

Progress of eDiscovery in Pharmaceutical Research

1. The Swiss IT company InfoCodex AG *Semantic Technologies* has developed a unique tool for information refinement and cross-lingual "understanding" of documents, opening up new possibilities in the support of pharmaceutical research.

2. The grand vision of discovering new relationships through text mining of large sets of biomedical publications has, to some extent, become reality. The key to success lies in the ability to categorize unstructured information and to extract the relevant facts and relationships from the vast flood of information. In a benchmark, InfoCodex reached the very high clustering accuracy of 88% (SEMTECH London 2011).

3. Breakthrough discoveries are often the result of combinations of small, seemingly unrelated facts and relationships into the bigger picture – like a puzzle which only reveals its picture once all the pieces are pulled together.

This is particularly true for pharmaceutical research where a correlation between disease and biomarker goes often unnoticed in isolated pieces of work, but taken as a whole, a clear relationship may emerge.

4. Two examples:

- A large pharma company has conducted a comprehensive experiment last year to investigate the power of InfoCodex in the *discovery of novel biomarkers* for specific diseases by analysing large sets of medical publications (PubMed etc.). The experiment has proven that the InfoCodex technology indeed can identify new potential biomarkers, enabling accelerated and targeted research (submitted for publication in a joint paper).
- A consortium of 5 University Hospitals has designed an FP7 research project to build a *repository of integrated medical data* gathered from distributed clinical records, containing masses of valuable but unexploitable patient data. InfoCodex was chosen as the technology partner for *extracting and cleansing* the concealed, unstructured data.

5. The reported approach, that employs autonomous self-organising semantic engines, is not restricted to discovering novel biomarkers. It could equally well impact pharmaceutical research, for example to shorten time-to-market of novel drugs, or speed up early recognition of dead ends (adverse events).