openM1N7ED
Open Mining Infrastructure for Text & Data

A TEXT MINING COMPANY TELLS ITS STORY

IMPROVING HEALTHCARE

Dr Jane Reed, Head of Life Science Strategy at Linguamatics

WHAT ARE YOU DOING WITH TEXT & DATA MINING?
We make TDM tools that help users to get value from their unstructured data, and to transform unstructured data into structured information that can be used for analysis and decision making in the healthcare and drug discovery areas.

Our tools are used from early drug discovery through to clinical trials, post-market work and regulatory affairs. Our U.S. healthcare customers use our tools to extract patient information from sources such as medical reports and nurses notes. Many of our pharma customers are also trying to get information from patents. These documents tend to be written to obscure rather than clearly communicate information. Being able to extract information on genes and diseases or complex information on who's innovating is quite valuable.

CAN YOU SHARE SOME REAL-LIFE EXAMPLES?
One of our biopharma customers said they brought two candidate drugs to late-stage clinical development in half the time they normally take because of our text mining tools. On the healthcare side, a user at Georgetown University Medical Centre was helping clinicians in the treatment of a child with three different conditions. The doctor wanted to quickly understand the relationship between the diseases and the impact on treatment. Our user was able to find out that there were only two journal articles published in the world on that topic. He gave that information to the doctor who could then make the best judgement for this child.

IS ACCESS TO CONTENT AN ISSUE FOR YOU?
In North America, the National Library of Medicine makes so much information openly available. We have a cloud offering that gives end-user access to those and other open content sources. For other content, we can't influence the publisher firewall but if you have access we have connectors and processing tools to handle different formats like PDFs, Word and HTML. We keep a close eye on discussions about Open Access and the pressures on big companies to loosen their firewalls. I think the publishing model will change in the coming years because of the pressure there is to make content more available in order to benefit everybody's research.

WHAT IS YOUR FUTURE VISION FOR TDM?
A lot of TDM tools demand technical expertise from their users but we look to provide solutions that empower people who are used to search engines to get the benefits from TDM and Natural Language Processing (NLP). We have to make interfaces for the next generation of workers fit with their use of mobile devices. That may seem far-fetched, taking something that's very complex and powerful and fitting that into the simplicity of a smartphone but I think that's what millennials will expect.

LEARN MORE
https://www.linguamatics.com

European Commission

www.openminted.eu

ATHENA Research & Innovation Information Technologies — AGRO-KNOW — EMBL-EBI — gesis Leibniz Institute for the Social Sciences — grnet Networking Research and Education — MANCHESTER 1824 The University of Manchester — TECHNISCHE UNIVERSITÄT DARMSTADT — UNIVERSITY of STIRLING

BSC Barcelona Supercomputing Center Centro Nacional de Supercomputación — EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE — frontiers — INRA SCIENCE & IMPACT — LIBER — The Open University — UNIVERSITY OF AMSTERDAM — University of Glasgow — CREATe