will discuss strategies, technologies and frameworks being employed to deal with this.

Roundtable 4: Improving usability for experienced text miners

Linguamatics is working on making I2E more accessible to occasional users, but what are the issues that experienced users face when developing large or sophisticated queries? What existing techniques do people use to version queries, avoid cut-and-paste and perform evaluation?

Roundtable 5: Text mining full text literature - challenges and opportunities

It is relatively easy to text mine abstracts from articles in scientific journals. MEDLINE is the one of the most commonly-used data sources for these abstracts and Linguamatics provides cloud-based access to this. However, in some cases, abstracts may not be good enough. Has any information been missed? The only way to be sure is to search through the complete article. In the past, this has been a time-consuming process involving many steps:
- Identify the full text articles for text mining
- Obtain permission from the publisher to text mine the content (publishers typically do not allow text mining by default)
- Download the articles
- Index the documents with I2E using the appropriate configuration files
- Text mine the articles and extract the appropriate information

This session will discuss existing processes and what would be desired in a future solution to address any hurdles.

Roundtable 6: Regulatory/IDMP

Identification of Medicinal Products (IDMP) is a new set 5 ISO standards to define, characterize and identify regulated medicinal products. The information gathering and submission of information requires all pharma companies to meet some immediate deadlines. This roundtable is to discuss strategies to extract information for submission from the variety of structured and unstructured data sources within an organization