

InSites Consulting helps a major pharmaceutical company understand consumer behaviour



Analysing patterns in web media with IBM SPSS Modeler

Smart is...

Trawling the web to answer questions other market research methods cannot ask

As an online full-service market research agency, InSites Consulting seeks to help its clients understand the needs, interests and desires of their consumers. In one recent project, a major pharmaceutical company hired InSites Consulting to analyse the way consumers think and communicate about epilepsy.

InSites Consulting uses IBM SPSS Modeler to analyse text from thousands of blogs, discussion forums and other internet sites and discover which aspects of a given topic are generating the most 'buzz' amongst consumers. The company has created a custom 'dictionary of emotions' which rates each comment to assess positive and negative responses to these topics.

InSites Consulting is an online market research and marketing consultancy. Founded in Belgium in 1997 as a cutting-edge internet research company, it has expanded to become a full-service agency that employs more than 100 people at offices in Ghent, Rotterdam, London and Geneva, and now operates in 37 countries. It generates revenues of approximately €13.5 million per year.

InSites Consulting specialises in online market research, performing surveys, online research communities and (n)ethnographic studies for clients in all industries. As one of the first companies to grasp the potential of the internet for market research, InSites Consulting has maintained its position as an innovator by constantly developing and refining its methods and analytical techniques.

Leveraging internet resources

"The Internet is an incredible resource for market research because it contains sites where hundreds of thousands of consumers express their views on every aspect of modern society," explains Niels Schillewaert, Managing Partner of InSites Consulting. "Traditional market research methods, such as constructing questionnaires or organising focus groups, require researchers to make certain initial assumptions about what aspects of a given topic are most important or interesting. But we realised that if we could take a 'bottom-up' analytical approach and use the internet to find out what people are actually talking about, we could help our clients build a much more resonant dialogue with their customers.

"The challenges, though, are the sheer volume of data available, the need to handle multiple languages, and the fact that the data exists mostly as unstructured text. To analyse it effectively, we need a highly sophisticated text mining engine – and for this, we use IBM SPSS Modeler. It's one of the most mature solution in terms of its dictionaries and analytical capabilities."



Business Benefits

- Helps to discover issues that people really care about, helping to identify gaps in the market for new products and offerings.
 - Reveals the language people use to describe these issues, helping companies to communicate using more consumer-oriented language.
 - Analyses huge data samples from a global consumer base, in complement to gathering information using traditional market research methodologies.
-

Collecting and analysing data

InSites Consulting uses a web-crawler/scrapper to collect data from thousands of web pages across the internet – blogs, discussion forums, social media, and other sites that allow users to leave comments. The text collected is then cleaned and de-duplicated and stored in a Microsoft SQL Server database. It can then be loaded into IBM SPSS Modeler for analysis.

“The best way to explain how it works is probably via an example of a real-world analysis project,” says Schillewaert. “We recently worked with a pharmaceutical company which wanted to investigate the way people talk about epilepsy. So we collected data from a variety of sites with medical discussions and used SPSS Modeler to perform a bottom-up analysis to identify patterns and find out which epilepsy-related topics were being discussed most frequently. We then used these findings to perform a top-down analysis, looking for relationships between topics, and then conducted a deep-dive analysis per topic.”

Finding the ‘golden nuggets’

These analytical processes yielded a number of ‘golden nuggets’: topics that the client had not previously considered important, but which generated a high level of ‘buzz’ (or activity) in social media. For example, it became clear that dealing with epileptic seizures at night was a key concern, as was the relationship between epilepsy and pregnancy.

“The next step is to ‘drill down’ into the selected topics to provide additional insight,” says Schillewaert. “We have developed a customised ‘dictionary of emotions’ in SPSS Modeler, based on Ekman’s theory of emotions. This allows us to rate each comment for positive or negative emotions, and even identify what the primary sentiment is: surprise, anger, fear, and so on. Topics with a lot of buzz and negative sentiment tend to point to frustrations, which can help our clients understand customer dissatisfaction, or even identify gaps in the market where new product or service offerings can be developed.”

Business Analytics and Optimization: Finding the ‘hot topics’ for consumers



Instrumented

People’s opinions (in the form of unstructured text data) are automatically ‘scraped’ and collected from hundreds of blogs, discussion forums and other websites across the internet.



Interconnected

The solution performs sophisticated ‘bottom-up’ and ‘top-down’ analyses on the collected data, identifying patterns and comparing language against a ‘dictionary of emotions’.



Intelligent

Understanding consumers’ real and spontaneous concerns, rather than what companies assume they are interested in, enables the formulation of more effective marketing and product development strategies.

Solution Components

Software

- IBM® SPSS Modeler
-

“The sophisticated text mining capabilities of IBM SPSS Modeler made a decisive contribution to our ability to deliver value to our client... Ultimately, we would like to extend our use of its data mining capabilities too, giving us a complete solution for web analytics.”

— Niels Schillewaert, Managing Partner of InSites Consulting

For example, in the epilepsy study, one of the key talking points was birth control. In many cases, female epilepsy sufferers were concerned that their epilepsy medication would reduce the efficacy of their birth control medication – giving them the dilemma of choosing between the risk of seizures and the risk of pregnancy. These findings identified a key point of need in the general population that the client was able to feed back to its marketing and product development teams.

Better communications

The solution can also help clients improve their communications strategy through linguistic analysis, by revealing the vocabulary consumers actually use to talk about the issues that concern them. In the epilepsy example, InSites Consulting helped its client understand that instead of technical terms like ‘petit mal’ and ‘grand mal’, sufferers tend to use more descriptive, subjective expressions to describe their seizures. This feedback is helping the client revise its communications strategy to build a closer, more natural dialogue with its customers.

Schillewaert concludes: “The sophisticated text mining capabilities of IBM SPSS Modeler made a decisive contribution to our ability to deliver value to our client in the epilepsy research project, and we have used it in similar projects in many other industries – from television through consumer packaged goods to energy and utilities. Ultimately, we would like to extend our use of its data mining capabilities too, giving us a complete solution for web analytics. The major challenges to realise our future goals are related to the multi-linguistic capabilities and the ability to process enormous amounts of data.”

► The Inside Story: Getting There

InSites Consulting had been using SPSS Statistics for data analysis, and decided it needed a tool for text mining too. Following a detailed review of the products available on the market at the time, the company realised that IBM SPSS Modeler was the most mature and flexible option.

“The installation of the software itself was straightforward, just like any software installation; the main challenges are user training and running the analysis on really huge data files and specific languages,” explains Schillewaert. “InSites Consulting really pushes the IBM SPSS Modeler software to its limits and the support offered by the IBM SPSS Netherlands and Belgium team is helpful. At the start we acquired a three-day intensive training course, and then rolled the software out to an internal team of ten people. We got our team up to speed quickly, and since then, SPSS Modeler has been a key tool for our ethnography studies.”

For more information

To learn more about IBM SPSS software, contact your IBM sales representative or visit ibm.com/spss/

To learn more about InSites Consulting, visit insites.eu



© Copyright IBM Corporation 2011

IBM Belgium / Luxembourg
Avenue du Bourget / Bourgetlaan 42
1130 Brussels
Belgium

Produced in Belgium
January 2011
All Rights Reserved

IBM, the IBM logo, and ibm.com are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. A current list of other IBM trademarks is available on the Web at “Copyright and trademark information” at: ibm.com/legal/copytrade.shtml.

Other company, product or service names may be trademarks, or service marks of others.

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program or service is not intended to imply that only IBM's product, program or service may be used. Any functionally equivalent product, program or service may be used instead.

All customer examples cited represent how some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

IBM hardware products are manufactured from new parts, or new and used parts. In some cases, the hardware product may not be new and may have been previously installed. Regardless, IBM warranty terms apply.

This publication is for general guidance only.

Photographs may show design models.



Please Recycle
