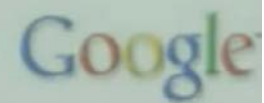



The Structured Search Engine

Andrew Hogue
Google

January 19, 2011



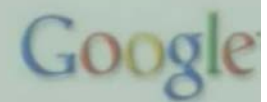
Google



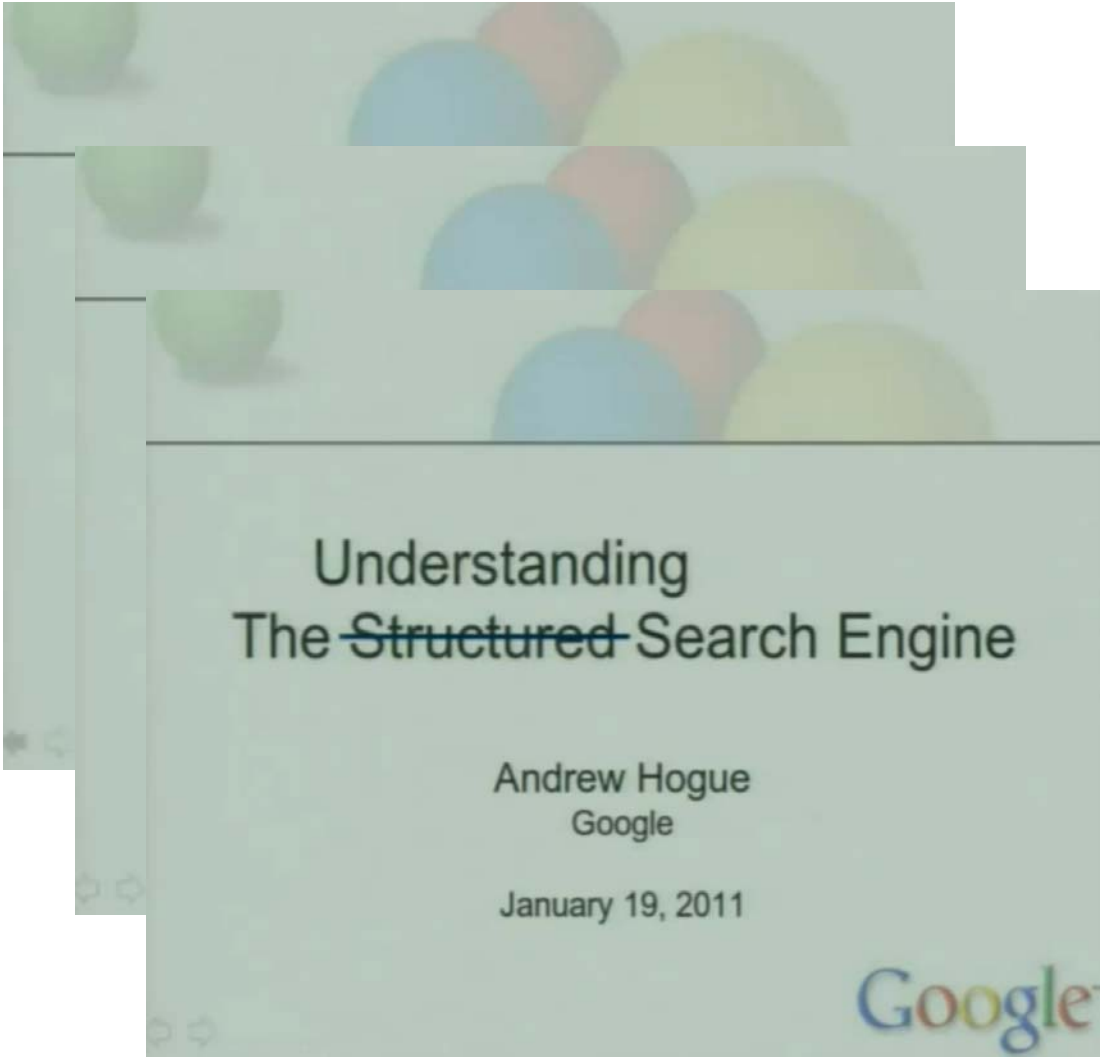
Semantic The ~~Structured~~ Search Engine

Andrew Hogue
Google

January 19, 2011



Google

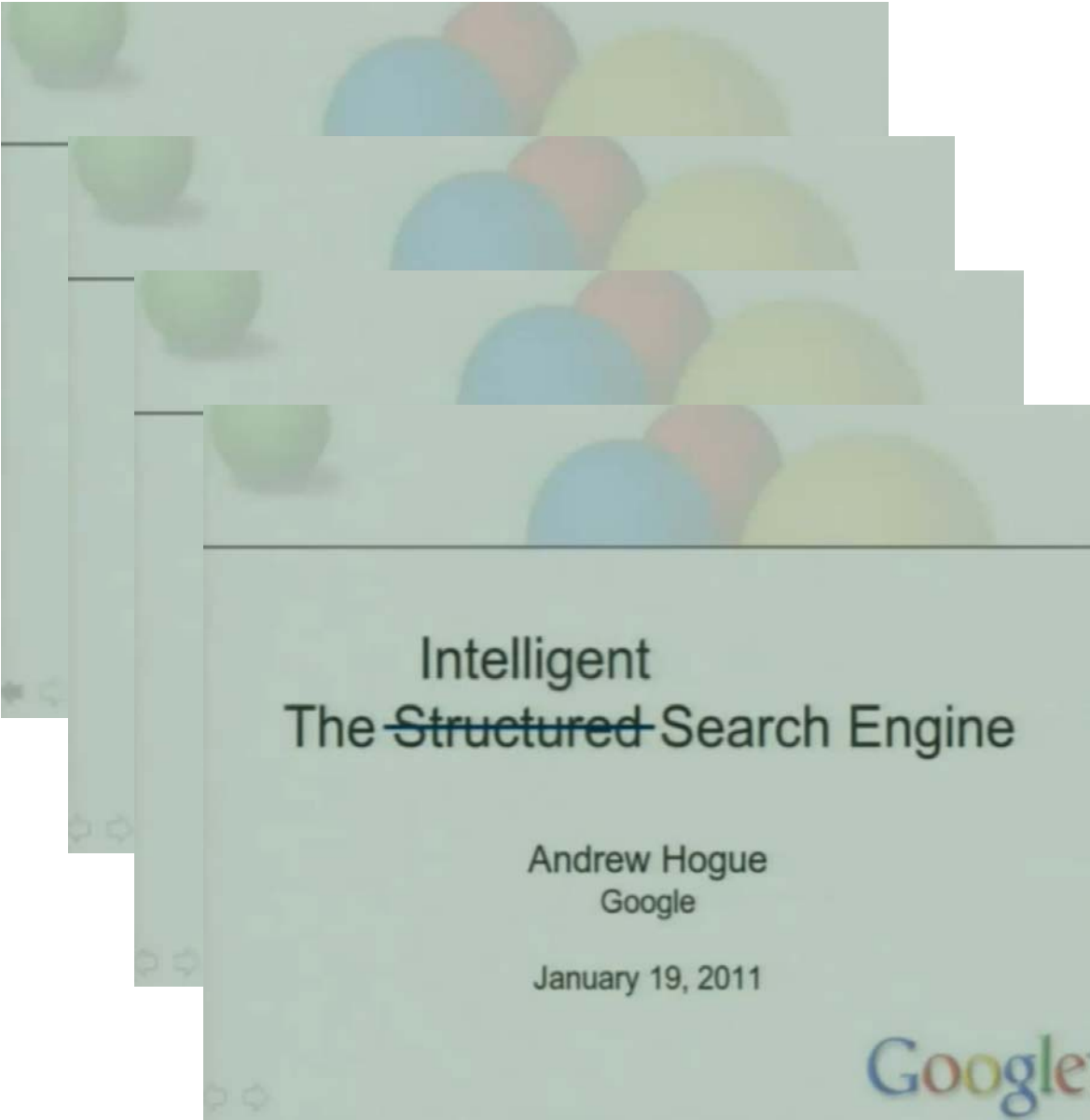


Understanding The ~~Structured~~ Search Engine

Andrew Hogue
Google

January 19, 2011

Google

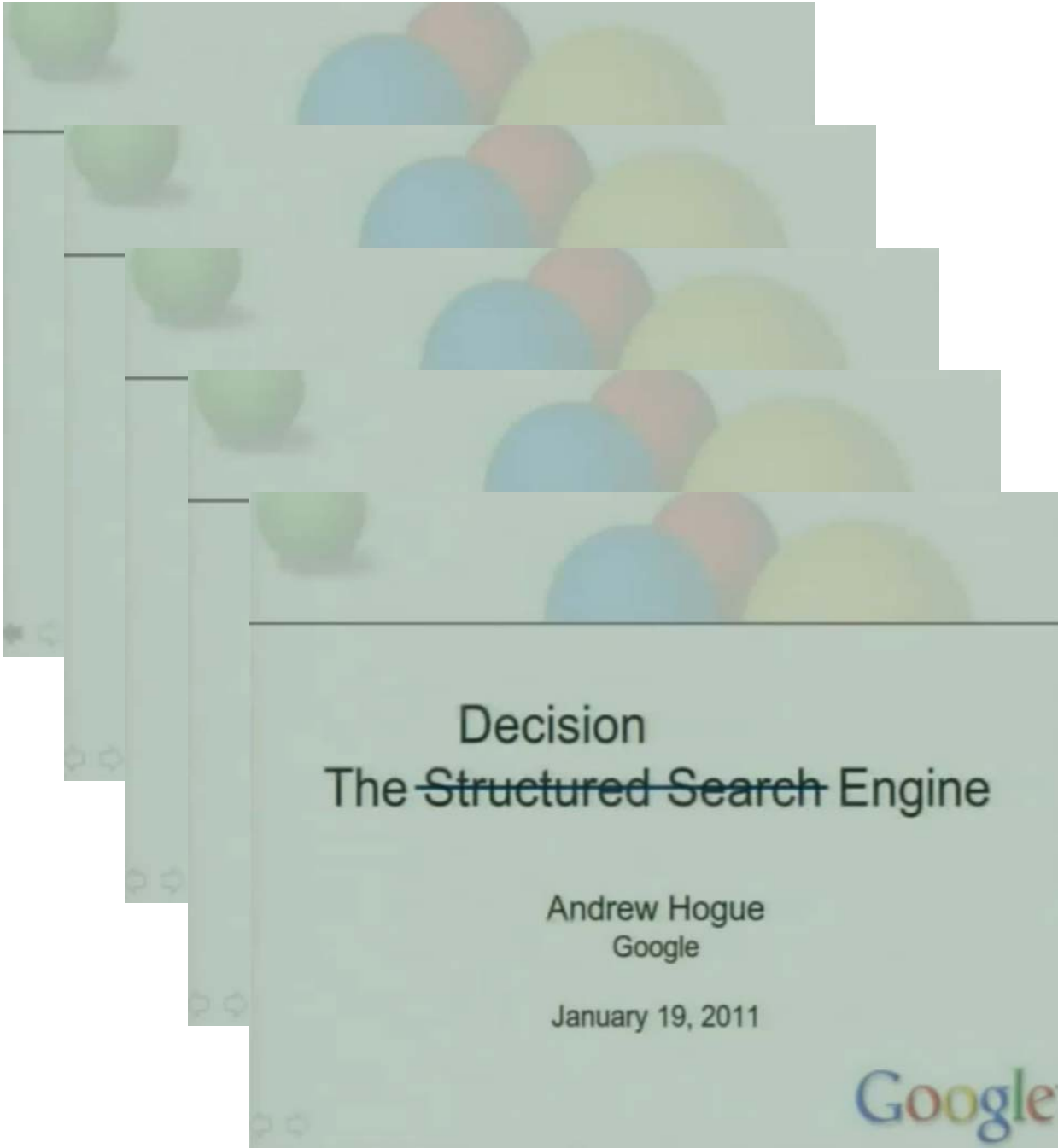


Intelligent The ~~Structured~~ Search Engine

Andrew Hogue
Google

January 19, 2011

Google



Decision

The ~~Structured Search~~ Engine

Andrew Hogue
Google

January 19, 2011

Google

Today

- Can we understand the world outside the web?
→ Freebase and Extraction
- Can we better understand queries?
→ Query parsing and Question Answering
- Can we better understand content?
→ Sentiment Analysis
- Can we make Google do the work for you?
→ Google Squared

Google

Today

- Can we understand the world outside the web?
→ Freebase and Extraction
- Can we better understand queries?
→ Query parsing and Question Answering
- Can we better understand content?
→ Sentiment Analysis
- Can we make Google do the work for you?
→ Google Squared

Google

January 2011

- 20M topics, >400M connections
- all publicly available

Today

- Can we understand the world outside the web?
→ Freebase and Extraction
- Can we better understand queries?
→ Query parsing and Question Answering
- Can we better understand content?
→ Sentiment Analysis
- Can we make Google do the work for you?
→ Google Squared

Google

January 2011

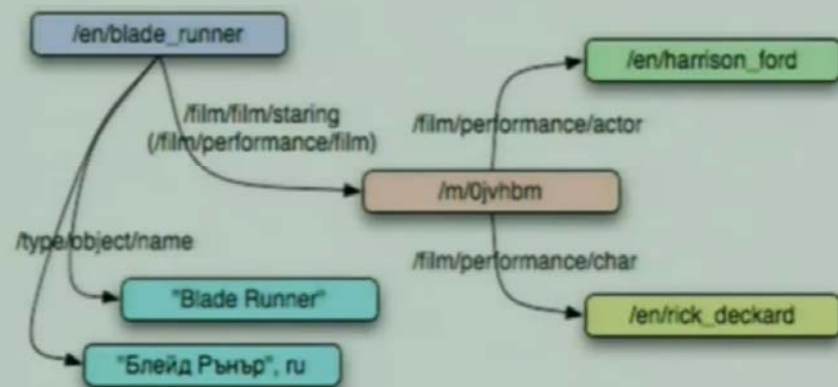
- 20M topics, >400M connections
- all publicly available

among objectives

- make Freebase canonical reference point for entity topic data on the web
- > billion entities

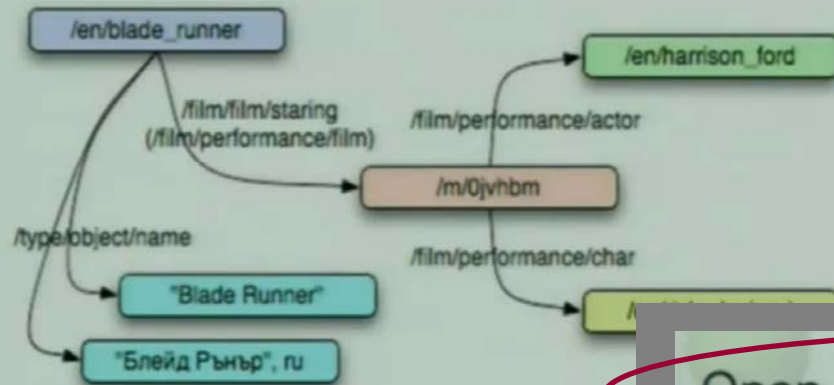
Freebase

entity schema



Google

Freebase



Open-Domain Fact Extraction

- Approach: lots of overlapping techniques with **high coverage, low precision**
 - Table extraction
 - "Clear prompts" (foo: bar)
 - Known value types (e.g. "depth" always takes a measurement of length)
 - Page segmentation
 - Wrapper induction
- Lots of machine learning
- Lots of evaluation
- Lots of iteration!

confidence

entities

Candidate Names

Attributes

- 0.753 **Manufacturer:** Sony
- 0.753 **Part number:** SS-F7000
- 0.747 **Phone:** 1-877-865-SONY
- 0.747 **Email:** contact@sei.sony.com
- 0.747 **Fax:** 941-768-7790
- 0.653 **Weight:** 33.1 lbs
- 0.630 **Max (RMS) Output Power (Total):** 180 Watt
- 0.630 **Response Bandwidth:** 35 - 50000 Hz
- 0.630 **-3dB Frequency Response:** 35 - 50000 Hz
- 0.630 **Input Impedance:** 8 Ohm
- 0.630 **Sensitivity:** 90 dB
- 0.630 **Magnetic Shield:** Yes
- 0.630 **Connectivity Technology:** Wired
- 0.630 **Recommended Placing:** Floor-standing
- 0.629 **Detachable Grilles:** Yes
- 0.619 **Width:** 11.4 in
- 0.619 **Depth:** 11.9 in
- 0.619 **Height:** 41.9 in
- 0.610 **System Components:** 2 speakers
- 0.610 **Speaker Type:** 4-way - Passive
- 0.420 **Average user rating:** 4.5 stars out of 2 reviews
- 0.420 **My rating:** 0 stars



useful threshold ?

[Show/hide low confidence attributes](#)

[Show/hide low confidence attributes](#)

0.298 **Pros:** accurate 4-way design

0.298 **Cons:** "Sony" name will run off most persnickety audiophiles

0.288 **Right/left channel speaker:** 1 x Midrange driver - 3.25 in.,

0.288 **Right/left channel speaker:** 2 x Woofer driver - 8 in

0.202 **Sort by:** Most helpful Highest rating Lowest rating Most recent

0.181 **Ships:** Today

0.181 **out:** 2

0.181 **rating:** 5

0.181 **Average:** 5

0.181 **Average:** 4

0.181 **rating:** 5

0.163 **reviews:** 5

0.163 **reviews:** 4

0.163 **Average:** 5

0.163 **Average:** 4

0.162 **out:** 2

0.156 **Showing:** 1

0.156 **out:** 2

0.152 **out of:** 10

0.152 **out:** 1

0.149 **Updated on:** Aug 28, 2009

0.149 **Updated:** Aug 28, 2009

0.148 **Showing:** 1

0.148 **out:** 2

0.145 **lbs:** 0

useful threshold ?

