

3rd International Conference on

Big Data Analysis & Data Mining

September 26-27, 2016 London, UK



Carlo A Trugenberger

InfoCodex AG-Semantic Technologies, Switzerland

Scientific discovery by machine intelligence: A new avenue for drug discovery

The majority of big data is unstructured and of this majority the largest chunk is text. While data mining techniques are well developed and standardized for structured data; numerical data, the realm of unstructured data is still largely unexplored. The general focus lies on information extraction, which attempts to retrieve known information from text. The Holy Grail however is knowledge discovery, where machines are expected to unearth entirely new facts and relations that were not previously known by any human expert. Indeed, understanding the meaning of text is often considered as one of the main characteristics of human intelligence. The ultimate goal of semantic artificial intelligence is to devise software that can understand the meaning of free text, at least in the practical sense of providing new, actionable information condensed out of a body of documents. As a stepping stone on the road to this vision I will introduce a totally new approach to drug research, namely that of identifying relevant information by employing a self-organizing semantic engine to text mine large repositories of biomedical research papers, a technique pioneered by Merck with the InfoCodex software. I will describe the methodology and a first successful experiment for the discovery of new biomarkers and phenotypes for diabetes and obesity on the basis of PubMed abstracts, public clinical trials and Merck internal documents. The reported approach shows much promise and has potential to impact fundamentally pharmaceutical research as a way to shorten time-to-market of novel drugs, and for early recognition of dead ends.

Biography

Carlo A Trugenberger has earned his PhD in Theoretical Physics in 1988 at the Swiss Federal Institute of Technology, Zürich and his Master's in Economics in 1997 from Bocconi University, Milano. An international academic career in theoretical physics (MIT, Los Alamos National Laboratory, CERN Geneva, Max Planck Institute Munich) led him to the position of Associate Professor of Theoretical Physics at Geneva University. In 2001, he decided to quit academia and to exploit his expertise in Information Theory, Neural Networks and Machine Intelligence to design an innovative semantic technology and co-founded the company InfoCodex AG-Semantic Technologies, Switzerland.

c.trugenberger@infocodex.com

Notes: