RSA Conference2016

San Francisco | February 29 – March 4 | Moscone Center

Connect to Protect

SESSION ID: GRC-T10R

# The Newest Element of Risk Metrics: Social Media

**Ian Amit**

Vice President
ZeroFOX inc.
@iiamit

# Basic Motivation - hottest/easiest vector!

- "… in previous years, we saw phishing messages come and go and reported that the overall effectiveness of phishing campaigns was between 10 and 20%. This year, we noted that some of these stats went higher, with **23% of recipients now opening phishing messages and 11% clicking on attachments**. Some stats were lower, though, with a slight decline in users actually going to phishing sites and giving up passwords."
- "For two years, more than two-thirds of incidents that comprise the Cyber-Espionage pattern have featured phishing."

2015 DBIR

# Why do I want this?

- Everyone is on social media. Whether you tell your employees that they can or can't.
- Organizations find that they conduct business communications over social media.
- The gap between online and physical is very narrow when factoring in social media.
- Attackers target organizations through the path of least resistance. Social media is the easiest as:
  - There are less (if any) controls over it.
  - It provides a more personalized "experience" for the user (unlike email).
  - It is more interactive and attackers can quickly adapt their approach.
- It is easy to impersonate someone on social media and impact the organization.

# Who is potentially affected?

- Are you engaged in a “controversial” practice?

Financial Services

DIB

Healthcare

Pharma

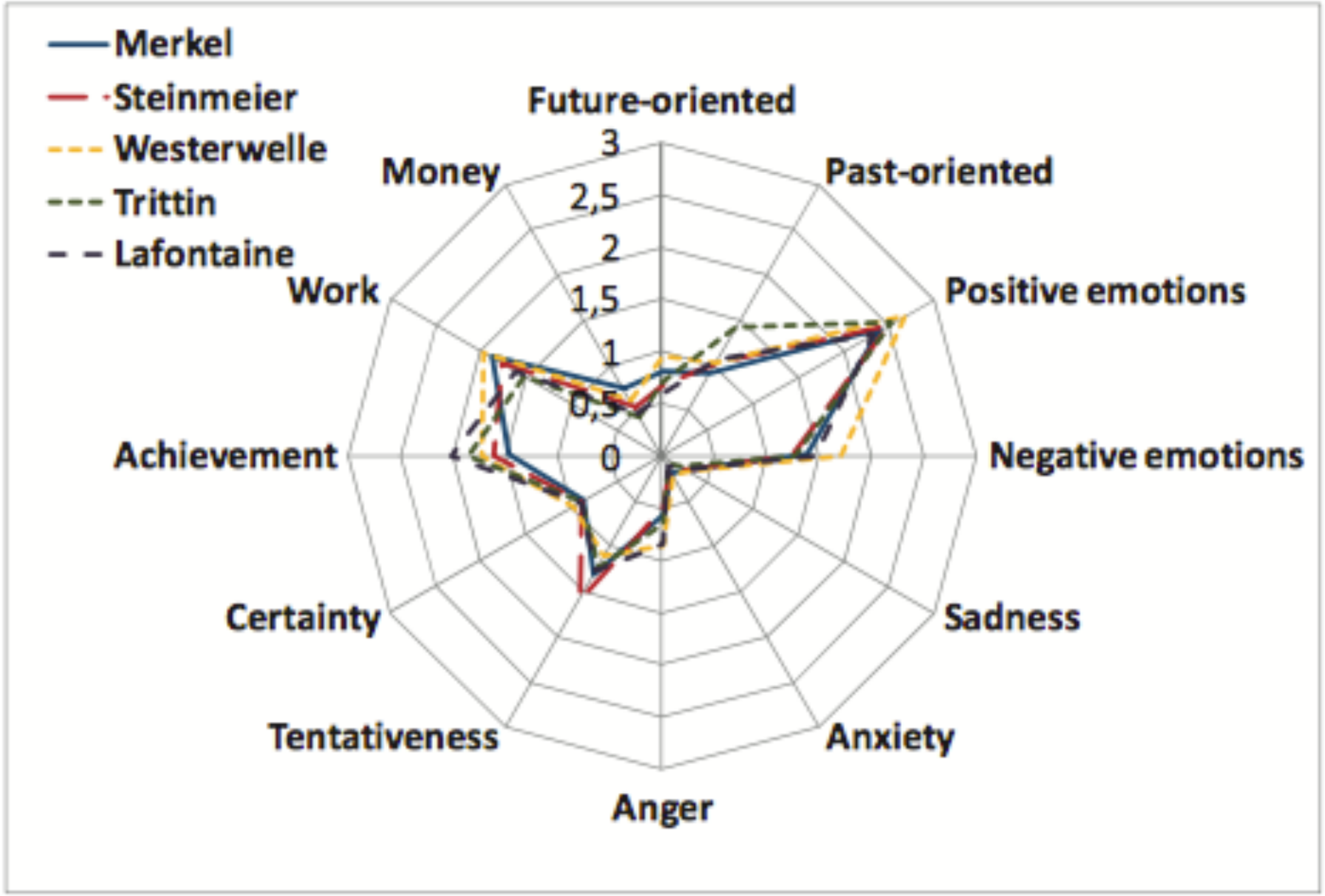
Agribusiness

LEA

Energy

# Can I really predict risk based on SM activity?

- Sentiment analysis and the German elections
- Predicting Elections with Twitter:  What 140 Characters Reveal about Political Sentiment
- “Twitter can be seen as a valid real-time indicator of political sentiment.”
- http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewFile/1441/1852


Merkel
Steinmeier
Westerwelle
Trittin
Lafontaine
Future-oriented
Past-oriented
Positive emotions
Negative emotions
Sadness
Anxiety
Anger
Tentativeness
Certainty
Achievement
Work
Money
3
2,5
2
1,5
1
0,5
0

# Coming up with a solution

Framework for measuring the risk of a person/organization's social media activity

# What is it that we need to address?

- A framework for you to look at how inflammatory or “risky” individuals in your organization are.  Individuals:
  - like executives,
  - technical contractors & employees who, you know, might have admin access, and/or
  - employees susceptible to other risk categories like Fraud, Reputation, and Strategic risk.

# What will I get out of this?

- The ability to build a scorecard allowing you to rank employee risk.
- The ability to drill down into the SM behaviors that contribute to risk
- And subsequently lower a risk profile through applying controls to select elements identified through the process.
- The ability to enhance OSINT functions with SM-focused functions

# Basic concepts behind the model

We utilized the GQM approach:

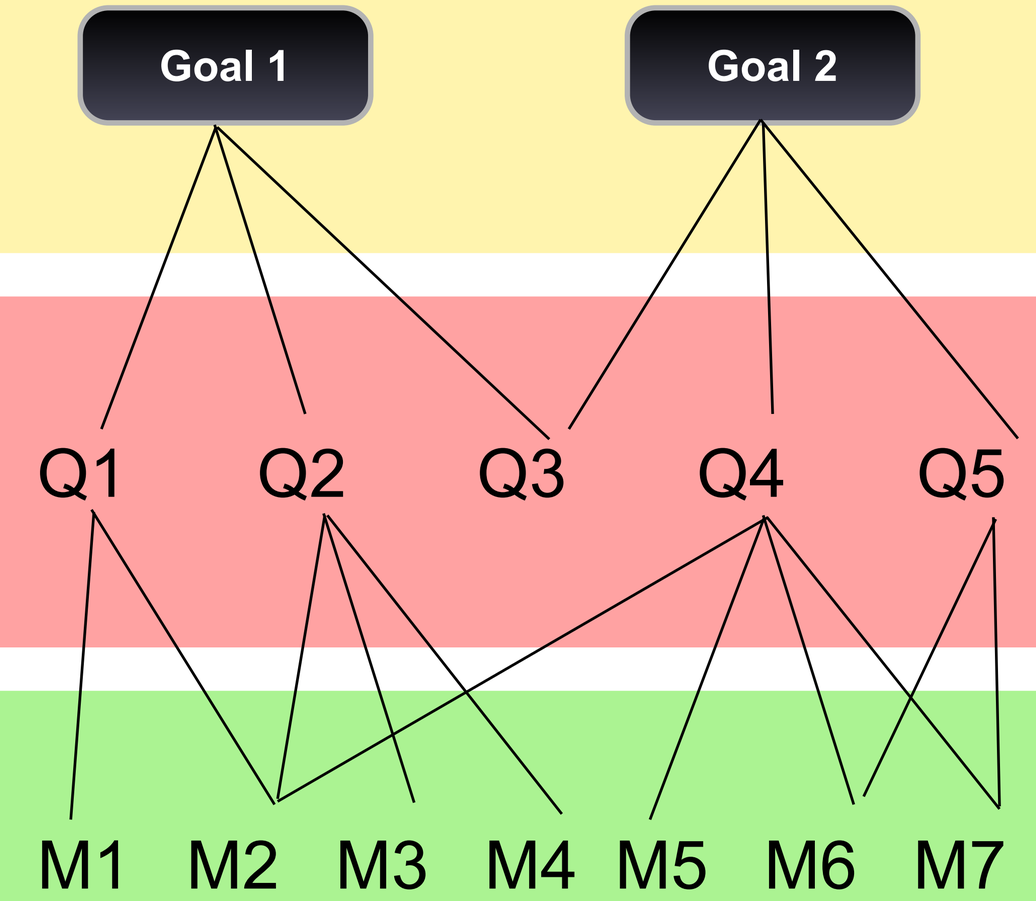
- Conceptual level (**goal**)
  - Goals defined for an object for a variety of reasons, with respect to various models, from various points of view.
- Operational level (**question**)
  - Questions are used to define models of the object of study and then focuses on that object to characterize the assessment or achievement of a specific goal.
- Quantitative level (**metric**)
  - Metrics, based on the models, is associated with every question in order to answer it in a measurable way.

# GQM

Goals establish what we want to accomplish.

Questions help us understand how to meet the goal. They address context.

Metrics identify the measurements that are needed to answer the questions.




Victor Basili

# Our GQM data

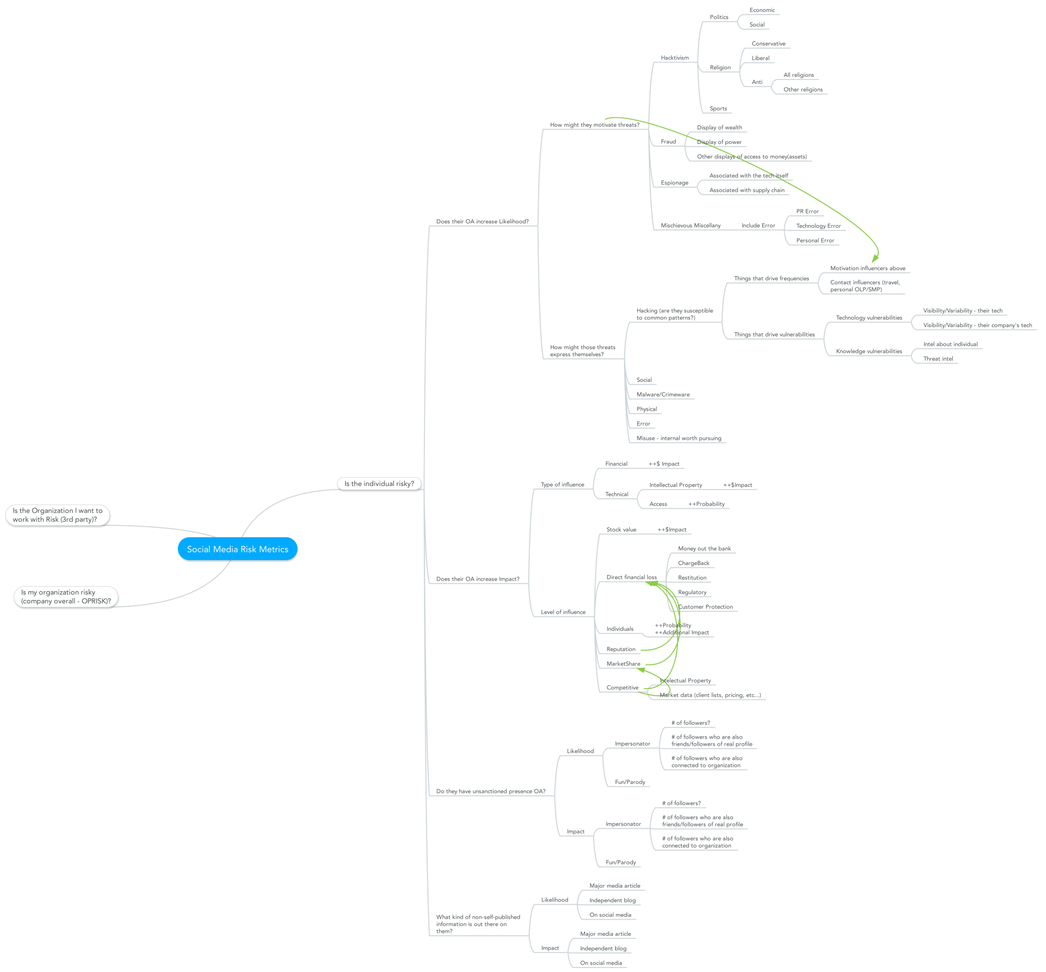
- Goal: Provide a social media risk scorecard for a person/ organization.
- Questions: How would one's OA affect the likelihood of a threat? How would one's OA affects the impact of a threat, and the areas of impact? How does unsanctioned presence of someone affect said threats?
- Metrics: Provide a qualitative* approach to measuring the overall risk, as well as specific aspects of the social media presence.

*And when we say qualitative we lie a little bit…

# More Goals

- 1. Provide a measurable way to quantify risk associated with online activity of the organization and it's employees.
- 2. Provide another measure for quantifying risk of working with 3rd parties and contractors.
- 3. Create a score for executives to measure their social media exposure (from an exec protection perspective, insider trading, etc...)
- 4. Create a score for measuring and comparing intra and extra industry social media risk ratings
- 5. Be able to quantify the effect of changing controls, processes and policies on the risk associated with social media.

# Mindmap (see external references)



# Developing the scoreboard

- Started with the basics, comparative measurements…
- Qualitative approach dictates trying to leave quantitative elements out (which we kind’a try to). So the compromise was to provide a fairly detailed breakdown of elements, and instead of measuring them on a scale, only indicate presence (1 or 0).
- Aggregation didn't work (per-se), Averaging would not take into account the full magnitude of the largest elements, MAX() would not factor in contribution from smaller ones. We have to provide more accurate weights…

# Scoring Approach

- Ended up with providing a weighting system for the major elements and their importance to the organization (context?!).
- Given X points to distribute between Y elements. Weight = Y'/X where Y' is the number of points given to each element.
- Sum(Y'…Y'')=1
- Apply weighting to the scorecard to get weighted risk score. (where weights are appropriate for the organization's operational context).

# Weighting

| **Threat types** | Priority (50 points overall) | Weighted |
|---|---|---|
| Fraud | 10 | 0.2 |
| Espionage | 12 | 0.24 |
| Error | 3 | 0.06 |
| Hacktivism - political | 5 | 0.1 |
| Hacktivism - economical | 5 | 0.1 |
| Hacktivism - religious | 15 | 0.3 |
| Hacktivism - sports | 0 | 0 |
| | 50 | |
| **Vulnerability** | Rank (20 points overall) | Weighted |
| Technology | 8 | 0.4 |
| IP | 5 | 0.25 |
| Physical | 3 | 0.15 |
| Social | 4 | 0.2 |
| | 20 | |
| **Influence** | Rank (50 points overall) | Weighted |
| Financial | 5 | 0.1 |
| Stock | 18 | 0.36 |
| PR | 0 | 0 |
| Technology | 2 | 0.04 |
| Intellectual Property | 8 | 0.16 |
| Competitive | 10 | 0.2 |
| Other individuals in org | 2 | 0.04 |
| Market Share | 5 | 0.1 |
| | 50 | |

Sample

# Scorecard roadmap

- Current:
  - Breakdown of **Likelihood**, **Manifestation**, **Impact**, and an estimated factor of the number of online **threats** (by compounding monitored instances of threat actors with the medium used).
- Future:
  - Add breakdown to personal vs corporate risk, and further semantics such as exposure to malicious content, negative sentiment, information leaks, etc…

# What kind of data is needed?

- Organization size, and the size of the management.
- How many in the organization are monitored (and in management)
- Locality information (HQ, offices, size per location)
- Chatter - mentioning, conversing with, and talking about monitored assets. Also - assets posting/conversing/mentioning others.
- Impersonations - who is being impersonated? What's the intent (nefarious vs. parody)
- Sentiment analysis - in chatter, broken down to management vs company vs individuals (per location), and by distance from asset (1st, 2nd, 3rd degree)

# How can YOU do it?

- STEP 1:  Determine how your organization will support profiling.
  - None at all
  - None but publicly available information
  - Volitional
  - Enforced

# How can YOU do it?

- STEP 2:  Determine who you might want to protect
  - privileged IT users
  - executives/board members
  - marketing/PR people
  - sales

# How can YOU do it?

- STEP 3:  Determine where you will profile them
  - Social Media sites
  - Web sites
  - Blog comments

# How can YOU do it?

- STEP 4:  Collect <ALL> the data.
  - Extract / Tranform / Load
    - Scrape / Transform / Load
      - Analysis post scrape
      - Analysis in real-ish time (storm ftw)
        (twitter api -> spout -> bolt for processing)

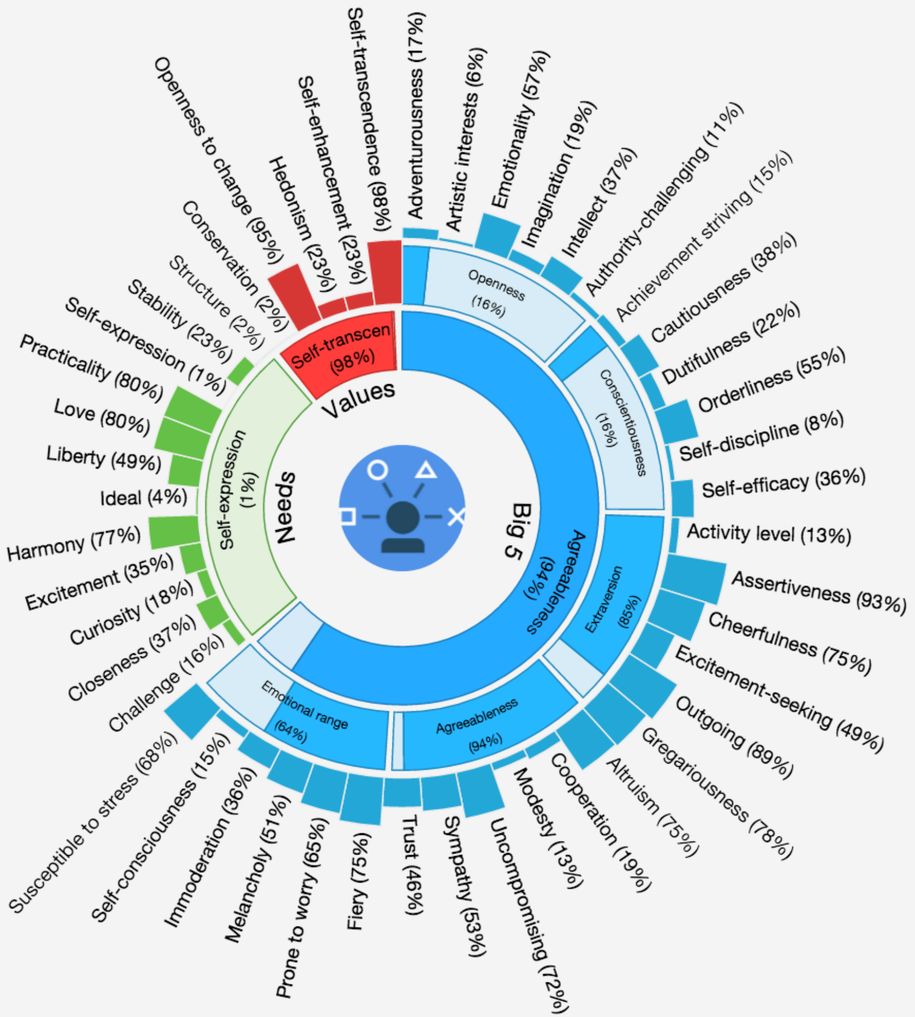
# How can YOU do it?

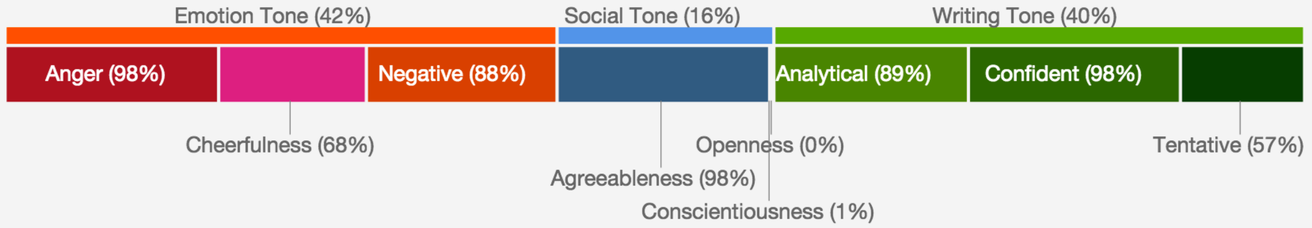
- STEP 5:  Store the data.
  - Whatever you want.

# How can YOU do it?

- STEP 6A:  Analysis
  - Amish hand crafted
    - Score comments regarding the factors that contribute to the  likelihood/ manifestation/impact elements of the model
    - Use freebie tools or do it yourself
    - tools like…
    - https://tone-analyzer-demo.mybluemix.net/
    - https://watson-pi-demo.mybluemix.net/
    - Score in our handy-dandy excel tool (or some variation thereof)

# How can YOU do it? (automated analysis)




You are social, boisterous and can be perceived as shortsighted.

You are unconcerned with art: you are less concerned with artistic or creative activities than most people who participated in our surveys. You are assertive: you tend to speak up and take charge of situations, and you are comfortable leading groups. And you are intermittent: you have a hard time sticking with difficult tasks for a long period of time.

Your choices are driven by a desire for efficiency.

You consider helping others to guide a large part of what you do: you think it is important to take care of the people around you. You are relatively unconcerned with tradition: you care more about making your own path than following what others have done.

**Compared to most people who participated in our surveys.*

# How can YOU do it?

- STEP 6B:  DIY BIG DATA MAGICS
  - Sentiment analysis (list from http://breakthroughanalysis.com/2012/01/08/what-are-the-most-powerful-open-source-sentiment-analysis-tools/)
  - Python NLTK (Natural Language Toolkit), http://www.nltk.org/, but see also http://text-processing.com/demo/sentiment/
  - – R, TM (text mining) module, http://cran.r-project.org/web/packages/tm/index.html, including tm.plugin.sentiment.
  - – RapidMiner, http://rapid-i.com/content/view/184/196/.
  - – GATE, te General Architecture for Text Engineering, http://gate.ac.uk/sentiment/.
  - Apache UIMA is the Unstructured Information Management Architecture, http://uima.apache.org/ — also sentiment classifiers for the WEKA data-mining workbench, http://www.cs.waikato.ac.nz/ml/weka/. See http://www.unal.edu.co/diracad/einternacional/Weka.pdf for one example.
  - Stanford NLP tools, http://www-nlp.stanford.edu/software/
  - LingPipe, (pseudo-open source). See http://alias-i.com/lingpipe/demos/tutorial/sentiment/read-me.html.

# How can YOU do it?

- STEP 7:  SCORECARD!
  - Output via model
  - Remember, it’s the factors of stress not necessarily a “risk score” that matters.
  - Ultimate goal is protect, be that via **technology** or **behavioral** controls.
  - Also applicable - legal, financial hedging, insurance, etc…

# Where can you get it?

The Society of Information Risk Analysts

http://www.societyinforisk.org

As well as on the SMRM site:

http://risk-metrics.com/

# Take-away

1. Check what is your current social media security policy (if you have one).
2. Do you have a current risk model that incorporates social media as part of it (attack surface / information leak / intelligence)
3. Measure your current social media risk posture for key individuals in your organization.

And then in 2-3 months - measure again to see whether any changes you have implemented in light of the initial measurement had the right impact.

# Thank you!

## Questions?

Ian Amit: @iiamit | ian@zerofox.com | http://www.iamit.org

# Resources

- Sentiment analysis and german elections: http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewFile/1441/1852
- Analyze tone of text: https://tone-analyzer-demo.mybluemix.net/
- Analyze personality based on text: https://watson-pi-demo.mybluemix.net/
- Sentiment analysis (list from http://breakthroughanalysis.com/2012/01/08/what-are-the-most-powerful-open-source-sentiment-analysis-tools/)
- Python NLTK (Natural Language Toolkit), http://www.nltk.org/, but see also http://text-processing.com/demo/sentiment/
  - R, TM (text mining) module, http://cran.r-project.org/web/packages/tm/index.html, including tm.plugin.sentiment.
  - RapidMiner, http://rapid-i.com/content/view/184/196/.
  - GATE, te General Architecture for Text Engineering, http://gate.ac.uk/sentiment/.
- Apache UIMA is the Unstructured Information Management Architecture, http://uima.apache.org/ — also sentiment classifiers for the WEKA data-mining workbench, http://www.cs.waikato.ac.nz/ml/weka/. See http://www.unal.edu.co/diracad/einternacional/Weka.pdf for one example.
- Stanford NLP tools, http://www-nlp.stanford.edu/software/
- LingPipe, (pseudo-open source). See http://alias-i.com/lingpipe/demos/tutorial/sentiment/read-me.html.