Good evening
Today’s Menu

1. Why I’m here
2. What we’re doing
3. What people are doing with it
4. My recommendations for commercial focus
Why I’m here

This guy is in between me and dinner. Why am I here?
What We’re Doing

We provide an open-access web service that uses NLP technology to extract meaning from text.

• What it does
• Why we focus on news
• Why we’re (mostly) giving this away
• What people are doing with it
Unstructured Documents (Text/HTML/XML)

CALAIS
Powered by Thomson Reuters

Categories
- Entertainment, Politics, Sports, Business, etc.

Named Entities
- People, Companies, Geographies, Albums, Authors, etc.

Facts
- Position, Alliance, Education, Political Affiliation, etc.

Events
- Management Change, IPO, Labor Action, Sporting, Entertainment, etc.
Results

• About 30,000 users (5K active)
• 3-6 million transactions / day
• Thousands of sites
• Dozens of businesses
Why Focus On News?
Why News?
Why Give This Away?

- Not a freemium model – 95%+ of users pay zero

- Battle hardening mission critical software
- Foraging for innovative ideas
- Metadata*
Metadata

• 2.5 years of operations
• Approximately 110,000,000,000 triples
• Perhaps 50,000,000,000 facts and events
• Extraordinary diversity of sources
• Increasingly near real time

• What newly appointed CFO’s of companies within my trading portfolio are headquartered in NY and have been associated with an SEC investigation in the last five years?
Immersion Learning

18 releases

About 150 presentations

Talked with 200+ customers

Heard a million great ideas

30,000 registered developers

9,000 emails
What We’ve Heard: Six Buckets

Tools

Social

Advertising

Search

Publishing

Insight Generation

Dark Matter
Consultant's Refuge

The 2x2 Grid

Marketplace Size

Distinctive User-Derived Value

- Medicine
- Accordions
- Dirt
- Soap
Consultant's Refuge

The 2x2 Grid

Distinctive User-Derived Value

Various Kinds of Good

Don’t go Here

“Brand” Domain

Marketplace Size
What We’ve Heard: Six Buckets

Tools (+ to ++)
Social (- to +)
Advertising (+)
Search (-- to +)
Publishing (+ to ++)
Insight Generation (+ to ++)

Dark Matter
#1 – Enabling Tools

Semantic data management
Semantic data generation
Databases
Integration & Workflow

[Diagram showing a circle with 'Distinctive User-Derived Value' and 'Marketplace Size' axes, marked with a green check mark.]
#2 - Social

Semantics-powered link sharing
Network mining
News sharing
Tweet mining
#3 - Advertising

Semantic ad placement
Contextual ad placement
Semantically-driven landing pages
Mashup Ads
#4 – Semantic Search

General “semantic search”

Vs..

Domain specific semantically-enhanced search
#5 – Publishing

A-Content Producers – from back office to user experience

B-Editorial + Aggregation
Publishing Models

C-Robotic publishing – aggregation only
Dark Matter #1

What Industry?

Leveraged new technology to deliver a seamless, tactile and compelling user experience…

And created a $57 Billion market?
Gaming 101 vs. Today’s State of the Web

Great story line

High interactivity, immediate responsiveness

No interruptions

Graphically engaging

Seamless

And

Fun
www.zigglydoodle.biz or
Hmmm...

I wonder why Google
Dark Matter #2

Disintermediation

Not just business models – but business processes

Model: Media Monitoring

Case: Workflow
Our Suggestions for Success

Put your idea in the 2x2. Be honest. Stay out of the deadzone no matter how cool

Decide if you’re care about semantics or about user value. If it’s semantics – be a tool vendor.

Don’t fund or spend money on semantic infrastructure beyond what’s necessary. The basic building blocks are available.

Think hard about user experience. Make it amazing.

Think hard about disintermediation – reducing cost base is the easiest sell there is
Thank you

ttague+mit@gmail.com