



**Gordon Anderson**

VP, Content

SIIA DataContent 2012

## InsideView: the company

- Software company, founded in 2005 in San Francisco & India.
- Mission: To turn large sets of structured & unstructured data into meaningful intelligence
- Audience: salespeople trying to understand customers and prospects
- Model: smart aggregation plus natural-language processing
- Work with all types of data – traditional editorial, regulatory filings, social media, Web-scraped, crowd-sourced information
- Goal is to use technology to make source content more relevant and more accurate

# InsideView: Aggregation model

- 30,000 sources drive content set, for both company & people info
  - Affect both current awareness insights and core database
- Entity triangulation
  - Back-end process to sort out differences between data sources
  - Essential to large-scale aggregation and analytics
- Social networks play a big part of our model
  - Vast scale, but incomplete, lumpy, and written in shorthand (e.g. Twitter)
  - The problem of “facts” in self-promotion
  - Best place for analytic “clues” (e.g. LinkedIn bio, or Facebook post)
- Unstructured data
  - Non-publishers publishing: Trend is toward free-form, unstructured content

# InsideView: Types of inferences made

- Companies
  - Corporate events
  - Firmographic info
  - Family tree
- People
  - Social network activity (tweets, blog posts, Facebook activity)
  - Background (e.g. employments not listed in official bio)
- Relationships
  - Between people (shared interests, common background)
  - Between companies

# Trends and Opportunities

- Increase in transparency
  - Regulatory filings
  - Corporate & personal self-promotion
  - Expansion of “real-time” analytical capacity
- Processing power & storage (Moore’s Law)
  - We currently analyze and parse 1 million news articles a day, up from 10K
  - We show tens of millions of relationships between people & companies
  - Our back-end technology costs are unchanged since startup, despite more than 10x expansion in user base and 100x expansion in data volume
- Audience receptivity
  - As more becomes available, more will be expected
  - Leveraging social media for intelligence