Harnessing Big Data for Social Media Intelligence: South African Case Study

1st Annual Society of Competitive Intelligence Professionals (SCIP – SA) Organisational Intelligence and Knowledge 11-13 November 2013

Dr Herman J. van Niekerk
Suritec (Pty) Ltd
• The Challenge brought about by Big Data ....
• Overview of Social Media Strategy
• Twitter – The Concept and Applications
• How Research was done...
• Case Studies – Examples of Social Media Tools in Practice
• Social Media Examples – Networks, Butchers and Bakers.
• Conclusion
• Questions
The Challenge of Big Data and Social Media

• What is Big Data and why is it a challenge for Social Media?

• What are the biggest challenges for Social Media Intelligence?
  o To obtain insights beyond the normal functionalities offered
  o To determine social networks and actors in these networks
  o To establish what are the key themes being discussed and the role hash tags play in social media.
Internet of People
1.11B People on Facebook Mar 2013

Internet of Information
30T Web Pages in Google Index

Internet of Places
3B Foursquare Check-ins, Jan 2013

Source: Gartner 2013

BIG DATA
CHALLENGE
Everything connected to everything
Internet of Everything: Big Data and Social Intelligence

Asset
- Thing
  - Big data analysis

Examples:
- Optimizing single industrial machine
- Monetizing pay-as-you-go usage

Network of Assets
- Thing
  - Big data analysis

- Thing
  - Optimizing a network of lights to minimize electricity
  - Optimizing network of traffic lights to improve traffic and safety

Everything
- System
- Place
- People
  - Big data analysis

- Optimizing hospital for maximized use of staff, equipment, consumables, beds, and rooms

Gartner
The Challenge of Big Data and Social Media

Big Data = Transactions + Interactions + Observations

Petabytes

Terabytes

Gigabytes

Megabytes

WEB

Sensors / RFID / Devices
Mobile Web
User Click Stream
Web logs
A/B testing
Dynamic Pricing
CRM
Segmentation
Offer details
Customer Touches
Support Contacts
ERP
Purchase detail
Purchase record
Support Contacts

Increasing Data Variety and Complexity

Source: Contents of above graphic created in partnership with Teradata, Inc.
What is Big Data?

Big data is the term for a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications.

Big Data is about building net-new analytic applications based on new types of data that (an organization) wasn't previously tracking [McJannet, VP of marketing for Hortonworks]

3V Model -- High
The 3V Model of Big Data

- **Volume**: Biggest challenge and biggest opportunity. Helps organisations understand people and social actors. Allocate resources more effectively. Relational databases are not scalable for data of this magnitude.

- **Velocity**: Rate of data inflow in the organisation exceeds the capacity of IT systems.

- **Variety**: Data types becoming increasingly diverse. Traditional data vs. other big data like photographs, audio & video, 3D models, simulations, geo-locations etc. Unstructured and uncategorised. Difficult to process with traditional computing techniques.
Big Data Technology Requirements

- Traditional data bases insufficient – enter in-memory databases such as Hadoop and Cassandra.
- Visualization tools are required.
- Deeper Analytical skills needed.
- Development of new algorithms.
- Challenge of merging structured and unstructured data.
- Difficulty in successfully and quickly extracting social networks.
Social Media Landscape

A challenge to extract social intelligence....
Social Media

- Social media introduces substantial and pervasive changes to communication and the relationships between organizations, communities, and individuals.
- This presents an enormous challenge for organizations on how to manage these relationships and how and what to communicate to your audience.
- Experts claim that effective relationship building is a crucial ingredient in business success, perhaps even the critical success factor. Despite this its importance is often under-acknowledged.
Twitter as a Social Information Source

- Fastest growing platform in South Africa. Twitter has grown by 129% from 2.4 million users in 2012, to 5.5 million in 2013.
- 54-million tweets a month, and the vast majority (85%) are sent from mobile devices.
- Only really open micro-blogging platform available – other platforms are closed to groups only.
- Limitations: 140 characters and at times content is not standard English
- Difficult to determine sentiment – sarcasm cannot be identified by computational linguistics.
- Hash tags play an important role.
- Social networks have huge potential if technology is capable of extracting networks.
- Twitter data is only accessible for seven days – individual tweets are kept.
WHAT'S THE BIG CAMPAIGN IDEA?

WE'RE GOING DIGITAL

FACEBOOK

YOUTUBE

A MOBILE APP

PINTEREST

WHAT ARE WE GOING TO DO IN ALL THOSE CHANNELS?

I DUNNO, WE'LL FIGURE IT OUT LATER

© marketoonist.com
Using Social Intelligence to Create a Competitive Advantage

- Build brand awareness / commitment / behaviours that strengthen this.
- Generate leads / Increase sales
- Establish thought leadership
- Build relationships with customers
- Engage customers as brand evangelists / drivers of participation
- Get customer input for product and/or service improvement

Key Drivers in Social Intelligence which will give competitive advantage

- What are people talking about?
- What are emerging topics and trends?
- Who shares my content online?
- Who is controlling my brand’s conversations?
- What are the dynamics around my network?

SM strategy to guide CI efforts
Social Media Intelligence Roadmap

**Step 1:** Define the companies, individuals, products, services, industries and markets to research.

**Step 2:** Identify the social media resources with which your competitors are engaged.

**Step 3:** Select a social media tool to start searching and tracking key words, and hash tags.

**Step 4:** Analyze results and identify trends, opinions and key players (butchers and ambassadors).

**Step 5:** Integrate research results with other, more traditional forms of CI. Look for outliers, gaps and additional questions.

**Step 6:** Disseminate your findings as appropriate.
Social Media Presence

"Twitter is not a technology but a conversation, and it's happening with or without you" - Charlene Li

Online conversations about you take place even if your organization is not around to hear it!
What is the Problem with Social Media Analytics?

• High adoption rate of SM analytics, but relative low maturity level of programmes.
• Technology backbone based on Big Data – and that is not yet mature.
• Focus on volume metrics, such as clicks, tweets, re-tweets which metrics don’t necessarily generate useful insights.
• Need to move away from real-time monitoring and volume metrics to intelligence analysis.
• Followers gained / lost and positive/negative sentiment: Need to move to a more granular level of analysis to understand the reasons behind these movements.
Leading Five Banks in South Africa:

- What is their Social Presence and Behaviour?
- What Intelligence can we obtain from that?
How Research was done....

• **Whisperer Social Media tool:**
  o An Analytical tool – not a reporting tool.

• **Data Collection**
  • Up to 200 000 tweets harvested and analysed at any given time for this project.
  • Data range: First two weeks of October 2013 as a comparison.
  • Overall view for October 2013

**Data Analysis**
• Using Whisperer.mobi which is a social intelligence tool under development.
• Extensive algorithms built into the search queries to provide social analytics and to compare with competitors.
Social Voltage is an indicator of how widespread a tweeter’s influence is from a messaging & re-tweeting perspective. We rank users from 1 to 3 million.

Social Energy reveals how effectively a tweeter’s messaging impacts on its network growth. Users are also ranked from 1 to 3 million.

Mood is an overall value of positive and negative word application by users based on an extensive library of emotional terms. The metric displayed vacillates between 5 and -5.
Social Intelligence – New Terminology

• **Voltage**: Influence of your tweets (Reach)
• **Social Energy**: Impact of your tweets on your Network (Depth)
• **Butchers and Bakers**: Tweeters who act as bridges or distractors.

This allows an analyst to Profile and Track the powerful Tweeters or Bloggers in your social network to determine their influence (voltage) on how far the message travels in the network.

Social energy relates to what happens to your network – do your Followers grow or contract?

In other words, a Tweeter may have a large following, but the message is not re-tweeted while a Tweeter with a smaller following has more voltage / power when the message is re-tweeted several times.

If your network then grows – your social energy increases.
Case Study
South African Bank’s Social Presence

Text based records – Close to 200 000 Tweets
- Tweets harvested during 27 September – 11 October 2013
- Complex and scientific algorithms built into the software
- New Concepts – Social Voltage and Energy
- Network Actors – Butchers and Bakers
- Networks extracted and visualised using Visual Analytics.
Step 1: Create a SCI Project
Project Detail

- Including Twitter Handles, relevant # tags and key words.
- Ability to edit project as it progresses.
• Specific period of analysis
  Top Tweets during period – relates to a competition by Standard Bank – ultimatebraaimaster
<table>
<thead>
<tr>
<th></th>
<th>HASHTAG</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Nedbank#digitaledge2013</td>
<td>291</td>
</tr>
<tr>
<td>2</td>
<td>Absa#humanspirit</td>
<td>48</td>
</tr>
<tr>
<td>3</td>
<td>Capitec#fnb</td>
<td>37</td>
</tr>
<tr>
<td>4</td>
<td>Capitec#teamabsa2014</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>Capitec#askwhy</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>Capitec#cwqauction</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>Capitec#lyrictheatre</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>Capitec#nothingformahala</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>Capitec#fashion</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>Capitec#safw</td>
<td>13</td>
</tr>
<tr>
<td>11</td>
<td>Capitec#keyonateam</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>Capitec#capitecbvc</td>
<td>8</td>
</tr>
<tr>
<td>13</td>
<td>Capitec#absaapp</td>
<td>8</td>
</tr>
<tr>
<td>14</td>
<td>Capitec#nedbank</td>
<td>6</td>
</tr>
<tr>
<td>15</td>
<td>Capitec#keyona</td>
<td>6</td>
</tr>
</tbody>
</table>

27 September – 5 October 2013

<table>
<thead>
<tr>
<th></th>
<th>HASHTAG</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Standard Bank#humanspirit</td>
<td>897</td>
</tr>
<tr>
<td>2</td>
<td>Standard Bank#fnb</td>
<td>457</td>
</tr>
<tr>
<td>3</td>
<td>Standard Bank#digitaledge2013</td>
<td>315</td>
</tr>
<tr>
<td>4</td>
<td>Standard Bank#askwhy</td>
<td>204</td>
</tr>
<tr>
<td>5</td>
<td>Standard Bank#cwqauction</td>
<td>84</td>
</tr>
<tr>
<td>6</td>
<td>Standard Bank#boks</td>
<td>70</td>
</tr>
<tr>
<td>7</td>
<td>Standard Bank#absacurriecup</td>
<td>56</td>
</tr>
<tr>
<td>8</td>
<td>Standard Bank#actscholarships</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>Standard Bank#absa</td>
<td>45</td>
</tr>
<tr>
<td>10</td>
<td>Standard Bank#ultimatebraaimaster</td>
<td>38</td>
</tr>
<tr>
<td>11</td>
<td>Standard Bank#springboks</td>
<td>37</td>
</tr>
<tr>
<td>12</td>
<td>Standard Bank#ssrugby</td>
<td>35</td>
</tr>
<tr>
<td>13</td>
<td>Standard Bank#nedbank</td>
<td>33</td>
</tr>
<tr>
<td>14</td>
<td>Standard Bank#bokke</td>
<td>31</td>
</tr>
<tr>
<td>15</td>
<td>Standard Bank#votesmallbusiness</td>
<td>31</td>
</tr>
</tbody>
</table>
Conference Digitialedge2013 – linked to Nedbank
Retweeted by BizCommunity

## Tags for Project: Banks Social Intel

<table>
<thead>
<tr>
<th>#</th>
<th>HASHTAG</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>#digitialedge2013</td>
<td>281</td>
</tr>
<tr>
<td>2</td>
<td>#tunnelspirit</td>
<td>48</td>
</tr>
<tr>
<td>3</td>
<td>#teams2014</td>
<td>37</td>
</tr>
<tr>
<td>4</td>
<td>#sh7wh7</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>#ownsag</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>#visionsght</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>#visionsght§</td>
<td>13</td>
</tr>
<tr>
<td>8</td>
<td>#fashion</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>#savvy</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>#swork</td>
<td>11</td>
</tr>
<tr>
<td>11</td>
<td>#sworkteam</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>#sworkspir</td>
<td>8</td>
</tr>
<tr>
<td>13</td>
<td>#sworksh</td>
<td>8</td>
</tr>
<tr>
<td>14</td>
<td>#sworkvb</td>
<td>8</td>
</tr>
<tr>
<td>15</td>
<td>#sworkar</td>
<td>8</td>
</tr>
</tbody>
</table>
Bizcommunity.com
@Bizcommunity
Daily B2B news for 26 industries in Africa! Sign up for press offices; subscribe to newsletters. @Biz_Marketing. Also @Biz_Africa/@Biz_Retail/@Biz_Lounge
South Africa

Voltage 847,216 out of 2,741,477
Mood Unchanged
Social Energy Unchanged

Tweets 4270 Retweeted 214 Mentioned 1073 Followers 17917 Followings 5669 #Tags
Retroviral – only **1120 Followers** but has the highest social influence in this project at that point in time. 35 Retweets by Ambassadors such as BrandsEye Difference with:

MasterCard – nearly **70 000 Followers** but not the same influence (Voltage).
Re-Tweet count of 35 resulted in a very high social voltage
Retweets for Retroviral by Power Tweeters

<table>
<thead>
<tr>
<th>USER</th>
<th>CREATED</th>
<th>LOCATION</th>
<th>TWEET</th>
</tr>
</thead>
<tbody>
<tr>
<td>@JenineDeklerk</td>
<td>2013-10-03 11:49</td>
<td>Johannesburg</td>
<td>Jenine de Klerk @JenineDeKlerk RT @retroviral: Using soccer to connect with people through their passion points - loving the Ke Yonal story @Nedbank #DigitalEdge2013</td>
</tr>
<tr>
<td>@cleo_gold</td>
<td>2013-10-03 11:25</td>
<td></td>
<td>maMiya @cleo_gold RT @retroviral: Using soccer to connect with people through their passion points - loving the Ke Yonal story @Nedbank #DigitalEdge2013</td>
</tr>
<tr>
<td>@brandseye</td>
<td>2013-10-03 11:11</td>
<td>South Africa</td>
<td>BrandsEye @brandseye RT @retroviral: Using soccer to connect with people through their passion points - loving the Ke Yonal story @Nedbank #DigitalEdge2013</td>
</tr>
<tr>
<td>@KallasB</td>
<td>2013-10-03 11:08</td>
<td>South Africa</td>
<td>Kallas @KallasB RT @retroviral: Using soccer to connect with people through their passion points - loving the Ke Yonal story @Nedbank #DigitalEdge2013</td>
</tr>
<tr>
<td>@mikesharman</td>
<td>2013-10-01 07:38</td>
<td>Johannesburg</td>
<td>Mike Sharman @mikesharman RT @retroviral: #WhyWeDoDigital: Our latest labour of love - the BraaiPhone <a href="http://ft.co/Qapfc1Up5">http://ft.co/Qapfc1Up5</a></td>
</tr>
</tbody>
</table>
• ABSA’s high social presence – but also driven by a high mention count related to bad internet service. Who are these Tweeters and what is their influence?
ABSA bank – 51 mentions. Some about the rugby, but many about poor internet service – see next slide
Example of Negative tweets which increased the Mentions

<table>
<thead>
<tr>
<th>USER</th>
<th>CREATED</th>
<th>LOCATION</th>
<th>TWEET</th>
</tr>
</thead>
<tbody>
<tr>
<td>@SABankingInfo</td>
<td>2013-10-06 09:15</td>
<td>South Africa</td>
<td>RT @EtienneM: Don't know who changed at @Absa in the last 3 months but it seems their internet division has gone on holiday! #offline</td>
</tr>
<tr>
<td>@SABankingInfo</td>
<td>2013-10-06 09:15</td>
<td>South Africa</td>
<td>SA Banking Info RT @EtienneM: Don't know who changed at @Absa in the last 3 months but it seems their internet division has gone on holiday! #offline</td>
</tr>
<tr>
<td>@Gail_Mail</td>
<td>2013-10-06 09:09</td>
<td>Johannesburg, South Africa</td>
<td>Love the double page @ABSA advert in the @SundayTimesZA “People across South Africa have been telling us what to prosper means to them!”</td>
</tr>
<tr>
<td>@EtienneM</td>
<td>2013-10-06 09:05</td>
<td>South Africa</td>
<td>Don't know who changed at @Absa in the last 3 months but it seems their internet division has gone on holiday! #offline</td>
</tr>
</tbody>
</table>
Negative Mention by one Tweeter then re-tweeted by a reputable institution resulted in a high mention count.
Hash Tags Example

Social Situation a week later.....

• #hash tags have changed significantly
10 October 2013

- Less activity than previous week.
- #knowledgebreakfast connected to Nedbank
Individual tweeting about Nedbank and #knowledgebreakfast
Very high social voltage and positive mood rating (35%) – positive influence on discussion
Not so many Followers, but high re-tweet and mention count.
Ideal Candidate to seed communication – act as Brand ambassador (Baker) for Nedbank
Nerina Visser’s message broadcast into the Banking Twitter network

NB: Only a sub-division of the entire tweets was used in this instance
Groupings around word scores
Nedbank has a wider / more dispersed impact
4-17 October 2013
ABSA’s #humanspirit had most tweets
Majority of Tweets are linked to the Currie Cup rugby season – ABSA is the main sponsor.
Same Period: 4-17 October 2013
Nedbank has by far the most mentions.
Reasons: Wide ranging topics, reference to new Wifi hotspots, poor security at Musgrave Centre, positive reference to Nedbanks’s Business Accelerator
Comparison Functionality:

FNB has consistently the highest social voltage of all the major banks.
<table>
<thead>
<tr>
<th>Top Mentions</th>
<th>Top # hash tags</th>
<th>Top Retweets</th>
</tr>
</thead>
<tbody>
<tr>
<td>@Nedbank - 9842</td>
<td>#fnb - 3381</td>
<td>Roar of crowd – 64 (rugby)</td>
</tr>
<tr>
<td>@absa - 9295</td>
<td>#humanspirit - 2236</td>
<td>Kolbe – player of the month - 46</td>
</tr>
<tr>
<td>@fnbsa - 3794</td>
<td>#nedbank - 893</td>
<td>Funny – scared of the chomp – FNB gator day - 45</td>
</tr>
</tbody>
</table>

Dashboard Summary:

- Date Range: 01/09/2013 00:00 to 29/10/2013 00:00
- Top Mentions: @Nedbank, @absa, @fnbsa
- Top # hash tags: #fnb, #humanspirit, #nedbank
- Top Retweets: Roar of crowd – 64 (rugby), Kolbe – player of the month - 46, Funny – scared of the chomp – FNB gator day - 45
Initial Observations

• The importance of **hash tags #** in the digital space
  
  o  Banks like FNB, ABSA and Capitec drive their communication around # hash tags of which they have taken ownership – part of the social DNA of the company.
  
  o  Nedbank links its social communication to specific events such as #knowledgebreakfast and #digitaledge and Standard Bank with #UltimateBraaiMaster.
  
  o  Nedbank has many more mentions than any other bank.
  
  o  ABSA’s sponsorship of the SA Rugby and Currie cup rugby is a highly successful social media campaign. The #humanspirit is a more successful than their brand name!
Conclusions

• The **importance of hash tags #** in the digital space.....
  - Companies / Organizations need to take ownership of a hash tag, e.g. #humanspirit, #fnb, #askwhy.
  - # tags need to be unique and clearly identifiable with the specific company – must become part of the company’s DNA.
  - # tags could also be linked to specific campaigns and need time to become part of the company’s social domain and presence.
  - Hash tags driving social media campaigns such as #digitaledge and #ultimatemaster

• **Authentic stories** – re-tweeted and mentioned hugely increase your social energy – e.g. Retroviral story about a soccer tournament.
Conclusions

• Social Data is dynamic and changes on a daily basis. Analysts must immerse themselves in the data on a daily basis but also have the ability to analyse trends over a longer period.
• Companies must not only focus on real-time monitoring but also on social intelligence analysis.
• Social Media tools provide practitioners / digital managers with opportunities to extract social intelligence from information sources which were previously difficult or impossible to analyse.
Conclusions

• Opens up new possibilities for practical application and theoretical research. Many possible applications:
  • Measure the effectiveness and exposure of Campaigns / Sponsorships
  • Monitor customer complaints and its impact.
  • Analysis and comparison of competitors social presence, activities and indication of the success of their campaigns.
  • Realign own social strategy to ensure relevance.
  • Targeted communications at brand ambassadors.
The Future......

• Cross-brand intelligence and Network Voltage – Establish the Power and reach of your network—essentially providing an improved metric on measuring the Networth (value of your network).
• Understanding the contextual potency of messages via entity extraction, tracking of actual participants and their affinity towards other brands.
• Social data intelligence opens the door to previously unconsidered opportunities.
• Results in improved decision-making regarding brand reputation and customer engagement
• Shaping of your digital message to maximize the effectiveness of your spend and ROI.
Questions?

All the answers we ever get are responses to questions

- Neil Postman

SURITEC

Unit F02A, Twickenham
Brookside
11 Lansdowne Road
Claremont

Tel: +27 (0)21 683-3624
Fax: +27 (0)21 683-3626

Web: http://www.suritec.co.za

Email: herman@suritec.co.za