

## KDPaine & Partners

### Beyond sentiment analysis, can we automate trust measurement?

**IPRRC Miami**  
 March 10, 2011  
 Katie Delahaye Paine  
 CEO



Contact info: [kdaine@kdpaine.com](mailto:kdaine@kdpaine.com)  
[www.kdpaine.com](http://www.kdpaine.com)  
<http://kdpaine.blogspot.com>  
 Buy my book: Measure What Matters  
 (<http://amzn.to/hvCNvP>)  
 Read my blog: <http://kdpaine.blogspot.com> or subscribe  
 to The Measurement Standard:  
[www.themeasurementstandard.com](http://www.themeasurementstandard.com)  
 For a copy of this presentation go to:  
<http://www.kdpaine.com>  
 Follow me on Twitter: KDPaine  
 Friend me on Facebook: Katie Paine  
 Or call me at 1-603-682-0735

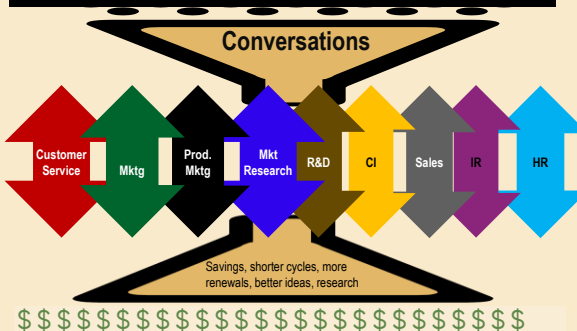
KDPaine & Partners

## The Premise

- ☒ Volume of conversations increasing
- ☒ Technology is improving
- ☒ Necessity for understanding the nature of the conversation is growing, & sentiment is only one element of what matters to an organization.
- ☒ RQ1: Can relationship elements accurately be measured using content analysis?
- ☒ RQ2: Can relationship elements accurately be measured using text mining and natural language processing?
- ☒ Problems anticipated:
  - ☒ Is it representative?
  - ☒ About 80% of all Twitter users have tweeted fewer than ten times.
  - ☒ 89% of the world does NOT use Facebook
  - ☒ 75% of "media" collected is either spam, unrelated to the topic, or generated by content farms

KDPaine & Partners

## It's not all about you, so get over it



KDPaine & Partners

## The Evolution of Social Media

2000-2004  
 Social Media/CGM  
 emerges as a technology  
 function

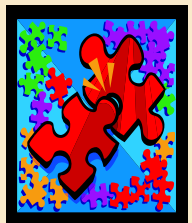


Social  
Media

2004-2010  
 Force-fitting Social Media  
 into Existing Structure



2010-2015  
 Integrating Social Media  
 into the Organizational  
 DNA



KDPaine & Partners

## The process

- Created coding instructions based on Grunig Relationship Theory
- Collected 2000 posts on 3 major Universities in February 2011
- Human coder trained on Relationship Concepts
- Human identified and coded 265 (13%) for relationship concepts

KDPaine & Partners

## University Data — Human coding

Colleges	Harvard	MIT	Stanford	Grand Total	
<b>Concepts</b>					
Commitment		4	4	8	3%
Communal Relationships	24	57	35	116	44%
Control Mutuality	6	13	2	21	8%
Exchange Relationships	1	4	3	8	3%
Satisfaction	8	19	6	33	12%
Trust	23	44	12	79	30%
<b>Grand Total</b>	<b>62</b>	<b>141</b>	<b>62</b>	<b>265</b>	

KDPaine & Partners

## Process Part 2: Machine Coding

- Used classifier and relationship modeling: frequency, importance, positional, directed positional
- Created taxonomy based on concepts, coding instructions and manual results and key words
  - Keeps promises, devoted, loyal — associated with entity
  - Reverse: Indifference, meaningless, disregard
- Machine coding
  - Positive precision: 0.001%
  - Negative precision: 0.00%
- Problems encountered:
  - Insufficient data that contains relationship concepts
  - Insufficient emotional content of data
  - Linguistic model requires positive or negative, valid/invalid constructs

KDPaine & Partners

## Machine Results Round 2

- Machine coding #2
  - 12,000 items from highly visible Corporation with emotional postings.
  - Continued taxonomy development
  - Communal and trust were most frequently found
  - Satisfaction and Trust are more accurately identified
  - Commitment, Control Mutuality & Exchange are more far more difficult

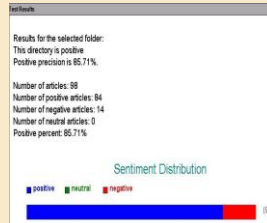
Total data	12000	
Valid Data	6778	56%
Items with Relationship Attributes	3153	47%
Commitment	295	4%
Communal Relationships	858	13%
Control Mutuality	285	4%
Exchange Relationships	199	3%
Satisfaction	365	5%
Trust	1151	17%
None	3625	53%

KDPaine & Partners

## The results so far

⊗ Preliminary Model for Communal Relationships validated against 858 items.

⊗ Challenges: Understand relationship between classifier and another word: i.e. "I'm falling in love AT Stanford" I'm falling in love WITH Stanford."



KDPaine &amp; Partners

## Where are we?

- ⊗ Have defined some words that define relationships but not enough.
- ⊗ Have first level definitions of words that relate to communal relationships
- ⊗ Can identify whether those words express positive or reverse communal relationships
- ⊗ Focusing on perfecting Trust taxonomy going forward

KDPaine &amp; Partners

## Conclusions

- ⊗ Linguistic tools sets can be used to pull concepts from data
- ⊗ Requires high quantity of valid data
- ⊗ Requires high quantity of emotional content data — i.e. something that people are passionate about
- ⊗ Requires lengthy iteration process
- ⊗ If you use a statistical model, taxonomy can't be generalized.  
Will only be specific to the organization studied.

KDPaine &amp; Partners

## Where do we go from here?

- ⊗ Need to model which words or phrases are more important
- ⊗ Must improve language understanding
- ⊗ Need greater quantity of human coded relationship data to build a statistical model that would validate relationship constructs in social media.

KDPaine &amp; Partners