Text Mining in the news

German Rigau i Claramunt

german.rigau@ehu.es

IXA group

Departamento de Lenguajes y Sistemas Informáticos

UPV/EHU

nature
International weekly journal of science

comments on this story

News

News mining might have predicted Arab Spring

Signs of impending social and political change may lie hidden in a sea of news reports.

Philip Ball

You could have foreseen the Arab Spring if only you'd been paying enough attention to the news. That's the claim of a new study that shows how data mining of news reportage can reveal the possibility of future crises well before they happen.

Technology

Supercomputer predicts revolution

9 September 2011 | Technology

Feeding a supercomputer with news stories could help predict major world events, according to US research.

A study, based on millions of articles, charted deteriorating national sentiment ahead of the recent revolutions in Libya and Egypt.

While the analysis was carried out retrospectively, scientists say the same processes could be used to anticipate upcoming conflict.

Sentiment mining showed a sharp change in tone around Egypt ahead of President Mubarak's ousting

The system also picked up early clues about Osama Bin Laden's location.
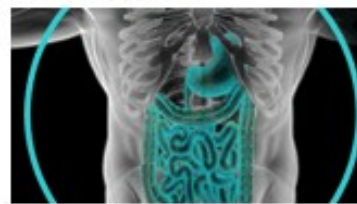
The Scientist » News & Opinion » Opinion

Opinion: Text Mining in the Clinic

Despite increasing use of electronic medical records, much patient data remains in text form, requiring text-mining techniques to make full use of patient information.

By Min Song | April 1, 2013

☉·PLOS | COMPUTATIONAL BIOLOGY

Browse   Publish   About

RESEARCH ARTICLE

Using Electronic Patient Records to Discover Disease Correlations and Stratify Patient Cohorts

Francisco S. Roque 🆑, Peter B. Jensen 🆑, Henriette Schmock, Marlene Dalgaard, Massimo Andreatta, Thomas Hansen, Karen Søeby, Søren Bredkjær, Anders Juul, Thomas Werge, Lars J. Jensen, Søren Brunak ✉

Article	Authors	Metrics	Comments	Related Content

Abstract

Author Summary

Introduction

Results

Discussion

Materials and Methods

Abstract

Electronic patient records remain a rather unexplored, but potentially rich data source for discovering correlations between diseases. We describe a general approach for gathering phenotypic descriptions of patients from medical records in a systematic and non-cohort dependent manner. By extracting phenotype information from the free-text in such records we demonstrate that we can extend the information contained in the structured record data, and

Blog

Archivo de categorías: Voz del Cliente (VoC)

Entradas sobre Voz del Cliente (VoC)

¿Pueden los antidepresivos causar malformaciones en el feto?

📅 28/julio/2015

🗂 API   Medios sociales   Procesamiento semántico   Recuperación de Información   Sector Salud   Voz del Cliente (VoC)

🏷 análisis de texto   embarazo   fármacos   medios sociales   monitorización   salud

Puede que no sea lo más común que en el *blog* de una empresa de tecnología de la información se hable de antidepresivos y embarazos, lo entendemos perfectamente. Pero en MeaningCloud hemos comprobado que los temas de salud tienen una fuerte repercusión en medios sociales y las empresas del sector, incluyendo las farmacéuticas, harían bien en tratar de entender la conversación que se genera alrededor. ¿Cómo? Mediante la tecnología de análisis de texto, como veremos a continuación.

Mirando los datos recogidos por nuestro prototipo para la monitorización de temas de salud en medios sociales nos sorprendimos del repentino incremento de menciones del término 'embarazo' el día 10 de julio. Para saber el porqué de este dato, acudimos a los tuits que hablaban sobre ese estado de *buena esperanza*. Resulta que ese mismo día se había publicado

Reuters   :: ✉ Enviar :: 🖨 Imprimir

Estudio relaciona uso de Prozac y Paxil en embarazo con malformaciones congénitas

Un estudio en miles de mujeres halló un vínculo entre los antidepresivos Prozac y Paxil y una serie de defectos congénitos.

Por Ellen Wulfhorst. Reuters | 13 de julio de 2015

Un estudio en miles de mujeres halló vínculos entre los antidepresivos Prozac y Paxil y una serie de defectos congénitos, al tiempo que descartó la relación en el caso de otros tratamientos populares como Celexa, Lexapro y Zoloft.

El medicamento Zoloft de Pfizer está siendo objeto de un juicio por demandas que lo

What do you think of our new site? Give us your feedback

0

DAILY NEWS 6 March 2013

Web searches uncover hidden drug dangers

You're feeling unwell, and wonder if the problem might be a side effect of the drugs your doctor just prescribed. So you turn to a search engine and type in your symptoms and the names of the drugs.

It happens every day, and now researchers at Stanford University in California have shown that search logs from millions of web users can be used to detect dangerous interactions between drugs – which are notoriously difficult to predict and monitor.

Sentiment Analysis Forecasts "No" Vote Will Prevail in Scottish Independence Referendum

September 18, 2014   Written by: Eileen McNulty   0 Comments

Category: Data Science, Featured, Front, News

permalink

Tagged under:

Regardless of the outcome of today's Scottish Independence Referendum, today is a landmark event in Great British history. Within the next 24 hours, Scotland will vote to decide whether or not to dissolve their 307-year-old union with England. Opinion polls have suggested the vote is too close to call; different polls have disagreed about whether or the "Yes" camp or the "No" camp have the edge, and either way the results have fallen within the margin of error.

Olery Review Data API now includes Sentiment & Traveler Data

July 31, 2015 — by Roman
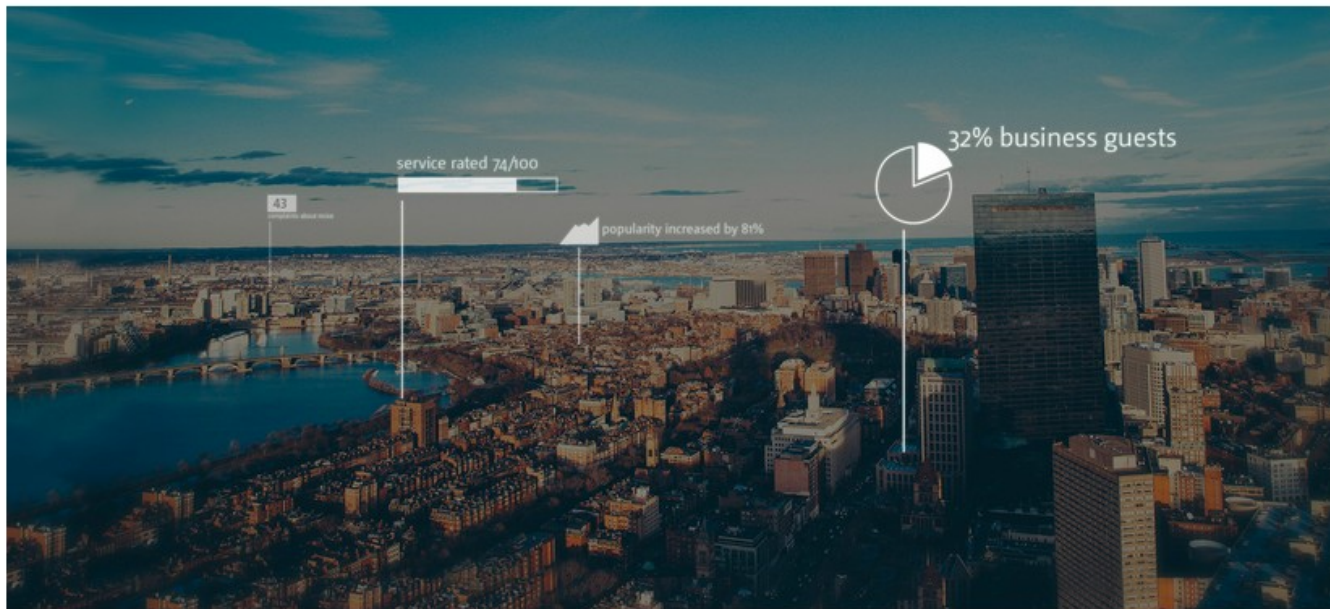
Last October the Olery Hotel Review Data API was announced. Since then we have experienced a lot of interest in structured, real-time performance data on hotels worldwide. Now the time has come to give the API the first major overhaul and to introduce a couple of exciting new features. Let's dive into it.

Artificial Solutions' Dr. Rebecca Jonsson to Speak at SpeechTEK

June 11, 2015 / admin / Events, Indigo, Natural Language Interaction, Teneo Mobile, Teneo Platform, Teneo Virtual Assistant / 1 comment

Artificial Solutions™ (www.artificial-solutions.com), the natural language interaction (NLI) specialist that enables users to have a

Why recruitment software has moved on from CV parsing

Posted on Fri, Oct 14, 2011

Tweet 4   in Share 0   f Like   Share 3   g+1 0

Recruiters are seeing more and more reason to move forward into the digital age when it comes to attracting and hiring the most qualified candidates online.

What once was an arduous task to review individual resumes by hand was quickly relieved by the advent of CV parsing software that greatly reduced the workload expected of many recruitment firms. CV parsing became the best thing since sliced bread when it came to going through and storing hundreds of resumes for the various positions being recruited for.

So what's the problem with CV parsing?

Although it was revolutionary, CV parsing has always had its limitations. Even with the newest and brightest parsing software programs there are still disadvantages to reviewing resumes in this manner. Some of the software was only able to search for keywords, some had the ability to search for grammatical

onrec

yu:talent

SmartRecruiters chooses Textkernel's resume parsing software

📄 Posted in News archive on 13th Jun 2014

➕ share   ✉ receive our newsletter

Jerome Ternynck, CEO of SmartRecruiters talks about his company and his decision to integrate Textkernel's resume parsing technology into its recruitment platform.

Who's SmartRecruiters?

SmartRecruiters is the cloud based platform to make hiring easy for managers and candidates. SmartRecruiters has helped find 12,000,000+ candidates for 65,000+ businesses.

30

Sept 2014

How Big Data is impacting the job-hunting and hiring experience today

'People analytics' is allowing companies to get the right candidate discovered and in the door

By Javid Muhammedali, Vice President of Product Management, Monster

Are you tired of hearing about 'Big Data' yet? If you're actively looking or open to a new job, learning more about it and what it might mean to you could get you to that next leap in your career.

In general, Big Data includes reams of information generated by organizations, the government, and individuals. Big Data is characterized by high volumes of unstructured data coming from several different sources at the same time. The information itself has existed for years (think of all the application forms you've filled out in your lifetime, then add all the tweets, web profiles, open source contributions and online discussion threads you've created), but the tools to efficiently sift through all that data have only recently become available.

Increasingly, companies are finding tools to organize and analyze these ever-growing pools of data. And it's not just sales and financial data being analyzed. Smart companies now use Big Data to make employment decisions, (sometimes referred to as "people analytics" or "Talent Analytics") — and it is already impacting your job search and employment opportunities.

Business

Does job success depend on data rather than your CV?

By Matthew Wall
Business reporter, BBC News

2 October 2014 | Business

EL PAÍS

ECONOMÍA

Jobandtalent levanta 23 millones de euros de un plumazo

- Pelayo Cortina Koplowitz, la Fundación Entrecanales y Nicolás Luca de Tena invierten junto a dos fondos en el portal de empleo español por segunda vez en menos de un año

CARMEN SÁNCHEZ-SILVA | Madrid | 27 MAY 2015 - 09:02 CEST

| f | 1.040 | | 74 | in | 465 | g^+ | 10 | |

Archivado en: Jobandtalent   Start-up   Tecnologías información   Internet   Empresas
Informática   Economía   Industria   Telecomunicaciones   Comunicaciones   Tecnología

El Confidencial

Julio Prada, director general en Europa, Oriente Medio y África de Inbenta

EFE   20/07/2015 (13:30)

A | A

Barcelona, 20 jul (EFECOM).- Inbenta, compañía especializada en procesamiento del lenguaje natural y en búsqueda semántica, ha nombrado a Julio Prada, hasta ahora director general de la firma en España, nuevo director general en las zonas de Europa, Oriente Medio y África.

La empresa catalana ha explicado en un comunicado que su nombramiento responde al rápido crecimiento que Inbenta ha experimentado en los últimos años y al objetivo de afianzar la posición de liderazgo mundial.

De esta manera, Julio Prada, que ha impulsado y consolidado el papel de la compañía catalana en España desde 2007, es el elegido para expandir el negocio en Europa, Oriente Medio y África.

Julio Prada es licenciado en Informática por la Universidad Autónoma de

0

0

0

0

Menéalo

Enviar

Imprimir

1

Blog

El papel de la Minería de Texto en el Sector de Seguros

📅 02/octubre/2014

📁 Análisis de Fraude   Análisis de Sentimiento   Sector Seguros   Tecnologías del Lenguaje   Voz del Cliente (VoC)

🏷 experiencia del cliente

¿Qué pueden hacer las aseguradoras para explotar toda su información no estructurada?

Un escenario típico de big data

Las compañías de seguros recogen ingentes volúmenes de texto diariamente, a través de múltiples canales (sus agentes, centros de atención a clientes, correo electrónico, redes sociales, web en general). Las informaciones recogidas incluyen pólizas, informes periciales y de salud, reclamaciones y quejas, resultados de encuestas, interacciones relevantes de clientes y no-clientes en redes sociales, etc. Es imposible atender, clasificar, interpretar o extraer la información esencial de todo ese material.

El Sector de Seguros es uno de los que más puede beneficiarse de la aplicación de las tecnologías para análisis inteligente de texto en formato libre (conocido como Analítica de Textos, Minería de Textos o Procesamiento de Lenguaje Natural).

Technology » August 2011 » Advances

💬 2 :: ✉ Email :: 🖨 Print

Parsing the Twitterverse: New Algorithms Analyze Tweets

Smarter language processors are helping experts analyze millions of short-text messages from across the Internet

By Francie Diep | Jul 13, 2011

Researchers have been trolling Twitter for insights into the human condition since shortly after the site launched in 2006. In aggregate, the service provides a vast database of what people are doing, thinking and feeling. But the research tools at scientists' disposal are highly imperfect. Keyword searches, for example, return many hits but offer a poor sense of overall trends.
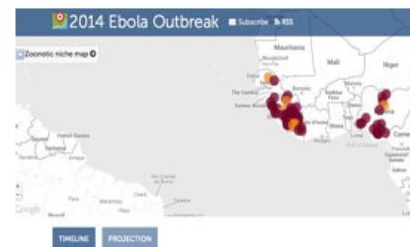
Technology » News

:: ✉ Email :: 🖨 Print

This article is from the In-Depth Report Ebola: What You Need to Know

Smart Machines Join Humans in Tracking Africa Ebola Outbreak

Web sites devour piles of disease-related information but still keep a person in the loop to help make sense of the data glut

By Larry Greenemeier | September 24, 2014

Health care officials and aid workers attempting to trace the progression of the Ebola virus disease outbreak that has claimed more than 2,800 lives so far (pdf) have come to rely heavily on a handful of disease-monitoring Web sites that act as pivotal hubs for processing information.

2014 Ebola Outbreak

Brown&
Meyers

Request a free transcription quote!

Get a Quote Now »

or call (800) 785-7505

Home   Transcription Services ⌄   Business Services ⌄   Testimonials   About ⌄   Blog   Contact Us

Home / Blog / Insurance Companies Use NLP Technology to Analyze Text and Reduce Fraud

Insurance Companies Use NLP Technology to Analyze Text and Reduce Fraud

⟨

⊘ March 19, 0204   🗀 General   💬 Leave a comment   👤 Kate Meyers

Identifying fraud and delivering outstanding, timely customer service to deserving claimants are long-standing challenges for insurance companies. How can insurance companies better use available data to service their quality clients?
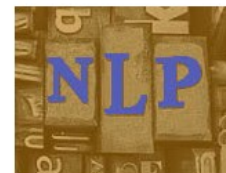
Insurance companies collect text-based data, in many languages and dialects, from an array of sources: applications, claims, email, social media, adjuster notes, medical records, police statements, surveys, research studies, underwriter notes, competitive intelligence and more. This vast quantity of unstructured and often conversational data piles up. While individual

Recent Blog

- 4 Easy Steps to Record Meetings for Transcription
- Scanning Legal Files for Practice Management and Remote Access
- Stop the Pain! There's a Better Way to Transcribe Legal Recordings
- Natural Language Processing (NLP) helps healthcare organizations mine transcribed data to enhance patient care, improve profitability, and comply with HIPAA

Connect with Us

Member Sign In ⌄
For Journalists ›
For Bloggers ›
Global Sites ⌄

Solutions | Knowledge Center | Blog | Browse News Releases | Contact PR New

See more news releases in Computer Electronics | Publishing & Information Services | Surveys, Polls and Research

North America Text Analytics (By Solutions(Competitive Intelligence, Fraud Det Software, Natural Language Processing, Enterprise Feedback Management, Keyv Linguistic Analysis))-Market Size And Forecast (2014-2019)

Share +1   in Pin it

LONDON, July 13, 2015 /PRNewswire/ -- The North Americantext analytics market is estimated to grow at a CAGR of 19.3% from 2014 to 2019. Although the U.S. holds the larger shares in the North Americantext analytics market, Canada is expected to grow considerably. The growth of text analytics market in this region is driven by technological advancement and the urge of getting useful data from the big data available on web. However, text analytics is time consuming when it comes to search for a particular information specific to a particular word. Moreover, emerging economies such as the U.S. and Canada are creating growth opportunities for the North Americantext analytics solutions in this region.

The North Americantext analytics market is segmented and forecast on the basis of submarkets such as data analysis applications, enterprise applications, search-based applications, web-based and software applications, and others. The North Americantext analytics market is also segmented on the basis of major countries in this region such as the U.S., Canada, and Mexico. Factors such as increase in usage of social media to analyze the consumer behavior by the organization to design the strategies accordingly and the rapid growth in the tech-savvy population are driving the growth of the text analytics market in SMB's in this region.

SECURITY**WEEK**

INTERNET AND ENTERPRISE SECURITY NEWS, INSIGHTS & ANALYSIS

Subscribe (Free) | **Security White Papers**

FREE WHITE PAPER

IBM

Fundamentals of Application Release and Deployment

Learn how automating application release and deployment can speed time to market, driv costs and reduce risk.

Malware & Threats   **Cybercrime**   **Mobile & Wireless**   **Risk & Compliance**   **Security Architecture**   **Managem**

Home > Virus & Threats

OpenDNS Uses Natural Language Processing to Detect APTs

By Eduard Kovacs on March 05, 2015

in Share ‹ 12   g+1 ‹ 5   Tweet ‹9 f Recomana-ho ‹2 RSS

OpenDNS has unveiled NLPRank, a new model that can be used to detect both opportunistic phishing campaigns and advanced persistent threats (APTs) by identifying certain patterns in DNS traffic.

OpenDNS security researcher Jeremiah O'Connor has found a way to combine natural language processing (NLP) techniques with the company's global network data to detect malicious activity. The algorithms used by **NLPRank** are usually seen in fields such as data mining and bioinformatics, but the researcher has demonstrated that they can also be useful

com/opendns-uses-natural-language-processing-detect-apts#

Using Natural Language Processing to Identify Malicious Domains

BY JAIKUMAR VIJAYAN • MARCH 16, 2015

f Share | 0 | 8+1 | 0 | 🐦 Tweet | 0 | in Share | 8 | f Like | 0 | ➕ Share | 22

Cybercriminals apparently have a tendency to use the same (or at least similar) lexical styles when establishing domains for phishing and advanced persistent threat (APT) attacks, making it possible for security researchers to identify sites using natural language processing (NLP) techniques.

That's according to OpenDNS Security Labs, which is prototyping a tool dubbed NLPRank to see if it can identify potentially malicious websites and phishing domains more quickly. Based on tests so far, the natural language processing tool could prove to be a "robust" method for defending against APTs, claimed OpenDNS security researcher

Jaikumar Vijayan

Independent Journalist & Tech Content Creation Specialist

🐦 Follow @jaivijayan

Jaikumar Vijayan is an award-winning technology journalist with nearly 25 years of experience in IT

bank innovation

Home   Content »   Lists »   Bank Innovation Israel '15   White Papers   Jobs

Beyond Big Data: Big Text Analytics

JULY 9, 2015
BY STEVEN RAMIREZ

Big Data is valuable for banks because it can identify trends, uncover emerging issues, and forecast business outcomes. Yet leading companies are thinking beyond Big Data, and using Big Text Analytics to create a more comprehensive, intelligent customer analysis. Businesses can enrich Big Data with Big Text to develop a better understanding of customers, and as a result, create better business outcomes.

What is Big Text Analytics?

Big Text is unstructured word-based data, and includes:

- Email and phone messages
- Online news and blogs
- Social media posts

Is "Big Text" next big thing?

Text analytics can reveal even more about customers

07/18/2015 - 18:00 | Written by John Ginovsky | Comments: 0 Comments

in y f g+ + 28

Move over, Big Data. Here comes Big Text.

Or, as it's more commonly known, "text analytics."

Or, sometimes, "text mining."

CCBS Handbook No. 33 - Text mining for central banks

29 June 2015

CCBS Handbook No. 33
by David Bholat, Stephen Hansen, Pedro Santos and Cheryl Schonhardt-Bailey

Although often applied in other social sciences, text mining has been less frequently used in economics and in policy circles, particularly inside central banks. This Handbook is a brief introduction to the field, discussing how text mining is useful for addressing research topics of interest to central banks, and providing a step-by-step primer on how to mine text, including an overview of unsupervised and supervised techniques.

📄 **CCBS Handbook No. 33 - Text mining for central banks** (2.3MB)

BANK OF ENGLAND

Search
Advanced Search | Search Help
► Contact Us
► Print this page
► Subscribe to alerts
► Feedback

Share

Latest!

- Detecting Risks in the Banking System by Sentiment Analysis Clemens Nopp and Allan Hanbury, 2015, EMNLP

- In November 2014, the European Central Bank (ECB) started to directly supervise the largest banks in the Eurozone via the Single Supervisory Mechanism (SSM). While supervisory risk assessments are usually based on quantitative data and surveys, this work explores whether sentiment analysis is capable of measuring a bank's attitude and opinions towards risk by analyzing text data. For realizing this study, a collection consisting of more than 500 CEO letters and outlook sections extracted from bank annual reports is built up. Based on these data, two distinct experiments are conducted. The evaluations find promising opportunities, but also limitations for risk sentiment analysis in banking supervision. At the level of individual banks, predictions are relatively inaccurate. In contrast, the analysis of aggregated figures revealed strong and significant correlations between uncertainty or negativity in textual disclosures and the quantitative risk indicator's future evolution. Risk sentiment analysis should therefore rather be used for macroprudential analyses than for assessments of individual banks.

Text Mining in the news

German Rigau i Claramunt

german.rigau@ehu.es

IXA group

Departamento de Lenguajes y Sistemas Informáticos

UPV/EHU