Maximize the Effectiveness of Your Text Analytics Initiatives

Webinar

Moderator: Meta Brown
Author, Data Mining for Dummies
metabrown@metabrown.com
PARTICIPANTS

Howard Lyeth  Stephen Scarr  Ramkumar Ravichandran

L.L.Bean  eContext  VISA
Howard Lyeth

Senior Marketing Analyst
L.L. Bean
Freeport, ME

“I am passionate about understanding consumer behavior and relationships with brands. I thrive on providing insights and recommendations to impact profitability based on innovative approaches and solutions. Past experience has included the development of portals for internal and client use, analytics tools, and a variety of advanced analytics.”
REACTIVE AND PROACTIVE APPROACHES TO TEXT ANALYTICS

February 19, 2014

Howard Lyeth
Senior Marketing Analyst, Customer Insights
hlyeth@llbean.com
The business

- 3 Channels – Web, Brick & Mortar, and Phone/Catalog
- Consumers ranging from babies to their great grandparents
- Products include hunting, fishing, camping, and casual apparel and gear

The data

- Customer satisfaction, service, and experience surveys
- Annual studies on service, brand image, and more
- Concept and product testing
- Retail “Comment Box”
- Compliments and complaints via phone and email
- Customer Letters
- Product ratings and reviews
- Social media
- E-Service Web Emails
- ForeSee
- Retail Receipt Survey
The Reactive Approach

What’s going on and what do we do about it?

Identify and resolve customer issues

1. Identify Existence
2. Identify Issue
3. Communicate internally
4. CS closes loop with customer

Identify unique or emerging issues

- Communicate to internal stakeholders
- Close loop internally

Identify potential operational issues

1. Identify trends
2. Communicate internally
3. Close loop with operations

Reporting

Report/inform
The Proactive Approach
What can we do to drive business?

- Identify experiences that drive KPIs (up or down)
- Identify experiences that drive behavioral loyalty and value
- Develop products and services
  - Find out what people want
  - Make the most of emerging trends
- Identify areas for product development
- Develop or better understand existing customer segments

Text sources
Survey data
Customer information
Transactional data
Text sources
Things to consider

• Rule-based categorization vs NLP

• Context of data – visitors or buyers, old/new customers, channel
  • Linking data sources with different contexts

• Linking to customer data

• Linking to other quantitative data

• Reporting vs analytical capabilities
REACTIVE AND PROACTIVE APPROACHES TO TEXT ANALYTICS

February 19, 2014

Howard Lyeth
Senior Marketing Analyst, Customer Insights
hlyeth@llbean.com
14th Text Analytics Summit

SAVE THE DATE!

June 15th & 16th, 2015 - New Yorker Hotel
New York City

- Drive business value with text analytics: identify the goals of text analytics and how to achieve them
- Integrate unstructured and structured data: get deeper business insights and improve decision-making in your organization
- Customer engagement and understanding: improve CEM through sentiment analysis and social listening techniques
- Create more effective marketing campaigns: use text analytics to improve segmentation and targeting

Project Director (Event Contact): Brian Parke
Tel: +1 (201) 234-4755
Email: brianparke@fc-bi.com
“As CEO of Info.com, I am responsible for all aspects of the company's development. An operator with 20 years experience in managing businesses with a strong marketing background and a passion for big data and analytics. Info.com is an independent search platform with 8 million unique users. From a single search query, Info.com provides results from the leading search engines. Info.com is also partnered with 8 vertical search providers.”
Curated Taxonomy vs. Machine Learning

THE AGE OF CONTEXT

Presented by

eContext

sscarr@econtext.com

Stephen Scarr, CEO

February 19th, 2015
To identify digital audiences and their constantly evolving interests the answers are in the data.

Textual data such as social media posts, blogs, instant messages or search keywords are notoriously difficult to classify in the correct context, particularly with machine learning.
THE NEED FOR CLASSIFICATION

Vital to this effort is a classification technology that can accurately structure text to provide actionable insights.
“People share and put billions of connections into this big graph every day. We don't want to just add incrementally to that. We want, over the next five or ten years, to take on a road map to try to understand everything in the world semantically and map everything out. These are the big themes for us and is what we are going to try and do over the next five or ten years. That is what I have tried to focus us on…”

Mark Zuckerberg, CEO, Facebook: TechCrunch Keynote, September 11, 2013
"The Web is so vast...you need to extend categorization and make sense of the content and have a Web ordered for you...One of the key pieces is you have to understand and decide what the Ontology of entities is. Meaning how things are named and how are they organized into hierarchies...By mapping people’s search habits you pull all their content together and have a feed of information that is the web ordered for you."

Marissa Mayer, CEO, Yahoo: Davos World Economic Forum Interview, January 24, 2013
TEXT CLASSIFICATION APPLICATIONS

- Social Insights
- User Profiles
- Contextual RTB
- Adult Filters
- Content Mapping
- Voice Classification
- Media Planning
- Message Classification
- Keyword Discovery & Local PPC
- Content Recommendation
On Content Recommendation:

“The Huffington Post’s personalization, which is built on top of AOL’s 2014 Gravity acquisition, increased article click-through rate by over 50 percent on average.”

THE CHALLENGES: STRUCTURE

(Sample path for electronics categories)
The Challenges: Structure

Word cloud visualizations may look nice, but provide few actionable insights.
**THE CHALLENGES: VERSATILITY**

- Many classifiers require prior training on corpuses of text

- Problematic to apply immediately to documents that vary from training corpus

- Especially difficult with social media

- Classification can take hours
THE CHALLENGES: ACCURACY

These real false positives highlight potential errors in machine classification:

<table>
<thead>
<tr>
<th>Text Sample</th>
<th>Incorrect Machine Mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>I've got a <strong>novel</strong> idea. How about you ask me first?</td>
<td>Books &amp; Literature</td>
</tr>
<tr>
<td>Arsenal and Chelsea ended in a <strong>tie</strong>. Football needs to fix that.</td>
<td>Neckties</td>
</tr>
<tr>
<td>I have <strong>no doubt</strong> you are wrong.</td>
<td>Rock Band</td>
</tr>
<tr>
<td>Winter is to relax. <strong>Happy days</strong>.</td>
<td>TV Shows</td>
</tr>
<tr>
<td>I'd love to go but I have no <strong>bread</strong>. Take lots of pics for me.</td>
<td>Food</td>
</tr>
<tr>
<td>They're going to <strong>stock</strong> up before this concert.</td>
<td>Stocks</td>
</tr>
<tr>
<td><strong>Suit</strong> yourself!</td>
<td>Men’s Apparel</td>
</tr>
<tr>
<td>She’s got too much <strong>baggage</strong> for blind dates.</td>
<td>Luggage</td>
</tr>
</tbody>
</table>
Curated Taxonomy: Top-down hierarchy of topics, where each node is assigned specific positive and negative vocabulary rules.

For example, the topic “RVs” would have positive vocabularies including ‘recreational vehicles’, ‘motor homes’ and ‘travel trailers’.

The topic “Shampoo” would have negative vocabularies including ‘carpet’, ‘dogs’, and ‘Warren Beatty’.
Management believes **eContext** is the **world’s largest taxonomy of commercial & social topics**. It was built over 7 years by 200 subject matter experts.

Without prior training, eContext categorizes unstructured text from any source, enabling businesses to **extract unique insights** from digital activity.

### Structure
- **415,000 Categories**, in a hierarchical structure up to **21 tiers deep**

### Versatility
- **25 Verticals**, including Cars, Arts & Entertainment, Home & Garden, and Health

### Accuracy
- **55 million vocabulary rules** create intelligent groupings of text by topic, classifying **55,000 words per second**
CLASSIFICATION WITH CURATED TAXONOMY

(Actual eContext sample path)

Computers & Electronics → Computers → Computer Products

Wearable Technology ← Internet-Enabled Devices ← Personal Computers

Smartwatches → Pebble Smartwatches → Pebble Steel Smartwatches
Classification with Curated Taxonomy

- With one hierarchy, clients can treat **web searches, browsing history, social posts**, and other activity as vital ingredients in **one holistic user experience**.

![Image of a diagram showing the hierarchy of categories like Computers & Electronics, Computers, Computer Products, Personal Computers, Internet-Enabled Devices, Wearable Technology, Smartwatches, Pebble Smartwatches, and Pebble Steel Smartwatches.]

![Image of a Twitter post and a Wikipedia article on Pebble smartwatches.]

**How much is Pebble Steel?**
**Classification with Curated Taxonomy**

eContext’s 55 million vocabulary rules instill “common sense” in an automated classification system.

<table>
<thead>
<tr>
<th>Text Sample</th>
<th>eContext Tier 1</th>
<th>eContext Category Path (Last 3 Tiers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hope to have long johns!</td>
<td>Apparel</td>
<td>Underwear::Casual Underwear::Long Underwear</td>
</tr>
<tr>
<td>I deserve it. What's the M3 convertible cost?</td>
<td>Vehicles</td>
<td>BMW::BMW 3 Series::BMW M3 Convertibles</td>
</tr>
<tr>
<td>Need to pay to remove moss on the roof</td>
<td>Home &amp; Garden</td>
<td>Roofing Services::Roof Cleaning Services::Roof Moss Removal Services</td>
</tr>
<tr>
<td>Listening to yellow submarine</td>
<td>Arts &amp; Entertainment</td>
<td>Music &amp; Musicians::Rock Music::The Beatles</td>
</tr>
<tr>
<td>Stomach pains after the match</td>
<td>Health</td>
<td>Health Conditions &amp; Treatments [No Condition Specified]::Pain::Abdominal Pain</td>
</tr>
<tr>
<td>Le Bristol hotel in paris is best</td>
<td>Travel</td>
<td>Hotels in France::Hotels in Paris France::Hotel Le Bristol in Paris France</td>
</tr>
</tbody>
</table>
“BEST OF BREED” REQUIREMENTS
Profile entire audiences to discover statistically high levels of interest.
eContext’s UI offers an intuitive, customizable interface to view results of real-time classification.
“Best of Breed” Requirements

When a text string is unrecognized (i.e. does not correspond to any of eContext’s 55 million vocabulary rules) the text can be compared against eContext’s corpus of classified search and social data in order to form a most-likely classification.

If the author of the text has been profiled by eContext, their interests can further inform this classification.

These corpus-based classification can be used to automatically suggest new vocabularies and/or taxonomy nodes.

“How tall is the efltwr?”

<table>
<thead>
<tr>
<th>Unrecognized Text</th>
<th>Possible Classifications</th>
<th>User Profile</th>
<th>Final Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>“How tall is the efltwr?”</td>
<td>Empire State Building</td>
<td>Eiffel Tower</td>
<td>Statue Of Liberty</td>
</tr>
</tbody>
</table>
FREE TRIAL

Free trial at: classify.econtext.com/signup

WE HOPE TO HEAR FROM YOU

Stephen Scarr

CEO

sscarr@econtext.com

To obtain this deck:

Chris Fendrich

Business Development Manager

cfendrich@econtext.com
14th Text Analytics Summit

SAVE THE DATE!

June 15th & 16th, 2015 - New Yorker Hotel
New York City

- Drive business value with text analytics: identify the goals of text analytics and how to achieve them
- Integrate unstructured and structured data: get deeper business insights and improve decision-making in your organization
- Customer engagement and understanding: improve CEM through sentiment analysis and social listening techniques
- Create more effective marketing campaigns: use text analytics to improve segmentation and targeting

Project Director (Event Contact): Brian Parke
Tel: +1 (201) 234-4755
Email: brianparke@fc-bi.com
“As Director, Analytics for the Digital Developed Markets department at Visa Inc., I am responsible for helping the Leadership & Stakeholders with actionable insights derived from Analytics. The questions span the whole spectrum across the Product, Marketing, Sales & Relationship. We leverage any of the various options - Strategic analysis, Advanced Analytics, Text Analytics or Mining depending on the problem being solved. The operating philosophy of our team is that ‘our success is the success of our Customer.’”
ACTIONABILITY OF INSIGHTS
What is actionability?

Why this confusion? And specific challenges for Text Analytics

Recommended approaches
WHAT DOES ACTIONABILITY MEAN TO ANALYTICS CONSUMERS?

- Relevancy to the business question
- Fit within overall strategy/business goals/big enough to act on
- Timely and available when, how and where needed
- Easy to understand
- Trustworthy & reliable
WHY THIS CONFUSION? PHILOSOPHICAL DIFFERENCES...

Analysts are from Mars and Analytics Consumers are from Venus...

<table>
<thead>
<tr>
<th>Language they speak</th>
<th>Numbers, graphs, lifts, accuracy, insights</th>
</tr>
</thead>
<tbody>
<tr>
<td>What excites them</td>
<td>Brilliance of approach</td>
</tr>
<tr>
<td>How they think</td>
<td>Detail oriented</td>
</tr>
<tr>
<td>What they can</td>
<td>Time for accuracy</td>
</tr>
<tr>
<td>compromise on</td>
<td>Scientists who deal with facts</td>
</tr>
<tr>
<td>Biggest difference</td>
<td>If we do this, then we get this...</td>
</tr>
<tr>
<td></td>
<td>Simplicity of the answer</td>
</tr>
<tr>
<td></td>
<td>Big picture</td>
</tr>
<tr>
<td></td>
<td>Perfection for timely action</td>
</tr>
<tr>
<td></td>
<td>Artists who deal with gut</td>
</tr>
</tbody>
</table>
...AND SOME TACTICAL REASONS

- Analytics has multiple value propositions
- Ambiguous questions, fit with overall corporate strategy and/or success criteria
- Instrumentation gaps; small counts; time & resource constraints
- Significant change in the type of Consumer of insights
- Waterfall/Agile approach of Project Management vs. Lean Approach
- Wrong hiring – statistician/data scientist for Business Analytics and vice versa

Specifically with Text Analytics...

- Source issue (Data availability, collection & preparation challenges)
- Too early a deployment of Text Analytics; lack of cross verification data points
- Impact Sizing and “thought through” sales pitch usually not done
- Tying it up with A/B Testing tricky
RECOMMENDATIONS ON HOW IT CAN BE MITIGATED

TYPICAL MATURITY CURVE OF AN ANALYTICS PRACTICE...

Plan
Inform
Act
Mine
Predict
Optimize

Maturity phases of Analytics Practice
## MULTIPLE VALUE PROPOSITION FROM ANALYTICS...

<table>
<thead>
<tr>
<th>Phase</th>
<th>Description</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan</td>
<td>Partner with Stakeholders to list &amp; prioritize key business questions</td>
<td>Prioritized list of metrics, analyses &amp; insights. Roll up &amp; monitoring framework</td>
</tr>
<tr>
<td>Inform</td>
<td>Size all the Key behaviors with KPIs and high level drilldowns</td>
<td>Baseline understanding of various key behaviors and levers</td>
</tr>
<tr>
<td>Act</td>
<td>Root cause analysis for addressing key issues via Business Analytics</td>
<td>Address specific issues; prevention; tactical responses</td>
</tr>
<tr>
<td>Predict</td>
<td>Causal relationship identifications via Advanced Analytics</td>
<td>Strategic and tactical understanding of drivers and “ahas”</td>
</tr>
<tr>
<td>Optimize</td>
<td>A/B Testing to optimize experience and response rates</td>
<td>Incremental engagement and quick fixes/redirects.</td>
</tr>
<tr>
<td>Automation/Mining</td>
<td>Automated Decisioning products (Recommendation engines) &amp; discovery</td>
<td>Broader and more efficient delivery</td>
</tr>
</tbody>
</table>
### Define Strategy for the Business and create metrics to monitor progress against Strategic Goals...

<table>
<thead>
<tr>
<th></th>
<th>Strategy 1</th>
<th></th>
<th>Strategy 2</th>
<th></th>
<th>Strategy 3</th>
<th></th>
<th>Strategy 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>a.</td>
<td>KPI 1</td>
<td>b.</td>
<td>KPI 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>a.</td>
<td>KPI 1</td>
<td>b.</td>
<td>KPI 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Current Month</th>
<th>MoM (%)</th>
<th>YoY (%)</th>
</tr>
</thead>
</table>

- HELPS LEADERSHIP
- MONITOR BUSINESS &
- TAKE PROACTIVE ACTION/RAPID RESPONSE

Intended for Knowledge Sharing only
RECOMMENDATIONS ON HOW IT CAN BE MITIGATED

CUSTOMIZED DELIVERY...

<table>
<thead>
<tr>
<th>Stakeholders</th>
<th>Needs</th>
<th>Reports, Insights &amp; OF Drilldowns</th>
<th>Datamarts</th>
<th>Documentation</th>
<th>Communication &amp; Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executives</td>
<td>- Reports</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- High level drilldown</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Unified summary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- &quot;On the go&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marketing</td>
<td>- Campaign performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Infographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Deep dives</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales / RM</td>
<td>- Sales performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Prospecting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Competitive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Infographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>- Product performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Deep dive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Mining</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Research</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology / AE /</td>
<td>- Platform performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operations</td>
<td>- Deep dive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Forecasting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Real time alerting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FP &amp; A</td>
<td>- Consolidated Initiative readouts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(E2E)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Deduping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Drill downs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Forecasting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Top Priority of any team needs to be to hire a User Experience Design (UED) Strategist
## RECOMMENDATIONS ON HOW IT CAN BE MITIGATED

### DEEPER STAKEHOLDER INVOLVEMENT...

<table>
<thead>
<tr>
<th>Steps</th>
<th>Responsible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective</td>
<td>Analyst, <strong>Stakeholder</strong></td>
</tr>
<tr>
<td>Translation to Analytical Framework</td>
<td>Analyst, Researcher, Data Instrumentation, &amp; Data Manager, Developer, Data Scientist</td>
</tr>
<tr>
<td>Data Collection and Preparation</td>
<td>Analyst, Data Manager, Data Scientist</td>
</tr>
<tr>
<td>Analysis, Validation &amp; Verification</td>
<td>Analyst, Data Scientist, <strong>Stakeholder</strong> and SME, Researcher</td>
</tr>
<tr>
<td>Actionable insights and impact sizing</td>
<td>Analyst, <strong>Stakeholder</strong>, Leader</td>
</tr>
<tr>
<td>A/B Testing</td>
<td>Analyst, A/B Testing, <strong>Stakeholder</strong>, Developer</td>
</tr>
<tr>
<td>Rollouts</td>
<td><strong>Stakeholder</strong>, Leadership &amp; Executives</td>
</tr>
</tbody>
</table>
SUMMARY OF TAKEAWAYS

• “Know” that not all Analytics is supposed to be actionable.

• “Must have” User Experience Design (UED) Strategist for the Analytics practice

• “Ensure” Deeper Stakeholder involvement in Analytics development & Test & Learn approach must

• “Develop” Outcome Focused Approach for Analytics

• “Prepare” for ever more increasing ask for analytics and related actionability issues
THANK YOU
14th Text Analytics Summit

SAVE THE DATE!
June 15th & 16th, 2015 - New Yorker Hotel
New York City

- Drive business value with text analytics: identify the goals of text analytics and how to achieve them
- Integrate unstructured and structured data: get deeper business insights and improve decision-making in your organization
- Customer engagement and understanding: improve CEM through sentiment analysis and social listening techniques
- Create more effective marketing campaigns: use text analytics to improve segmentation and targeting

Event Contact: Brian Parke
Tel: +1 (201) 234-4755
Email: brianparke@fc-bi.com
**Live Q&A**
**Participants**

Howard Lyeth  
Stephen Scarr  
Ramkumar Ravichandran  
Meta Brown

L.L.Bean eContext VISA
THANK YOU!