



The Centers for Disease Control and Prevention predicts spread of the flu with text analytics

The challenge

Part of the United States Department of Health and Human Services, the Centers for Disease Control and Prevention (CDC) serves the mission of protecting public health. As pandemics emerge, the CDC relies on predictive analytics models to measure and track the spread of diseases to understand their evolution and impact.

The CDC builds sophisticated models based on past quantitative data gathered from doctors, emergency rooms, and urgent care centers. Although reliable, these models are reactive and outdated by the time data is analyzed. To be truly predictive, the CDC needed the ability to understand the spread and severity of a wide range of illnesses, even those that were not yet known. During a recent flu season, the Situational Awareness Branch of the CDC looked to prove the value of analyzing real-time text data to create more accurate prediction models.

An effective solution would:



Examine public discussion to monitor and predict the spread of the flu



Analyze real-time text against historical qualitative data



Perform advanced analyses on incoming conversational text

The solution

With Luminoso, the CDC integrated its historical quantitative data with real-time qualitative data, including text from social media. The results showed that present flu references aligned with past reports from doctors and hospitals, enabling the team to deeply examine a more complete picture of the disease as it was evolving.

The CDC could now identify new flu cases from social media, even when a person did not explicitly mention – or know – what they were suffering from. For example, posts mentioning “shopping for Nyquil” were more likely to be flu cases than those mentioning “home sick from work”. And because Luminoso also analyzes emoji, the team discovered the pill emoji strongly indicated the flu, even if a post provided no other information – or actual text.

The benefits



Improved clarity and accuracy of the flu spread prediction model



Categorized and processed over 8,000 social posts per minute



Correlated real-time text with historical numerical data

The CDC could now identify new flu cases from social media, even when a person did not explicitly mention or know what they were suffering from.

The results

With Luminoso's text analytics, the CDC's disease models evolved from reactive to predictive, helping the organization:



Identify words and phrases distinguishing flu cases from other diseases



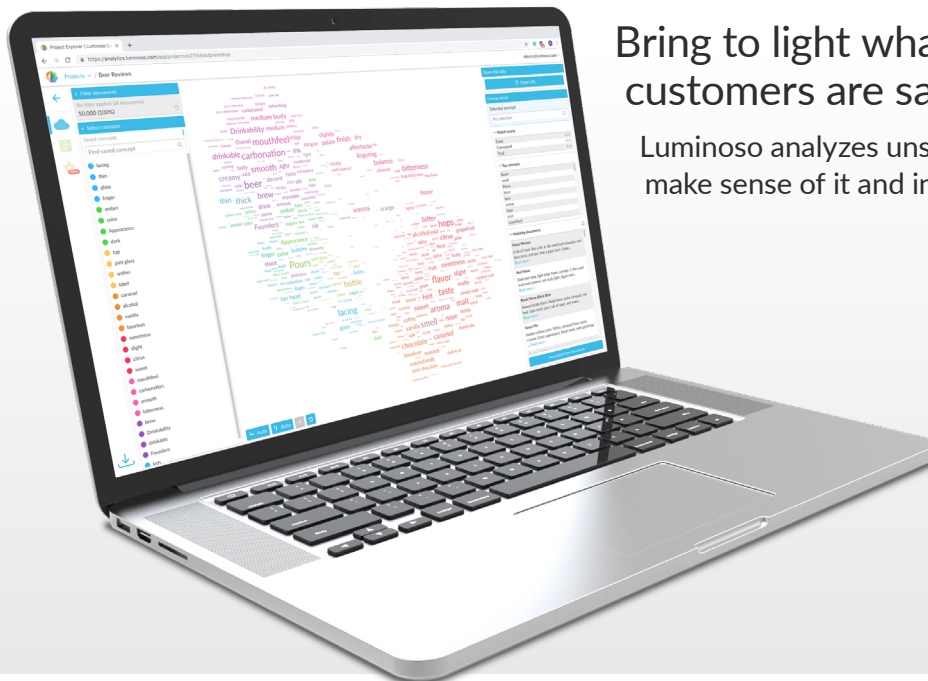
Understand the severity, duration, and virulence of the current flu strain



Determine flu vaccine and treatment efficacy



Track the number and spread of new flu cases



Bring to light what your customers are saying

Luminoso analyzes unstructured text in minutes, to accurately make sense of it and inspire your organization's actions.

Luminoso turns unstructured text data into business-critical insights. Using common-sense artificial intelligence to understand language, we empower organizations to discover, interpret, and act on what people are telling them. Requiring little setup, maintenance, training, or data input, Luminoso combines world-leading natural language understanding technology with a vast knowledge base to learn words from context – like humans do – and accurately analyze text in minutes, not months. Our software provides native support in over a dozen languages, so leaders can explore relationships in data, make sense of feedback, and triage inquiries to drive value, fast. Luminoso is privately held and headquartered in Boston, MA.

©2019 Luminoso Technologies. Luminoso is a registered trademark of Luminoso Technologies.



www.luminoso.com

info@luminoso.com

[@LuminosoInsight](https://twitter.com/LuminosoInsight)