Why Text Mining will Revolutionize Big Data in Biotech

Currently, searching information within 'big data' is performed via key-words, often resulting in a high number of results that are not always relevant to your initial question. Now, imagine a tool that gives a more precise answer to your question with the same ease of use. That's the power of text and data mining, and it's coming into Biotech.

Since the development of the Internet, search tools have vastly improved, giving you powerful solutions to find information. However, often there are limitations in their results. When you're looking for very specific answers, you simply obtain an extensive list of sources in which you have to dig around to find the relevant answer to your question.

Text and data mining tools use a specific algorithm capable of sifting through huge amounts of data to find the specific answer to an open question. Data mining tools use raw data sets, text mining tools use scientific literature. The text mining algorithm basically machine reads all articles full-text that are relevant to a topic, finds the right information and then provides you with just what you need, with the help of a thesaurus.

Let's take an example in the Biotech field. Imagine that you're looking for the molecules that can bind to a **receptor of interest**. With text mining tools you can obtain a list of relevant molecules mentioning the receptor, rather than a list of scientific papers that you would get with a 'normal' search engine that you then need to browse through, saving a huge amount of your research time. Text mining can be so effective because it can be set up to search through the full-text of articles, not just the abstract

Let's take another example; you want to know if a molecule you are investigating plays a role in treating a known disease. In the literature, an article shows that this molecule can bind to a specific receptor. Another article, published at a different period by a different team, highlights the fact that this same receptor is involved in this disease. The power of text mining is that the algorithm is able to combine these two unrelated articles and give you new insights. Uncovering these hidden gems in scientific literature may save you a lot of time doing literature study and bench-work.

With text-mining, finding the needle in a haystack becomes easy…

Here are three good reasons to consider text mining in your research project:

- **It enhances R&D efficiency**: a text mining project can be performed over potentially huge amounts of scientific articles that would take years for someone to read. By shortening the literature research process, R&D teams can focus on discovery, innovation and so accelerate delivery of results.

- **It increases discovery**: where standard keyword searches only scan the surface of documents, text mining reveals hidden relationships to help researchers identify and develop new hypothesis, attain knowledge and improve   understanding.

- **It monitors drug safety**: with text mining you can recognize potential adverse effects of a particular component. It will help you to assess the safety issues much earlier than if you had to investigate it yourself in a preclinical trial. This will therefore avoid unnecessary cost of development and so will save vast amounts of time and research budget.

As has been described above, text mining is a tool that will raise the quality of searching in scientific information and finding relevant insights. While classical literature research with keywords gives you a list of documents, text mining analyses the results for you and gives a more direct answer to a question. The potential for the Biotech industry is substantial in terms of finance and efficiency in research. Finally, with the development of text mining, research teams will be able to develop new drugs more quickly and ultimately may be able to develop benefits to the patient faster.

About Labiotech.eu:

Labiotech.eu is the leading digital media covering the European Biotech industry. It is used by over 45 000 people monthly to keep a watch on the business and innovations of biotechnologies. Join them for free and enjoy reading our stories!

To register to the newsletter: http://labiotech.eu/newsletter/