



**VISIBLE™**  
Socializing the Enterprise

**On Empirical Sentiment Accuracy Bounds**  
**Shawn Rutledge, Chief Scientist**

# About Visible

**Founded:** 2005

**Headquarters:** Seattle

**Offices:** NY & London

**Employees:** 85

**Venture backed:** Investor Growth Capital  
Ignition Partners  
WPP  
Centurion Holdings  
In-Q-Tel

**Awards:** Top vendor (#22) in 2010  
Deloitte Technology Fast 500  
#164 Inc 500 list, 2010



# Visible's Sentiment Approach

## Algorithms

- State of the art
- Beyond overhyped NLP

## Features

- Deep experience
- Social NLP & Context

## Data

- Massive proprietary data

➤ A sentiment model based on years of labeling *social* data for *enterprises*.

➤  $10^7+$  labels,  $10^5+$  topics,  $10^2+$  enterprises.



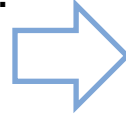




# Audit Findings, Large Financial Institution

## Double Blind, Multi-Reviewer Study:

1. Same posts labeled by both human labeling practice and automation.
2. At least two auditors grade each label. Blind to label source.



No statistically significant difference between human labeled and AI labeled sentiment

But...

Auditors agree with each other only 73% of the time [95%CI: 69%-77%].



# Key Audit Findings, Large Financial Institution

## Social Media Professionals Grading Human Annotations

Both auditors agree with label only **58%** of the time



At least one auditor agrees with label **91%** of the time



# True Across a Wide Variety of Problems

Multi-Reviewer 3<sup>rd</sup> party audits across a variety of Brands consistently show relatively low agreement rates.

**About 81% Inter-Annotator Agreement**  
[IQR: 78% - 83%]





