What is text analytics and why should I care?
An overview and analysis of the current market for text analytics software
Text analytics and AI—Moving into the practical realm

New techniques and technologies in text analytics from graph databases to deep learning
Case studies of multiple applications and the business value of text analytics
Integrating machine learning, taxonomy, and search

The use of text analytics in dealing with fake news and bad ad placement
Latest ideas in auto-categorization and data extraction from text
Machine learning vs. rule-based categorization

What is text analytics and why should I care?
An overview and analysis of the current market for text analytics software
Text analytics and AI—Moving into the practical realm

New techniques and technologies in text analytics from graph databases to deep learning
Case studies of multiple applications and the business value of text analytics
Integrating machine learning, taxonomy, and search

The use of text analytics in dealing with fake news and bad ad placement
Latest ideas in auto-categorization and data extraction from text
Machine learning vs. rule-based categorization

Go Deeper: Build Smarter Solutions
November 8–9, 2017
JW MARRIOTT // WASHINGTON, DC

text-analytics-forum.com
Welcome to the Text Analytics Forum, Information Today's newest conference.

Text Analytics is a platform technology that adds depth and intelligence to any organization's ability to utilize that most under-utilized resource - text. The field of text analytics has seen an explosion of interest and new techniques in recent years and this is a great time for introducing this new conference.

The inaugural theme, Go Deeper, invites all who deal with text to take a deep dive into this exciting field. The Forum has something for all: whether you are new to the field and want to understand how it can add new capabilities or you are an experienced text analyst and want to see what the latest techniques and tools can add to your repertoire.

The Text Analytics Forum is designed to be a place for sharing ideas from how to get started to how to make the business case to the latest development best practices to use cases showcasing a range of cutting edge applications like tagging fake news and adding new structure to "unstructured" text.

I've been working in the field for over 10 years and I'm excited to be able to be part of introducing a new conference and taking text analytics to the next level. I look forward to meeting new and old friends and hope all who attend learn a lot and enjoy the talks and panels we have put together. In addition, all attendees have a chance to win a signed copy of my new book, Deep Text: Using Text Analytics to Overcome Information Overload, Get Real Value from Social Media, and Add Big(ger) Text to Big Data.

CONFERENCE CHAIR
Tom Reamy
Chief Knowledge Architect
KAP'S Group, LLC

---

**WIRELESS INTERNET**
Complimentary Wi-Fi is available in all meeting rooms.
1) Network: JWMarriott_CONFERENCE
2) Click on browser to open log-in page
3) Enter passcode: searchblox

**ENTERPRISE SOLUTIONS SHOWCASE GRAND OPENING RECEPTION**
Monday, November 6 • 5:00 p.m. - 6:30 p.m.
Y'all join us for chow and drinka as we celebrate the grand opening of the Enterprise Solutions Showcase. Featuring the top companies in the KM, CM, search, taxonomy, intranets, and text analytics marketplace, the Showcase offers attendees an opportunity to explore the latest product and service solutions. If you are looking for a particular product, evaluating competing systems, or keeping up to date with the latest developments, be sure to visit the Enterprise Solutions Showcase.

**SHOWCASE HOURS**
Monday, November 6 ......................................................... 5:00 p.m. – 6:30 p.m.
Grand Opening Reception

Tuesday, November 7 .......................................................... 10:00 a.m. – 6:00 p.m.
Networking Reception ......................................................... 5:00 p.m. – 6:00 p.m.

Wednesday, November 8 ..................................................... 10:00 a.m. – 4:00 p.m.

**NETWORKING RECEPTION**
Tuesday, November 7 • 5:00 p.m. - 6:00 p.m.
Enterprise Solutions Showcase
Join us in the Enterprise Solutions Showcase to mix and mingle with other conference attendees, speakers, and our conference sponsors.

**CONTINENTAL BREAKFAST & BREAKS**
A continental breakfast will be provided outside of the keynote room each morning from 8:00 a.m. – 8:45 a.m. Check your program for additional morning and afternoon break times and locations.

**CONFERENCE PRESENTATIONS**
Many speaker slides are available for download from the conference website (text-analytics-forum.com). Click on “PRESENTATIONS” and enter Username/Password: DC2017.

**KMWORLD BOOKSTORE**
Don’t forget to stop by the KMWorld Bookstore located in the Enterprise Solutions Showcase offering a great collection of KM and KM-related titles at discount prices! Open during Showcase hours.

**GET CONNECTED!**
Facebook, Twitter, #TextAnalyticsDC
### Wednesday, November 8

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00 a.m. –</td>
<td>CONTINENTAL BREAKFAST</td>
</tr>
<tr>
<td>8:30 a.m.</td>
<td>KEYNOTE • Capitol Ballroom • Wow, Woo, Win: KM for Customer Delight • Tom Stewart &amp; Patricia O’Connell</td>
</tr>
<tr>
<td>8:30 a.m.</td>
<td>KEYNOTE • Capitol Ballroom • Relevance Maturity Model: Revolutionizing With AI-Powered Search • Diane Tetraaf</td>
</tr>
<tr>
<td>9:00 a.m.</td>
<td>KEYNOTE • Capitol Ballroom • Energizing Communities of Practice • Naomie Manypeeny</td>
</tr>
<tr>
<td>10:00 a.m.</td>
<td>COFFEE BREAK in the Enterprise Solutions Showcase</td>
</tr>
<tr>
<td>10:15 a.m.</td>
<td>KEYNOTE • Hart (Meeting Room Level) • A Deep Text Look at Text Analytics • Tom Reamy</td>
</tr>
<tr>
<td>11:15 a.m.</td>
<td>KEYNOTE • Hart (Meeting Room Level) • Text Analytics Market Insights: What’s Working &amp; What’s Next? • Seth Grimes</td>
</tr>
<tr>
<td>12:30 p.m.</td>
<td>ATTENDEE LUNCHEON in the Enterprise Solutions Showcase</td>
</tr>
<tr>
<td>1:30 p.m. –</td>
<td>TRACK 1 • Technical • Hart (Meeting Room Level)</td>
</tr>
<tr>
<td>2:00 p.m.</td>
<td>COFFEE BREAK in the Enterprise Solutions Showcase</td>
</tr>
<tr>
<td>2:30 p.m. –</td>
<td>COGNITIVE COMPUTING &amp; GRAPH DATABASES</td>
</tr>
<tr>
<td>3:00 p.m.</td>
<td>AI vs. Automation: The Current State of Automated Content Tagging • Buch</td>
</tr>
<tr>
<td>3:15 p.m.</td>
<td>Graph Stores Combined With Text Analytics • Fried</td>
</tr>
<tr>
<td>3:30 p.m.</td>
<td>NLP &amp; Entity Extractors for Cognitive Computing &amp; Semantic Graph Databases • Asaman</td>
</tr>
<tr>
<td>3:45 p.m.</td>
<td>COFFEE BREAK in the Enterprise Solutions Showcase</td>
</tr>
<tr>
<td>4:00 p.m.</td>
<td>KEYNOTE • Grand Ballroom • Creating Unified Views of Data With Semantic Graphs • Joe Hilger</td>
</tr>
<tr>
<td>4:00 p.m.</td>
<td>CLOSING KEYNOTE • Grand Ballroom 3/4 • Knowledge Management in the Age of Digital Transformation • Jeanne Holm</td>
</tr>
</tbody>
</table>

### Thursday, November 9

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00 a.m. –</td>
<td>CONTINENTAL BREAKFAST</td>
</tr>
<tr>
<td>8:45 a.m.</td>
<td>KEYNOTE • Grand Ballroom 3/4 • KM Buy-in: Proven Practices • Stan Garfield</td>
</tr>
<tr>
<td>9:45 a.m.</td>
<td>KEYNOTE • Grand Ballroom 3/4 • Beyond the Box: How Search is Driving Data Access in a Hybrid World • Wil Hayes</td>
</tr>
<tr>
<td>10:00 a.m.</td>
<td>COFFEE BREAK</td>
</tr>
<tr>
<td>10:15 a.m.</td>
<td>TRACK 2 • Technical • Capitol Ballroom, Salon G</td>
</tr>
<tr>
<td>11:15 a.m.</td>
<td>MACHINE LEARNING VS. RULES</td>
</tr>
<tr>
<td>12:00 p.m.</td>
<td>MACHINE LEARNING, TAXONOMY, SEARCH</td>
</tr>
<tr>
<td>1:00 p.m.</td>
<td>AUTO-CATEGORIZATION &amp; SUMMARIZATION</td>
</tr>
<tr>
<td>2:00 p.m.</td>
<td>TEXT &amp; DATA TOGETHER</td>
</tr>
<tr>
<td>3:00 p.m.</td>
<td>MEASURING THE RESULTS</td>
</tr>
<tr>
<td>4:00 p.m.</td>
<td>KEYNOTE • Grand Ballroom 3/4 • Creating Unified Views of Data With Semantic Graphs • Joe Hilger</td>
</tr>
<tr>
<td>4:00 p.m.</td>
<td>CLOSING KEYNOTE • Grand Ballroom 3/4 • Knowledge Management in the Age of Digital Transformation • Jeanne Holm</td>
</tr>
<tr>
<td>5:00 p.m.</td>
<td>Text Analytics Forum</td>
</tr>
</tbody>
</table>
Wednesday, November 8

**KEYNOTE**

**Text Analytics Market Insights: What’s Working & What’s Next?**
11:45 a.m. – 12:30 p.m. • Hart (Meeting Room Level)
Seth Grimes, Principal Consultant, Alta Plana Corporation
Text analytics emerged in the mid-2000s, a collection of technologies, solutions, and practices aimed at meeting a diversity of business challenges. A decade in, what’s new and promising, what’s tried-and-true, and what’s on the horizon? How has the market evolved—both demand and supply—and how should practitioners, solution providers, business analysts, and investors stay on top of developments?

**Attendee Luncheon**
12:30 p.m. – 1:30 p.m.

**TRACK 1**

**AI & TEXT ANALYTICS**
1:30 p.m. – 2:15 p.m.

**Real-Time, Actionable Analytics**
Catherine Havasi, CEO, Luminoso
Leveraging AI to instantly analyze and respond to customer feedback is quickly becoming critical to business success, but making this leap is easier said than done. Many companies which move to real-time customer feedback systems struggle because either their analytics or business processes—or both—aren’t synchronized or fully established. This session shows how AI-based analytics systems can successfully be adopted and merged with appropriate business practices and examines case studies to reveal how these approaches differ across industries and use cases.

**AI vs. Automation: The Current State of Automated Content Tagging**
Joseph Busch, Principal Consultant, Taxonomy Strategies
Some technologies such as IBM Watson are being touted as AI. In response, there are new AI offerings from the large enterprise software companies as well as many startup companies. But is this AI or automation? This talk discusses the difference and argues that these offerings use entity extraction and business rules rather than AI. However, there are real opportunities to use this new technology to automate content tagging.

**COGNITIVE COMPUTING & GRAPH DATABASES**
2:30 p.m. – 3:15 p.m.

**Graph Stores Combined With Text Analytics**
Jeff Fried, CTO, BA Insight
Text analytics can discover and add underlying structure to content, providing some remarkable new capabilities for the enterprise. This session focuses on the discovery of relationships between data and the population of graph databases and graph search. There are now more than 30 graph databases on the market, ranging from neo4j to SQL Server 2017, and graph search has become mainstream (including Lucene 6, the Microsoft Graph and many more). Text analytics is important for these to provide value to organizations.

**NLP & Entity Extractors for Cognitive Computing & Semantic Graph Databases**
Jans Aasman, CEO, Franz Inc.
NLP and entity extractors make up an important part of our use cases in cognitive computing. We discuss how terminology systems and knowledge bases are used in combination with NLP and entity extractors to greatly enrich the contents of our data infrastructures.

**Coffee Break**
3:15 p.m. – 4:00 p.m.
Last Chance to Visit the Exhibits in the Enterprise Solutions Showcase
**Ask the Experts Panel**

4:00 p.m. – 5:00 p.m.

A panel of four text analytics experts answer questions that have been gathered before the conference, during the conference, and some additional questions from the program chair and sponsors.

**TRACK 2 • Business & Applications**

**Rayburn (Meeting Room Level)**

**CASE STUDIES**

1:30 p.m. – 2:15 p.m.

**Building Text Analysis Models to Understand & Predict Adverse Events Occurring With FDA-Approved Drugs**

**Qais Hatim**, Operation Research and Data Analyst, U.S. Food and Drug Administration (FDA)

Various drugs approved by the FDA have been associated with adverse medical events in patients. These events are reported to the FDA through different sources, including physicians at hospitals and clinics, pharmacists, and patients. Most of these adverse events are documented in texts. However, manual analysis of these texts for root cause are time consuming and produce qualitative results at best. This session shows how building text analysis models in order to understand the themes related to drug adverse events can help us quantitatively understand which adverse events are most common within a cluster of FDA-approved drugs.

**A New Way of Working Graph & Semantics, Text Analytics, & Linked Data**

**Jeremy Bentley**, CEO, Smartlogic

Using case studies of real-world client projects, Smartlogic’s CEO presents, discusses, and demonstrates how post-relational databases, text analytics, AI, semantics, and linked data are delivering rapid returns on investment in data intensive industries. Cases range from predictive analytics and financial risk assessment to compliance, superior customer service, and unified enterprise intelligence within industries including banking, life sciences, media, and healthcare. The talk looks at the technology, the opportunity, lessons learned, and the keys to project success.

**SEARCH & TEXT ANALYTICS**

2:30 p.m. – 3:15 p.m.

**Leveraging Text Analytics to Build a Personalized Information Retrieval Environment**

**Pengchu Zhang**, Technical Staff, Sandia National Laboratories

To improve the effectiveness of information findability and usability, we are developing a new mechanism to understand users’ interests and predict the information that will be most relevant to their needs. We analyze the technical documents published by members of the workforce and build models that can be used to match user’s requests with the best available content. We utilize an existing hierarchical taxonomy as part of the clustering effort in order to provide preliminary labels for the clusters. The information retrieval environment we are building will not only support retrieval of relevant corporate information upon request, it is designed to proactively notify targeted members of the workforce when relevant information becomes available.

**Using Text Analytics, Taxonomy, & Search to Probe Ignorance & Risk**

**Patrick Lambe**, Principal Consultant, Straits Knowledge

In this talk, Patrick Lambe takes an unconventional look at how text analytics, taxonomies and search can be used in concert to probe areas of ignorance, not just uncover and organize what is already known, via three problem cases from the areas of public health and public transport. We demonstrate how elements of the search and discovery technology stack can be used to detect patterns in the environment to address or mitigate these types of problems.

**Coffee Break • 3:15 p.m. – 4:00 p.m.**

Last Chance to Visit the Exhibits in the Enterprise Solutions Showcase

**Ask the Experts Panel**

4:00 p.m. – 5:00 p.m. • Hart (Meeting Room Level)

A panel of four text analytics experts answer questions that have been gathered before the conference, during the conference, and some additional questions from the program chair and sponsors.

---

**Transform the Flood of Social Media and Big Data Into a Useful Asset**

“One of the most thorough walkthroughs of text analytics ever provided.”

—Jeff Catlin, CEO, Lexalytics

Deep text is an approach to text analytics that adds depth and intelligence to our ability to utilize a growing mass of unstructured text the world is drowning in. Here, author Tom Reamy explains what deep text is and surveys its many uses and benefits. He describes applications and development best practices; discusses business issues, including ROI; provides how-to advice and instruction; and offers guidance on selecting software and building a text analytics capability within an organization.

This is an important book for anyone who needs to be on the text analytics cutting edge, from developers and information professionals who create, manage, and curate text-based and Big Data projects to entrepreneurs and business managers looking to cut costs and create new revenue streams. Whether you want to harness a flood of social media content or turn a mountain of business information into an organized and useful asset, Deep Text will supply the insights and examples you’ll need to do it effectively.

---

Tom Reamy will be signing copies of Deep Text at the KMWorld Bookstore at the following times:

**Monday, November 6, 5:30–6:00 P.M. & Tuesday, November 7, 12:30–1:00 P.M.**

---

(800) 300-9868 • text-analytics-forum.com
**Thursday, November 9**

### Continental Breakfast  
8:00 a.m. – 8:45 a.m.

### KEYNOTE

**KM Buy-In: Proven Practices**  
8:45 a.m. – 9:45 a.m.  
Grand Ballroom 3/4  
Stan Garfield, Author, Proven Practices for Promoting a Knowledge Management Program

For a KM initiative to be successful, knowledge managers must secure the support of senior leaders before implementation. Early top management buy-in results in funding, resources, advocacy, usage, broad organizational support, and success—the program yields its expected benefits. KM is spoken of and written about positively by leaders, stakeholders, and users. Hear from our long-time KM practitioner about proven practices illustrated by real-world examples for securing resources, active participation, and ongoing advocacy from top leadership. Get lots of tips for leading an effective, sustainable KM program that is seen as essential to the success of companies in different industries, of different sizes, and with different cultures.

**Beyond the Box: How Search Is Driving Data Access in a Hybrid World**  
9:45 a.m. – 10:00 a.m.  
Grand Ballroom 3/4  
Will Hayes, CEO, Lucidworks

For more than a decade, search technology has been used as the primary access point to the mountains of knowledge and data sitting behind an organization's firewall. As environments evolve to account for private and public clouds, search is evolving beyond just the box to an API for human information. Will Hayes explores that evolution and talks about how search technologies and professionals play a key role in the enterprise cloud migration strategy.

### Machine Learning, Taxonomy, Search

10:15 a.m. – 11:00 a.m.

**Combining Machine Learning, Text Analytics, & Semantic Web for Automated Tagging**  
Dan Segal, Corporate Taxonomist, IBM

This talk describes work that the IBM Taxonomy Squad has done to develop an enterprise-scale service that automates the extraction of entities and the generation of meaningful metadata. We cover the approach that was taken to design a solution architecture that leverages a corporate knowledgebase and integrates best-of-breed services in taxonomy and ontology management, NLP, machine learning, text annotation, and entity extraction.

**The Savior Machine: Text Analytics, Machine Learning, & the Role of Taxonomy**  
Dave Clarke, CEO, Synaptica LLC

The promise of machine learning has become a practical reality in today’s enterprise, but companies often struggle with implementation or reliable results. One fundamental issue is the common “garbage in, garbage out” problem. Poor input stems from the lack of clean data or unclear results from unstructured data analysis feeding machine learning models. Well-built taxonomies powering clear analytics rules are an important infrastructure need often overlooked in data science activities. Come learn more about the role of taxonomy and text analytics as sources of clean data for machine learning.

### Attendee Luncheon & Keynote

**Cognitive Search & Analytics: What It Is & Why You Should Care**  
12:00 p.m. – 1:00 p.m.  
Grand Ballroom 3/4  
Scott Parker, Senior Product Marketing Manager, Cognitive Search & Analytics, SINEQUA

If you are a believer in the data-driven organization (or even just curious) and have ever wondered what could happen if you cleverly combined the power of data collection, indexing, text mining, search, and machine learning into a unified platform and applied it within the enterprise, this talk is for you! Come learn about the state of cognitive search and analytics technology and how it is enabling great companies across a wide swath of industries to amplify mission-critical expertise within their business in a surprisingly short amount of time. Our speaker illustrates the technology in action with real-world examples.

**Auto-Categorization & Summarization**  
1:00 p.m. – 1:45 p.m.

**Auto Categorization by Taxonomy: Pros, Cons, & Pragmatics**  
Dave Clarke, CEO, Synaptica LLC

The terminologies that form taxonomies, thesauri, classification schemes, and name authorities aim to define all concepts unambiguously. These conceptual definitions are, however, primarily written for a human audience and are only partially meaningful to automated categorization processes. This talk explores how automated categorization rules can be synthetically generated by mining the terminology and semantic relationships found in traditional knowledge organization systems. We examine the pros, cons, and limitations of using categorization rules derived from KOS and discuss how they can then be refined and extended using human-curated categorization rules.

**Search, Semantic Analysis, Text Mining**  
Dmitriy Soubbotin, CEO & Founder, Semantic Engines

This talk presents an original approach to processing search results. Rather than showing the usual 10 blue links to webpages, the software creates a text summary of those webpages—a narrative on the topic of the user’s query. The narrative gives the user a quick way to understand the key information on his query. This approach is best applicable to queries that are informational in nature, i.e., those where the user wants to understand a particular subject and get a quick grasp of a concept, an event, a product, or a public figure. The talk focuses on the merits and drawbacks of the approach and comparison with other techniques of presenting the answer to the user’s query.

**Automatic Classification: Rules-Based vs. Training-Set-Based Bakeoff**  
Jeff Fried, CTO, BA Insight

Machine learning techniques can be used effectively for a wide variety of text analysis scenarios, such as reputation monitoring on social media, fraud detection, patent analysis, and e-Discovery. But to apply them well, you need to understand where the limits and pitfalls are in the technology, and you need to understand your data and the problem you are trying to solve. This session outlines an approach that uses text analytics to help understand the characteristics of your data, followed by selection and tuning of linguistic and statistical processing and machine learning parameters to address the application at hand. We highlight three real-world projects that used this approach and show how they worked, what went right and wrong, and how they evolved over time.
**TEXT & DATA TOGETHER**

2:00 p.m. – 2:45 p.m.

*Extracting Content for Linked Data Triples*

*Margie Hlava, President, Access Innovations*

There is much talk about building triple stores from source content, but most of the models are just that: without content to back them up. This session covers a case study of building a triple store to support search and other use cases from nearly 6 million documents. It also looks at the extraction or mining process for pulling 22 types of triple sets for full text and redeploying them for search queries. Lessons learned are also covered.

*Text Analytics in the Context of Semantic Datasets & Ontologies*

*Barislav Popov, Head, Text Analytics and Annotation, Ontotext AD*

This session addresses the main principles of extracting entities and relationships from unstructured content against ontologies and semantic data sets. We give industry examples of business cases and key components of semantic technology architectures including text analytics and supporting data and metadata governance workflows. Finally, we demonstrate semantic annotation, talking about the challenges organizations face in this regard and some of the important lessons learned in more than 15 years of industry experience.

**MEASURING THE RESULTS**

3:00 p.m. – 3:45 p.m.

*Measuring Auto-Categorization Quality*

*Larry Lemport, Director, Product Research and Planning, Bloomberg BNA*

Auto-categorization is “auto” in part—there is much in the process that still requires old-fashioned human judgment. One critical step on the human side of the fence is to evaluate the quality of results so refinements can be developed and fed back into the process. But how do you measure quality when human indexers themselves apply topics inconsistently and often differ over applicability of topics? This case study explains how one publisher approached quality assessment in light of human variability and details how classic recall and precision measures were adjusted to provide a user-focused sense of auto-categorization quality.

*Information Retrieval Performance Measurement Using Extrapolated Precision*

*Bill Dimm, Founder & CEO, Hot Neuron LLC*

In search, there is often a trade-off between recall and precision, and this impacts any evaluation of approaches: If one system achieves higher recall but lower precision, is it better? Traditionally, this situation has been addressed by using a measure that combines precision and recall into a single number, such as the F1 score. F1 makes strong assumptions about the amount of precision you can trade for a little more recall, and those assumptions are not always appropriate. In some contexts, recall and precision have very different significance. This talk presents a novel performance measure called the extrapolated precision, which avoids making such strong assumptions about allowed trade-offs between precision and recall.

**KEYNOTE**

Creating Unified Views of Data With Semantic Graphs

*Joe Hilger, COO, Enterprise Knowledge, LLC*

In recent years, document-centric search over information has been extended with the use of graph-based content and data models. The implementation of semantic knowledge graphs in enterprises is not only improving search in a traditional sense, but opens up a path of integrating all types of data sources in a most agile way. Linked data technologies have matured in recent years and can now be used as the basis for numerous critical tasks in enterprise information management. Hilger discusses how standards-based graph databases can be used for information integration, document classification, data analytics, and information visualization tasks. He shares how a semantic knowledge graph can be used to develop analytics applications on top of enterprise data lakes and illustrates how a large pharmaceutical company makes use of graph-based technologies to gain new insights into its research work from unified views and semantic search over heterogeneous data sources.

**CLOSING KEYNOTE**

**KM in the Age of Digital Transformation: Magic Sauce for a Successful Future**

*Jeanne Holm, Senior Technology Advisor to the Mayor, Assistant General Manager at City of Los Angeles, & former Knowledge Architect, Jet Propulsion Lab*

At the cross-section of innovation, open data, and education, our speaker, a former government KM practitioner, shares her thoughts about the challenges and opportunities for organizations and communities in the coming years. She discusses empowering members of our communities and improving services using new tech like AI, machine learning, virtual and augmented reality, Internet of Things, predictive analytics, gamification, and more. How will people interact and share knowledge over the next decade? Are we moving toward anticipatory knowledge delivery (just enough, just in time, just for me), being in the flow of work at the teachable moment, establishing trust in a virtual environment, and learning from peer-to-peer marketplaces like Airbnb and Uber? Our longtime KM practitioner shares her insights about the evolving digital transformation of every part of our world and hints at the magic sauce we need for a successful future!

**FAKE NEWS & BAD AD PLACEMENT**

**10:15 a.m. – 11:00 a.m.**

*News Analytics System*

*Lipika Dey, Principal Scientist, Innovation Labs, Tata Consultancy Services Limited*

For modern digital enterprises, the key to survival is held by real-time predictive analytics done with heterogeneous data gathered from multiple sources—layered with contextual intelligence. The data is a mix of structured and unstructured data. Establishing contextual relevance requires systems imbued with deep reasoning capabilities that can link relevant pieces of information from within and outside the organization. This talk presents the outlines of a framework that can gather news events in real time, classify them, reason with them, and finally link them to an enterprise information repository and thereby generate alerts or early warnings for subscribed users. The framework is presented through a number of case studies.

**Content Meets Interest—Contextual Ad Targeting by Means of Cognitive Computing**

*Heiko Beier, CEO, MORESOPHY*

The globally increasing tendency for political populist and media criticism has raised the sensitivity of brands to avoid misplacement of their own campaigns in negative and compromising contexts (bad ads). However, ad targeting is predominantly based on behavioral targeting techniques that heavily rely on (cookie-based) user profiling. The talk showcases a solution for real-time contextual targeting that is exploiting the full power of cognitive computing to match campaigns to online users’ real interests. The approach abandons tracking of any kind of user data and at the same time increases the precision of ad targeting on a real semantic level—beyond what can be achieved with keyword-based methods.

**CASE STUDIES II—BANKS & PUBLISHING**

**11:15 a.m. – 12:00 p.m.**

*Text Analytics & KM*

*Daneila Collaguazo, Consultant, & Kyle Strand, Knowledge Management Specialist, Inter-American Development Bank*

The bank as a nonprofit financial cooperative, which acts as the main source of multilateral financing for the execution of projects in Latin-America, intends to provide solutions to development challenges and support in the key areas of the region. The PMR (Project Monitoring Report) and the PCR (Project Completion Report) are two documents that include a section, known as Lessons Learned, which gathers the challenges and lessons learned of each operation. In order to make those lessons more accessible, reusable, and personalized to the several users, a proof of concept using machine learning and natural language processing technologies, was fulfilled. The scope of the proof of concept consisted of the extraction of the documents’ lessons and their corresponding classification. The aim of this talk is to present the insights gained on behalf of technology during the fulfillment of the proof of concept and specifically to present the results of the different classification algorithms.
**ATTENDEE LUNCHEON & KEYNOTE**

**Cognitive Search & Analytics: What It Is & Why You Should Care**

**Scott Parker**, Senior Product Marketing Manager, Cognitive Search & Analytics, SINEQUA

See page 56 for complete description.

**TEXT ANALYTICS & TAXONOMY**

**1:00 p.m. – 1:45 p.m.**

**Bringing It All Together (At Last): Integrating Structured & Unstructured Information With Text Analytics & Ontologies**

**Zach Wahl**, Principal, Enterprise Knowledge

Organizations are always looking for better ways to integrate their structured (databases and reports) and unstructured (documents and webpages) information. This concept is not new; in fact, it has been the primary information management goal for many years. The difference is that the text analytics are performed in the native language of the original text, not just English. This talk shares real-life examples of how this is done in large repositories using text analytics and ontologies. Session attendees will understand what an ontology is and how it can be merged with text analytics tools to provide better analytics for their data scientists.

**Taxonomies & Text Analytics**

**Gary Carlson**, Principal, Factor

This presentation discusses two recent projects where enterprise projects have benefited from direct interactions between taxonomies/ontologies and text analytics. While these are often seen as competing work streams, our recent work continues to build on the idea that complex information-rich projects require both, and that pursuing one while abandoning the other often leads to poor results or project failure.

**NEW APPLICATIONS**

**2:00 p.m. – 2:45 p.m.**

**Human-Like Semantic Reasoning**

**Bryan Bell**, EVP, Expert System

To address the complexity of language ambiguity requires a technology that can read and understand text the way people do. This session explains the concepts behind linguistic analysis, word disambiguation, and semantic reasoning to read and understand content the way people do. It explains the concepts that support a semantic platform, demonstrates a semantic engine, and how one mobile phone carrier deployed a self-help solution that automatically answered 24,000,000 customer questions annually with 94% precision, and shows a knowledge platform that automatically organizes hundreds of data sources and millions of unstructured documents around multiple corporate taxonomies and entity clusters using dynamically generated metadata in a precise and complete way.

**Breaking Down Silos With Text Analytics**

**Fiona McNeil**, Global Technology Product Marketer, SAS

The next phase of how we communicate has already started. Popularized by Siri, Alexa, and the like, natural language interaction (NLI) has achieved commercial Q&A success. For organizations looking to adopt new experiences with their customers, NLI holds promise. But there is a big difference between AI applications—the distinction is the degree to which they are intelligent. This talk examines the considerations for enterprise application of NLI and how to avoid applications that just drive more white noise.

**APPLICATION ISSUES**

**3:00 p.m. – 3:45 p.m.**

**Leveraging Text Analytics to Build Applications**


In the world of scholarly publishing (as well as many other industries—such as KM/information conferences), meeting organizers are inundated with submissions for inclusions in conference programs. Given a large set of submissions, how can we develop tools to cluster submitted manuscripts into tracks based on topical similarity? This talk describes a project that used a subject taxonomy, NLP, and other text analytics tools as well as a large corpus of documents to construct an application to cluster submitted manuscripts based on topical similarity, including a GUI interface to interact with and analyze the results. This is not intended as a detailed technical talk (no slides of code!), nor is it intended as a product spotlight; the focus is on using known/existing text analytics tools to construct purpose-built applications to solve specific document-centric problems.

**Maximizing Analytic Value From Multi-Language Text Feeds**

**Christopher Blow**, SVP, U.S. Public Sector, Basis Technology

The universe of text analytics is largely constrained to the output of the entire human race. This can and does result in huge, petabyte-scale problems. Technologies for this scalability, computational distribution, deep learning, resolution, and semantic expression are all new within the last 10 years, and their combination is revolutionary. Key to putting all of this together is that the text analytics are performed in the native language of the original text, prior to the inevitable loss of fidelity in machine or human translation. This talk covers a number of use cases including counterterrorism, knowing your customer, border security, disease tracking and detection, and countering fake news and conspiracy theories.

**KEYNOTE**

**Creating Unified Views of Data With Semantic Graphs**

**Joe Hilger**, COO, Enterprise Knowledge, LLC

See page 57 for complete description.

**CLOSING KEYNOTE**

**KM in the Age of Digital Transformation: Magic Sauce for a Successful Future**

**Jeanne Holm**, Senior Technology Advisor to the Mayor, Assistant General Manager at City of Los Angeles, & former Knowledge Architect, Jet Propulsion Lab

See page 57 for complete description.