Meetings International SciTechnol

J Forensic Toxicol Pharmacol 2018, Volume: 7
DOI: 10.4172/2325-9841-C2-012

Discovery of biomarkers by artificial intelligence

Carlo A Trugenberger
InfoCodex Semantic Technologies AG, Switzerland

The majority of big data is unstructured text. While data mining techniques are well developed and standardized for structured, 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 the new AI approach to drug research first pioneered by my company InfoCodex Semantic Technologies AG in collaboration with Merck and Thomson Reuters. This consists in employing a self-organizing AI engine to text mine extremely large repositories of biomedical research papers in search of unnoticed correlations enabling the discovery of relevant information for new drugs. 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.

c.trugenberger@infocodex.com