BIG DATA in Medical Affairs

“Text mining and Analytics”
Big Data as a source for Text Mining and Analytics

Text mining is a process of gaining essential and important information out of a text. This information can be derived and further enhanced by analyzing the patterns and trends. The term text analytics describes a set of linguistic, statistical and machine learning techniques that model and structure the information content of textual sources for business intelligence, exploratory data analysis, research, or investigation.

Most of data that life sciences industry generated were in unstructured formats more so text, growing rapidly in size and importance. This data with volume, velocity and variety over the time is taking a shape of Big Data. As text data are mostly semi-structured or unstructured, to bring out meaningful information out of such data, becomes a huge challenge for many companies. Use of analytical approaches and statistics like weighting, term frequency, document frequency, scaling, classification, clustering, segmentation and regression we can give some direction to unstructured text.

Text analytics an emerging technology is still being discovered everyday by organizations, so that meaning out of large volumes of unstructured text can be derived. It is being used in applications by numerous industries, which also include Life Sciences companies, who are mining text data sources to find answers to their subject of interest like efficacy issues, safety signals, discovering unknown patterns and causes or to identify promising drug candidates. With the emergence of web and social media, availability of information from various sources has increased rapidly which if mined properly can uncover valuable insights and probable solutions which were never thought of before.

Importance of Text Analytics

Vast amounts of new information and data generated everyday have significant potential economic and social value. Techniques such as text mining and analytics are being used to enhance this potential. Companies use such techniques to analyze industry, customer and competitor data in order to improve upon various key areas like profitability, market share and competitiveness among others.

Various ways in which text analytics is helping in finding greater insights are:

- Information Retrieval
- Text Classification
- Pattern Recognition
- Information Extraction
- Text mining
- Integration
- Visualization
- Prediction
- Entity recognition
- Finding trends, causes and patterns
- Sentiment Analysis

One area that can benefit from literature mining is the automatic identification of a large number of potential drug-drug interactions (DDI). Biomedical literature mining can aid by extracting relevant DDI
signals from either the published literature or from available large clinical databases. Pharmacogenomics is another area where text analysis can aid in knowing how genetic makeup affects an individual’s response to drugs. Text analysis can also come handy in bringing out relations between biomedical entities identified by Biological named entity recognition.

**Role of Medical Affairs**

Most of the Pharmaceutical companies have positioned their Medical Affairs groups in a leadership role for bringing high efficacy and safe products to market successfully. As a result, the group’s scope of responsibility has stretched beyond providing medical and scientific information to participating in all aspects of the drug process from discovery to post product launch.

In the past few years medical affairs vertical has matured and evolved to add new responsibilities. As per survey which was conducted by Cutting Edge Information, in recent time 35% of medical affairs groups have been adding the responsibilities of health economics and outcomes research (HEOR). Now even Drug companies are looking at providing more detailed patient outcome results/data to prove the value of their products. With raising cost of healthcare it is becoming invaluable for all the stakeholders.

Pharmaceutical companies are required to adhere to pharma-covigilance regulations and guidelines. Specific to detecting safety signals from the published literature, these guidelines state that a stipulated number of internationally recognized literatures from the databases should be searched at least monthly. It is also required that periodic updated reports on safety should have abstract and references from available literature, including relevant published summary from meetings or discussions containing important positive or negative findings on safety. The process requires that the published scientific and medical literatures must be assessed on a regular basis.

**Use of Big data in Medical Affairs**

In order to achieve the performance excellence that gives a company the competitive edge over others, effective use of Big Data is needed in today’s marketplace. Many Big Data sources like literature data base, interview content, web content, social media are key tools for Medical affairs and can help them achieve the following:

- Real world data / Evidence
- Valuable medical information
- Unique patient insights
- Spot drug safety signals
- Data management relationships
- Identify Key opinion leaders
- Scientific publication / reporting strategy
- Health economics and outcomes
- Relationships and interaction between entities e.g. care provider and patient
- Real time monitoring
- Risk Analysis
There is an increasing interest in text mining and information extraction strategies being applied to life sciences literature due to increasing number of electronically available publications stored in databases. Above Diagram 1 shows the use of literature databases and word indexes to search the data in order to get most relevant articles or publications pertaining to their scope of search. Since it is required that only authentic source of data should be used for justifying the efficacy or safety, social media is not considered as potential data source so far. Also most of the companies do not utilize advance analytical tools, as they restrict themselves to reporting only.

But the growing amount of information available in conversations, interviews or social media cannot be ignored for long. As such the same is included in Diagram 2. Many companies have already started taking note of the conversations happening about their product in professional social networks Facebook, Twitter and in Blogs. The explosion of use of social media is highly relevant to Medical Information. It’s not just about marketing and promotion campaigns, it’s about getting more insightful information about drugs, new diseases, chemical compositions, genes, trace & track a dialogue which is currently happening or a subject of interest and discussions. We need to be where the conversations about medicines are taking place.

In today's virtual world, social media conversations are having a direct impact on pharmaceutical sales and brand reputations, especially when conversations are off-label. Every day patients are questioning their treatment results and making product choices based on online conversations, many of which consist of un-edited and un-qualified user generated content.
Patients and their caregivers are able to interact with each other, as well as healthcare providers and Life Sciences companies from across the globe in real-time. They are doing so in unprecedented numbers. For many individuals now the first place they seek medical information is online. On the other hand, patients are also sharing valuable information about their thought processes and behavior.

As shown in Diagram 2, if Big Data sources are utilized correctly with the use of analytical tool and proper algorithms, the insights that we can get out of the results can potentially help all functions of drug development - from R&D to pharmacovigilance. Also it can bring in predictability in the system whereby companies will be more prepared.
Conclusion

With emerging social media and interactive medical blogging sites on the rise, it becomes imperative that Medical Affairs has to give importance to both using and monitoring this scientific exchange. Determine what may be expected on efficacy, what new developments can change the usage of your drug, and how you can do your part to prevent consumers receiving false information has become all the more important.

As the usage of BIG DATA becoming imperative with each passing time, medical affairs will also evolve and will use this opportunity to use greater source of data, so as to provide value based inputs to its stakeholders.

About Makrocare

Expert Strategic development and commercialization global partner for pharmaceutical, biotechnology and medical device industries. Our experience, programs and processes bring a new dimension to development strategy, regulatory/risk planning & management, clinical research, medical/scientific support and emerging region expansion.

MakroCare has operations and presence in US, UK, EU, and Asia. MakroCare is well positioned to meet the demands of the expanding product portfolio needs of our life sciences clients worldwide. We also understand the demands of an increasingly competitive market.
Resources:

Emerging Technology Series: Big Data Analytics in Health by Canada Health Infoway Inc. 2013

Medical Affairs Groups Take on New Roles by Laura Mortkowitz www.hcplive.com


Biological literature mining from information retrieval to biological discovery BY Lars Juhl Jensen EMBL, Germany

Text Analytics for Surveillance (TAS) An Interactive Environment for Safety Literature by Camilla Christensson, Ph.D. Geoffrey Gipson, PhD, Tracey Thomas, James Weatherall, PhD

Frontiers of biomedical text mining: current progress by Pierre Zweigenbaum

Biomedical literature mining for pharmacokinetics numerical parameter collection by Zhiping Wang

Text mining: From Wikipedia, the free encyclopedia

Evaluation of linear classifiers on articles containing pharmacokinetic evidence of Drug-Drug interactions by A. Kolchinsky, A. Lourenclo, L. Li and L. M. Rocha

Big data revolution in healthcare, accelerating value and innovation by McKinsey & Company.

Big data: The next frontier for innovation, competition and productivity by McKinsey & Company.

What's Really Different about Big Data Analytics? By Fern Halper

In 2013, big data will drive big pharma, By Zack King / Co-founder, Trinity Pharma Solutions


Big Data in the Pharmaceuticals Industry, Originating Author: Jeff Kelly


Hosted big data demands big performance, SoftLayer

How big data can revolutionize pharmaceutical R&D, by Jamie Cattell, Sastry Chilukuri, and Michael Levy

Pharma's Great Hope: Big Data, McKinsey & Company
USA

New Jersey (HQ)
California
Illinois
Pennsylvania

EUROPE

Basel (Switzerland)
Frankfurt (Germany)
London (UK)
Warsaw (Poland)

ASIA

Hyderabad, Mumbai (INDIA)
SINGAPORE
Beijing (CHINA)
Tokyo (JAPAN)

Disclaimer

This publication only contains general information and none of MakroCare Clinical Research ltd, its member firms, or their related entities etc. (collectively, the "Makro Group") is, by means of this publication, rendering any professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you are requested to consult a qualified professional adviser or consultant. No entity in the MakroCare company / Makro group shall be responsible for any loss whatsoever sustained by any person/company who/which relies on this publication.

© 2013 MakroCare