"Text Mining and Natural Language Processing Technologies to Support Competitive Intelligence Efforts"

Chris Beguel
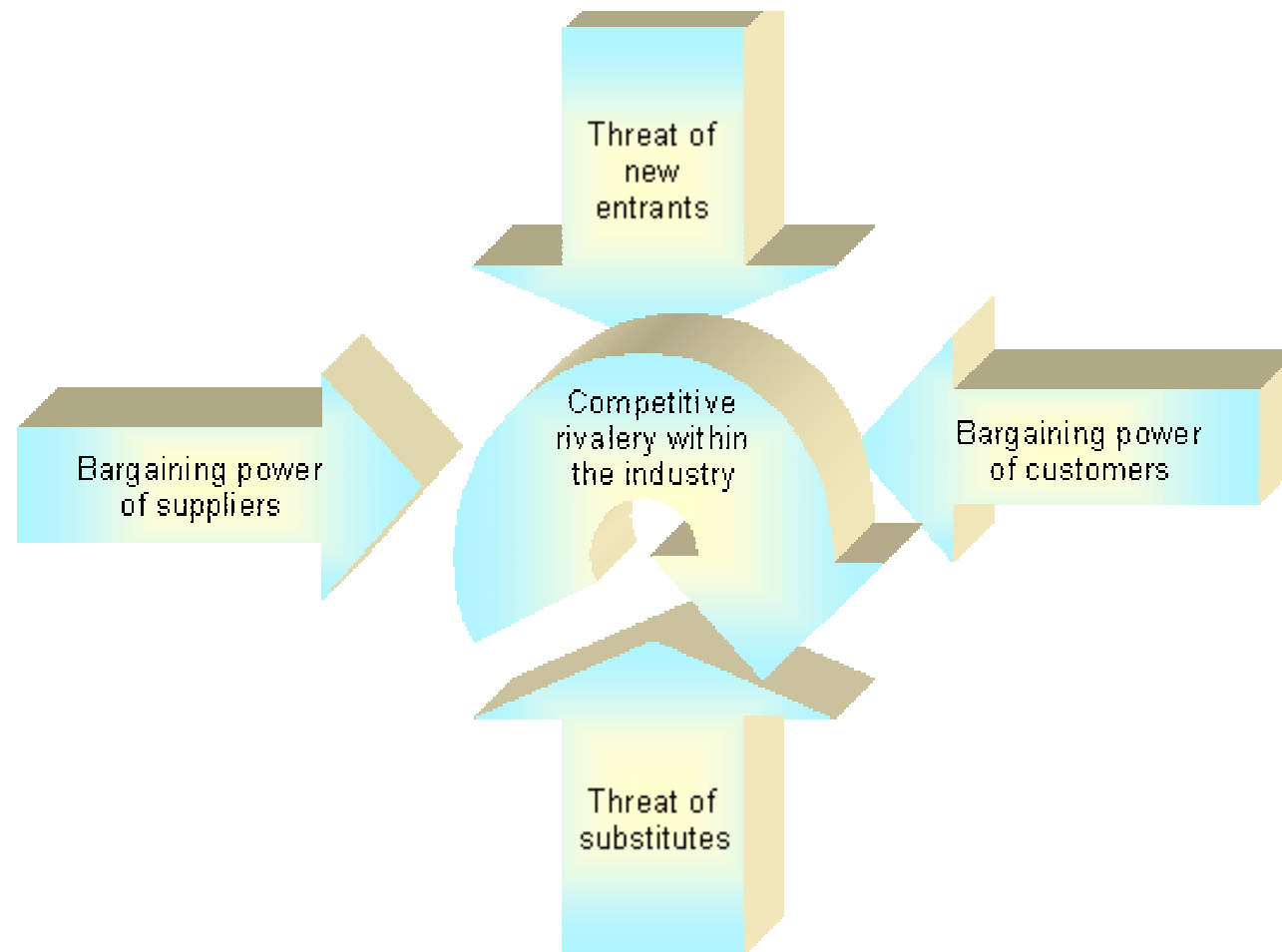*Business Development Manager*

Agenda

- What is Competitive Intelligence?

- What market does this represent today?

- Case study 1: Telecom Italia Mobile

- Case study 2: Total

- Case study 3: Pharma

- How does TEMIS address those issues?
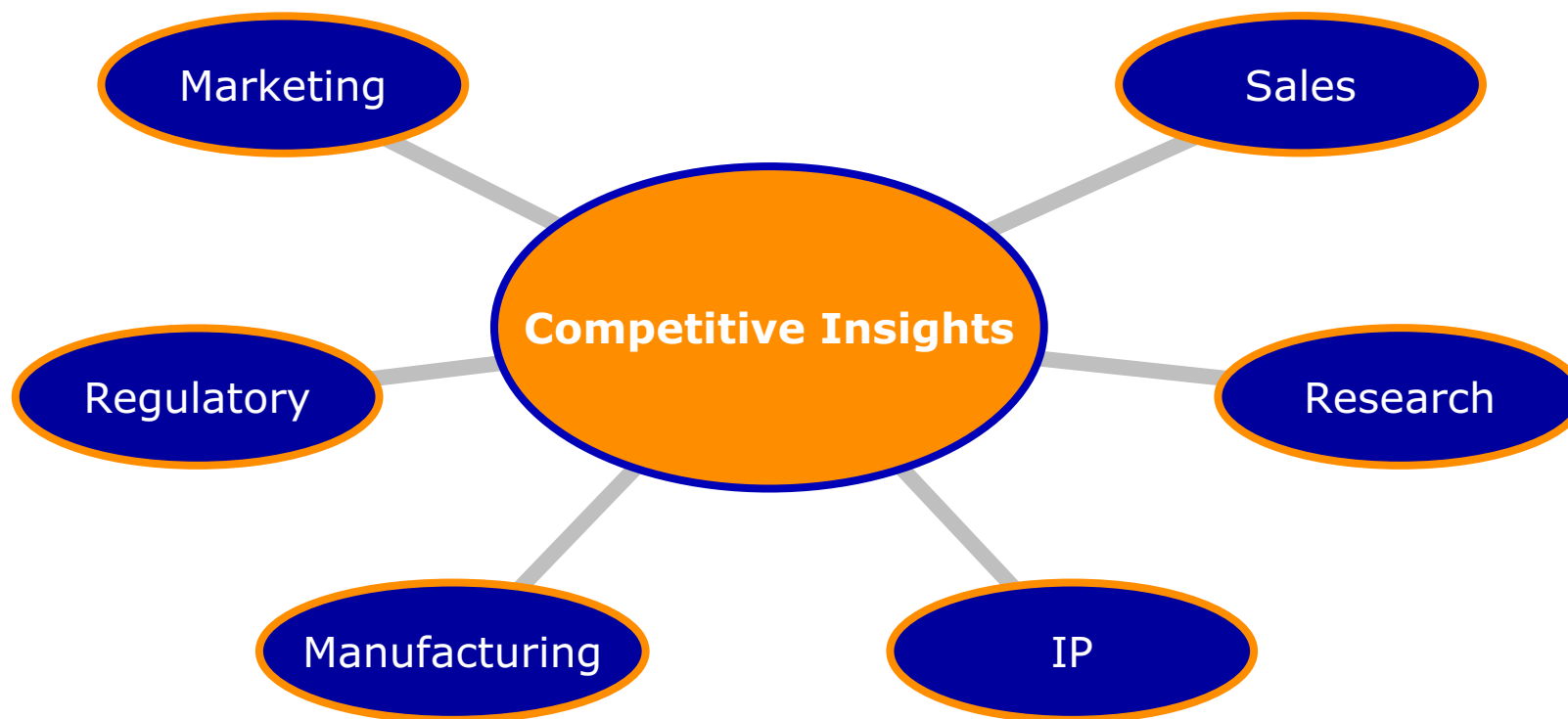
- Conclusion

Porter's five forces

What is Competitive Intelligence?

"Competitive Intelligence is a systematic and ethical program for gathering, analyzing, and managing external information that can affect your company's plans, decisions, and operations."

Society of Competitive Intelligence Professionals, http://www.scip.org/

CI is cross-functional

The Competitive Intelligence Cycle
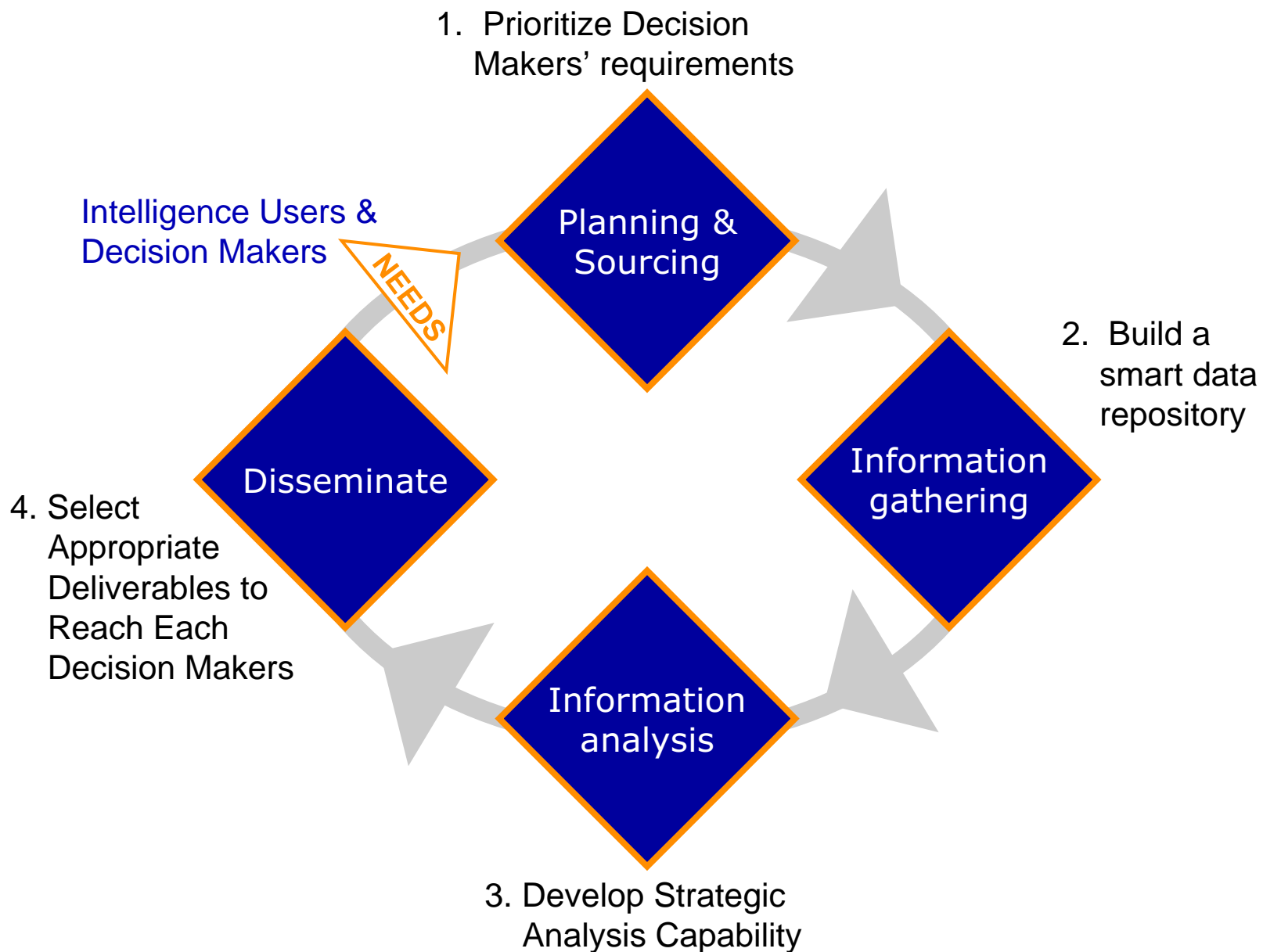
1. Prioritize Decision Makers' requirements

Intelligence Users & Decision Makers

NEEDS

Planning & Sourcing

2. Build a smart data repository

Disseminate

4. Select Appropriate Deliverables to Reach Each Decision Makers

Information gathering

Information analysis

3. Develop Strategic Analysis Capability

Agenda

- What is Competitive Intelligence?

- What market does this represent today?

- Case study 1: Telecom Italia Mobile

- Case study 2: Total

- Case study 3: Pharma

- How does TEMIS address those issues?

- Conclusion

Who is doing Competitive Intelligence?

- 90% of Fortune 500 firms in the U.S.

- Almost 50% of leading European companies

- 9% of U.S. firms with formal processes

- Chemical and telecommunications firms

- Firms with high R&D expenditures

- Firms that own many patents

- U.S. & U.K. firms: leading intelligence producers

Agenda

- What is Competitive Intelligence?

- What market does this represent today?

- Case study 1: Telecom Italia Mobile

- Case study 2: Total

- Case study 3: Pharma
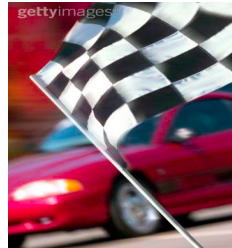
- How does TEMIS address those issues?

- Conclusion

Slide 9

Case Studies

Life Science
Monitor competitive activities & increase productiveness

- Competitive Intelligence
- Patent Analysis
- Preclinical Research
- Development (Therapeutic Areas)
- Safety Assessment & Pharmacovigilance

Automotive
Improve car quality & customer relationship

- Solution for Quality Departments
- Analyze customer & employee reports
- Manage faults & customer satisfaction
- Minimize warranty costs
- Terminology & documentation

Energy
Customer & Corporate Intelligence

- Competitive Intelligence
- Customer feedback analysis
- & customer Intelligence
- Terminology extraction

Telecom & Internet
News Monitoring

- People & companies reputation
- Paradigm shift detection
- Competitive Intelligence

Case Study 1: TIM

- Profile
  - #1 in Italy (~26,2 Million subscribers)
  - Leader in Greece, Turkey and Brazil
  - Corporate and individual customers

- Context
  - Telecom industry in Europe and South America: deregulation, M&A, concentration, new players, new technologies

- Sources
  - Competitors' websites, Industry portals, weblogs, newsgroups

- Solution
  - Competitive Intelligence Skill Cartridge™
  - Online Miner™ Portal+Webcrawler

Case Study 2: Total

- **Context**

  - Total examines and analyzes geopolitical, economic, technical, financial, social, macro & micro information
  - One division must supply quality information to all business units

- **Expectations**

  - One system that collects and analyzes documents with the goal to provide information sorted and organized to be used effectively

- **Solution = Online Miner™ + CI Skill Cartridge™**

  - 5,000 users connected

- **Benefits = productivity gains**

  - Time savings of about ½ hour per day per user
  - Gain = 5,000 * 1/16 * 400 = 125,000 € per day

Case Study 2: Total

List of areas of interest

Text annotation

Relationship details

Detected relationships (in one area of interest)

Organisations | People | Categories | Documents

Acquisition and Selling (133)
Company Relation (139)
Competition Regulation (44)
Customer (15)
Export Import (169)
Field (75)
Financial (77)
Infrastructure (332)
Investment (108)
JointMarketing (2)
Market Share (9)
Mergers (9)
Partnership (196)
Privatization (15)
Stake Information (101)
Stock Index Information (4)
Stock Information (71)
Person - Head (320)

0 1

Acquisition and Selling

Alon Inc.	<=>	Petrom
Arco British	<=>	Arco
ArmRosgazprom	<=>	Armenia-Azerbaijan
• ASM • Endesa	<=>	Enel's
BELGIAN-RUSSIAN	<=>	MOLDOVA GRES
BELGIAN-RUSSIAN	<=>	MOLDOVA GRES
Bermudas	<=>	Gazprom
BG Group	<=>	Gm3/yr
BP	<=>	
BP	<=>	Southern Basin

Copyright 2003. Platts. All Rights Reserved.

Vilnius, Lithuania (Platts)–21Nov2003/228 pm EST/1928 GMT Poland's state oil company PKN Orlen said Friday submitted a preliminary non-binding bid for a majority stake in Romanian oil and gas company SNP Petrom, PKN Orlen said in a statement. Nov 21 was the last day to submit preliminary bids for Petrom to the Romanian Ministry of Economy and Commerce. In October, 11 companies were shortlisted to bid, including: OMV; Gazprom; MOL; Eni of; PKN Orlen; Hellenic Petroleum; Occidental Petroleum; TNK-BP; ConocoPhillips; Glencore International from Switzerland; and Alon Inc. Romanian government officials plan to sell a 51% stake in Petrom in a two-stage process. In the first stage, a 33.34% stake will be sold to a strategic investor will acquire newly issued Petro... take to 51%. Petrom operates the Arpechim an... ...s 695 service stations and 128-mil tonnes of ...

Acquisition and Selling
Stake Information

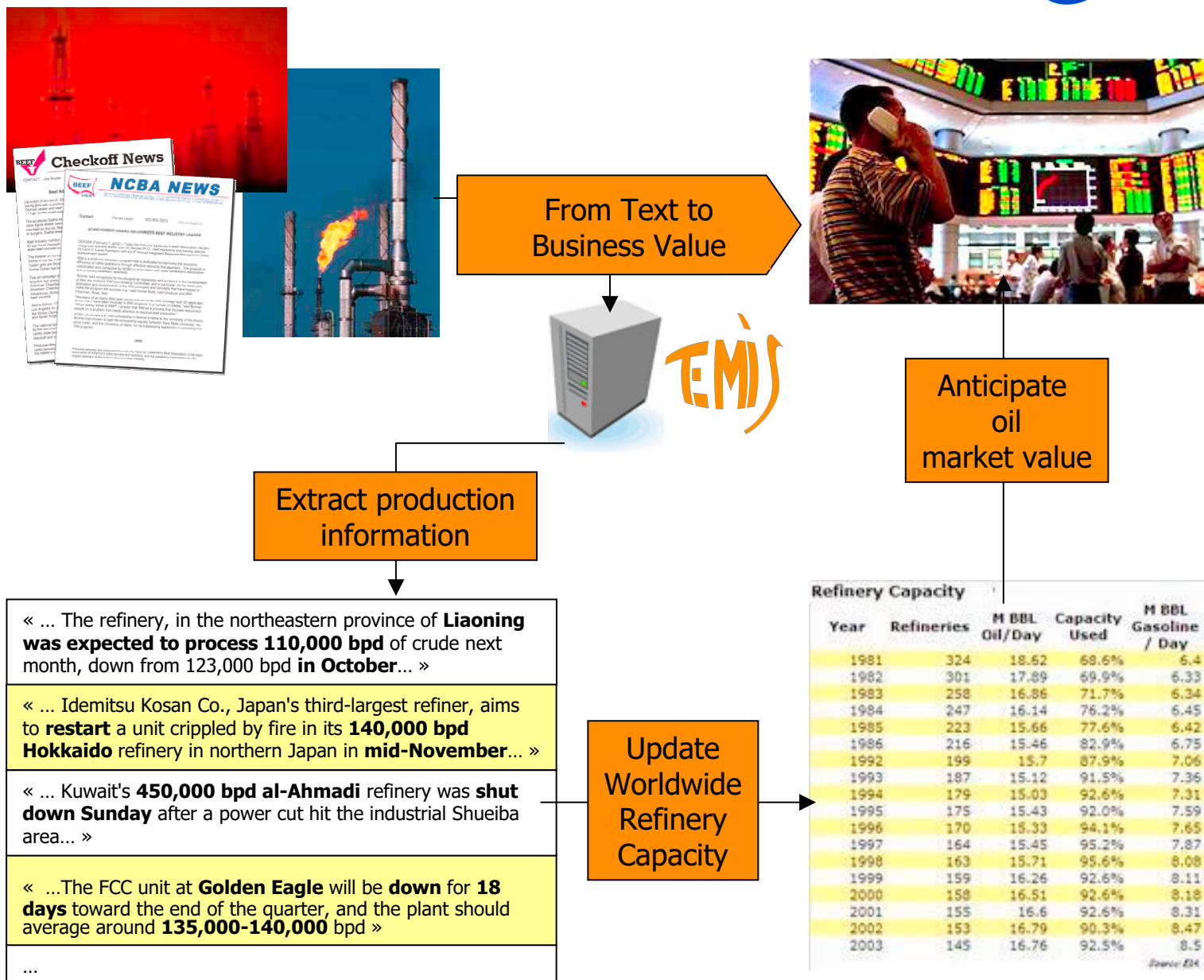| /Where | |
/Europe/Switzerland	Switzerland
	Alon Inc.
who	Romanian government officials
/guessed actor	Alon Inc.
Rumor	plan to
/CI Stake	sell a 51% stake
/selling cession	sell
whom	in Petrom
/guessed actor	Petrom

Lietuvos Dujos' Jan-Oct supplies of Russia gas to consumers up.

250 mots

Prime-TASS Energy Service (Russia)

Anglais

Slide 13

Case Study 2: Total

From Text to Business Value

Extract production information

Anticipate oil market value

« ... The refinery, in the northeastern province of **Liaoning was expected to process 110,000 bpd** of crude next month, down from 123,000 bpd **in October**... »

« ... Idemitsu Kosan Co., Japan's third-largest refiner, aims to **restart** a unit crippled by fire in its **140,000 bpd Hokkaido** refinery in northern Japan in **mid-November**... »

« ... Kuwait's **450,000 bpd al-Ahmadi** refinery was **shut down Sunday** after a power cut hit the industrial Shueiba area... »

« ...The FCC unit at **Golden Eagle** will be **down** for **18 days** toward the end of the quarter, and the plant should average around **135,000-140,000** bpd »

...

Update Worldwide Refinery Capacity

Refinery Capacity

| Year	Refineries	M BBL Oil/Day	Capacity Used	M BBL Gasoline / Day
1981	324	18.62	68.6%	6.4
1982	301	17.89	69.9%	6.33
1983	258	16.86	71.7%	6.34
1984	247	16.14	76.2%	6.45
1985	223	15.66	77.6%	6.42
1986	216	15.46	82.9%	6.75
1992	199	15.7	87.9%	7.06
1993	187	15.12	91.5%	7.36
1994	179	15.03	92.6%	7.31
1995	175	15.43	92.0%	7.59
1996	170	15.33	94.1%	7.65
1997	164	15.45	95.2%	7.87
1998	163	15.71	95.6%	8.08
1999	159	16.26	92.6%	8.11
2000	158	16.51	92.6%	8.18
2001	155	16.6	92.6%	8.31
2002	153	16.79	90.3%	8.47
2003	145	16.76	92.5%	8.5

Source: EIA

Case Study 3: Pharma

New chemical compound is discovered and tested in the labs

Drug is tested in healthy humans to assess safety
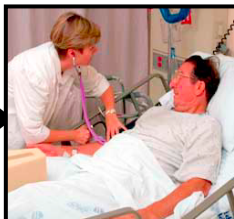
Drug is tested in patients to assess effectiveness

Product is marketed

Patent expires

12 Years
~ $802M

Case Study 3: Pharma

- **User Group**
  - Disease Area Managers, Product Managers, Business Analysts, etc.

- **Objective**
  - Overhaul the Corporate Competitive Intelligence system with Analytics capabilities

- **Key benefits**
  - **Enables search by concepts** (e.g. "which companies signed licensing agreement with Pfizer in the past ten months?", "what are the product for hypertension in phase II and III?")
  - **Provides the "big picture"** on complex topics with clustering and mapping technology (e.g. ~2000 docs about Type 2 Diabetes summarized in 2 sec.)
  - **Enables trend analysis** in a few seconds (e.g. the VIOXX crisis week after week)

Case Study 3: Pharma

- **Project**
  - Leverage Competitive Intelligence portal retrieval and analytics capabilities (information extraction, clustering, categorization, navigation)

- **Sources**
  - IMS Health, NewsEdge, Factiva, IDDB3, PJB Scrip…
  - Company's websites, news portal, blogs,…
  - Internal documents and reports

- **IT-Context:**
  - Deployment of TEMIS Online Miner™ with customized versions of the CI Life Science Edition and Medical Entity Relations Skill Cartridges™.
  - Integration with Oracle DB, Document Management System (Documentum) and Search Engines (Convera RW, Verity, FAST)
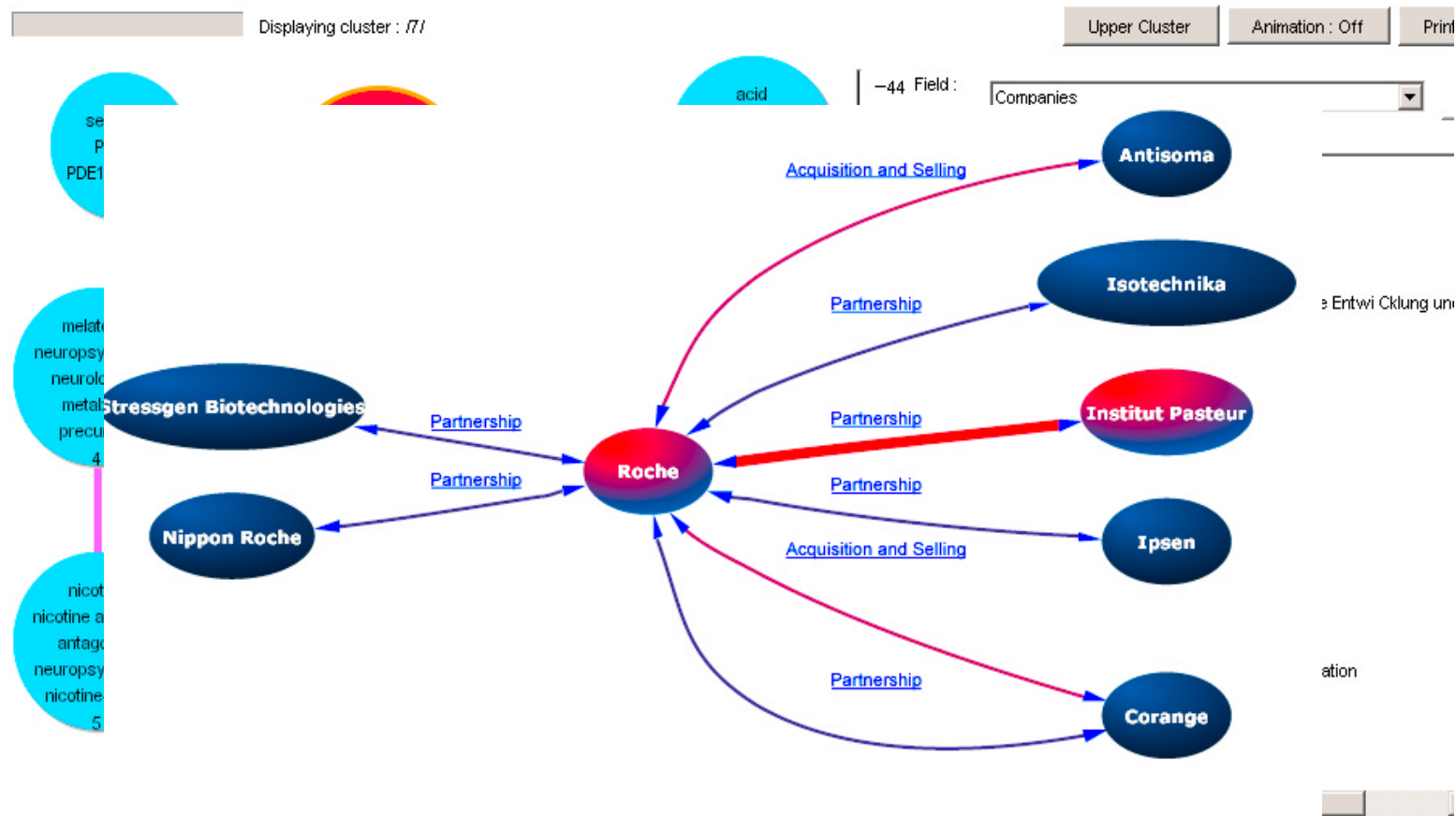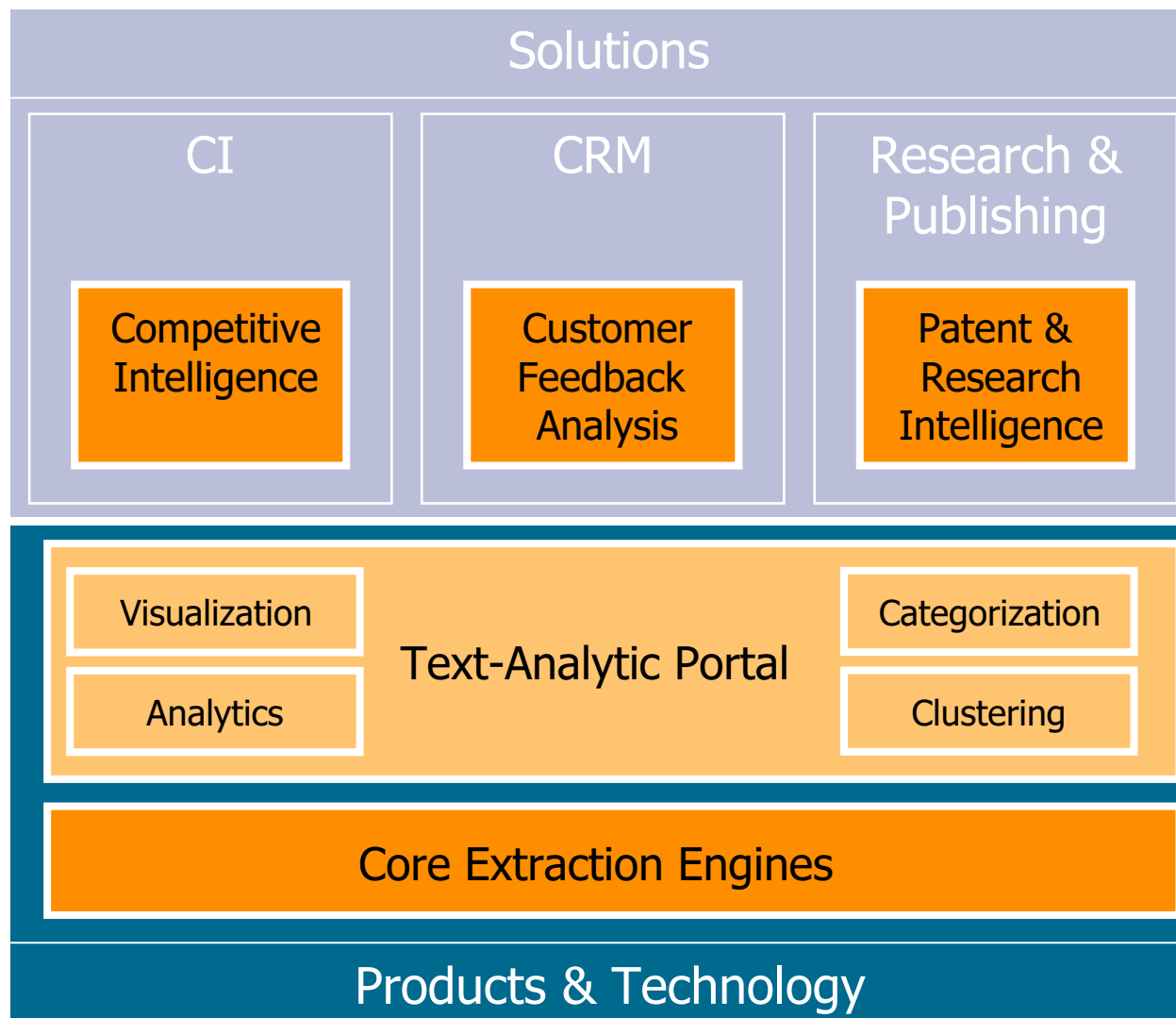
Agenda

- What is Competitive Intelligence?

- What market does this represent today?

- Case study 1: Telecom Italia Mobile

- Case study 2: Total

- Case study 3: Pharma

- How does TEMIS address those issues?

- Conclusion

Product & Solution Stack

Solutions

CI

Competitive Intelligence

CRM

Customer Feedback Analysis

Research & Publishing

Patent & Research Intelligence

Text-Analytic Portal

Visualization

Analytics

Categorization

Clustering

Core Extraction Engines

Products & Technology

Core Extraction - Meaning & Metadata

Standard text analysis will turn Text into Words

Pharmacia Corp. was acquired by Pfizer Inc. in 2003

XeLDA™

Pharmacia Corp. [1-14]:	?	+*PROP (guessed)*
was [16-18]:	be	+*VBPAST*
acquired [20-27]:	acquire	+*VPAP*
by [29-30]:	by	+*PREP*
Pfizer Inc.[32-42]:	Pfizer	+*PROP*
in [44-45]:	in	+*PREP*
2003 [47-50]:	2003	+*CARD*

TEMIS turns Text into Meaning !

Pharmacia Corp.   was acquired   by Pfizer Inc.   in 2003

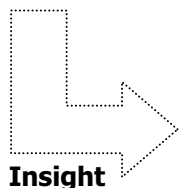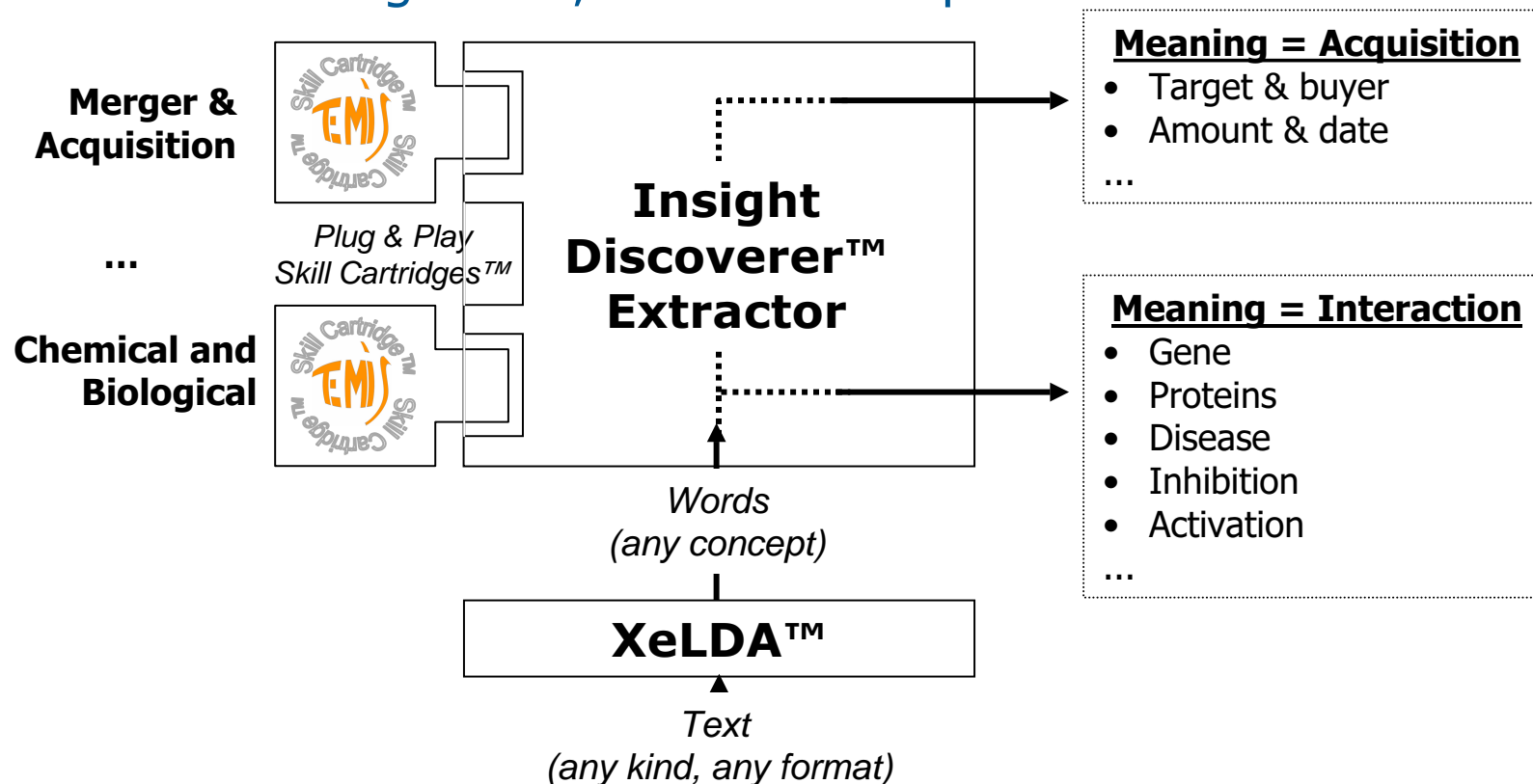| **Target**	**Action**	**Buyer**	**Date**
Company	*Acquisition*	*Company*	*Year*

Insight Discoverer™ Extractor

Meaning = Acquisition (action)
- Target = Pharmacia Corp.
- Buyer = Pfizer Inc
- Amount = *unknown*
- Date = 2003

...

- **Plug & Play annotation components**
  - Open architecture
  - Each defines areas of interests & extraction rules

- **Capture the domain knowledge**
  - Some is generic, but lot of it specific

Merger & Acquisition

...

Chemical and Biological

Plug & Play Skill Cartridges™

Insight Discoverer™ Extractor

Meaning = Acquisition
- Target & buyer
- Amount & date
...

Meaning = Interaction
- Gene
- Proteins
- Disease
- Inhibition
- Activation
...

Words (any concept)

XeLDA™

Text (any kind, any format)

Real Semantic Analysis

Company A announced the **acquisition** of **Company B** **in 2005**

> **Meaning = Acquisition**
> - Target: Company B
> - Buyer: Company A
> - Amount: $20 millions
> - Date: June 2005
> ...

is different from ...

Company A said they would want to **acquire** **Company B** **by 2007**

> **Meaning = Acquisition**
> - ***rumor***
> - Target: Company B
> - Buyer: Company A
> - Amount: *Unknown*
> - Date: 2007
> ...

Real Semantic Analysis

Company A announced the acquisition of Company B …

> **Meaning = Acquisition**
> - Target: Company B
> - Buyer: Company A

is the same as…

Company B has been acquired by Company A …

> **Meaning = Acquisition**
> - Target: Company B
> - Buyer: Company A

…but different from…

Company B has *not* been acquired by Company A because …
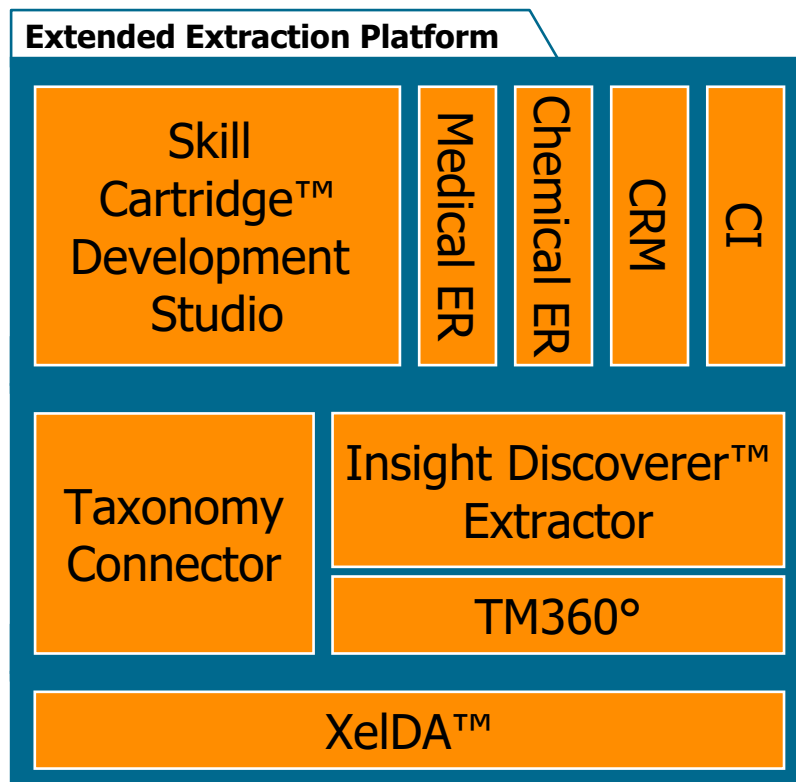
Detect reaction
Find start and end of one chemical reaction

310 g of **2,6-dimethoxyphenol**, 300 g of **formalin**, and 810 g of **aqueous sodium hydroxide** were reacted for 60 hours at 25° C. under a pressure of 0 kg/cm^2-G, the reaction **mixture** was cooled and then neutralized with **sulfuric acid**, to obtain 80 g of **2,6-dimethoxy-4-(hydroxymethyl)phenol**. (Yield 22 mol%).

(Yield 22 mol%). **25 g of the obtained product was reacted in 320 ml of methanol in the presence of platinum-alumina catalyst and hydrogen for 4 hours at 200° C. at 199 kg/cm2-G, and 15 g of 2,6-dimethoxy-4-methylphenol was obtained. (Yield 6 mol%)**

Extended Extraction Platform

Skill Cartridge™ Development Studio

Medical ER

Chemical ER

CRM

CI

Taxonomy Connector

Insight Discoverer™ Extractor

TM360°

XelDA™

- **Extended Extraction Platform**
  - Library of existing SC™
  - Ability to enrich Skill Cartridges™

- **Universal extraction**
  - Ability to develop domain specific Skill Cartridges™
  - Graphical Development Environment
  - Full APIs

- **Linguistic Services**
  - Language identification
  - Tokenization
  - Lemmatization
  - Part of Speech tagging
  - 16 languages supported
  - 4 more to come in H2

  - Out of the box extraction
  - Any universal named entities (currency, date, location, people name, title, phone, company, product,…) [see]
  - Finance & business relationships (People to companies, companies to companies, Products & companies, …) [see]
  - 5 languages
  - 2 additional in H2

Text Analytics Portal – Online Miner™
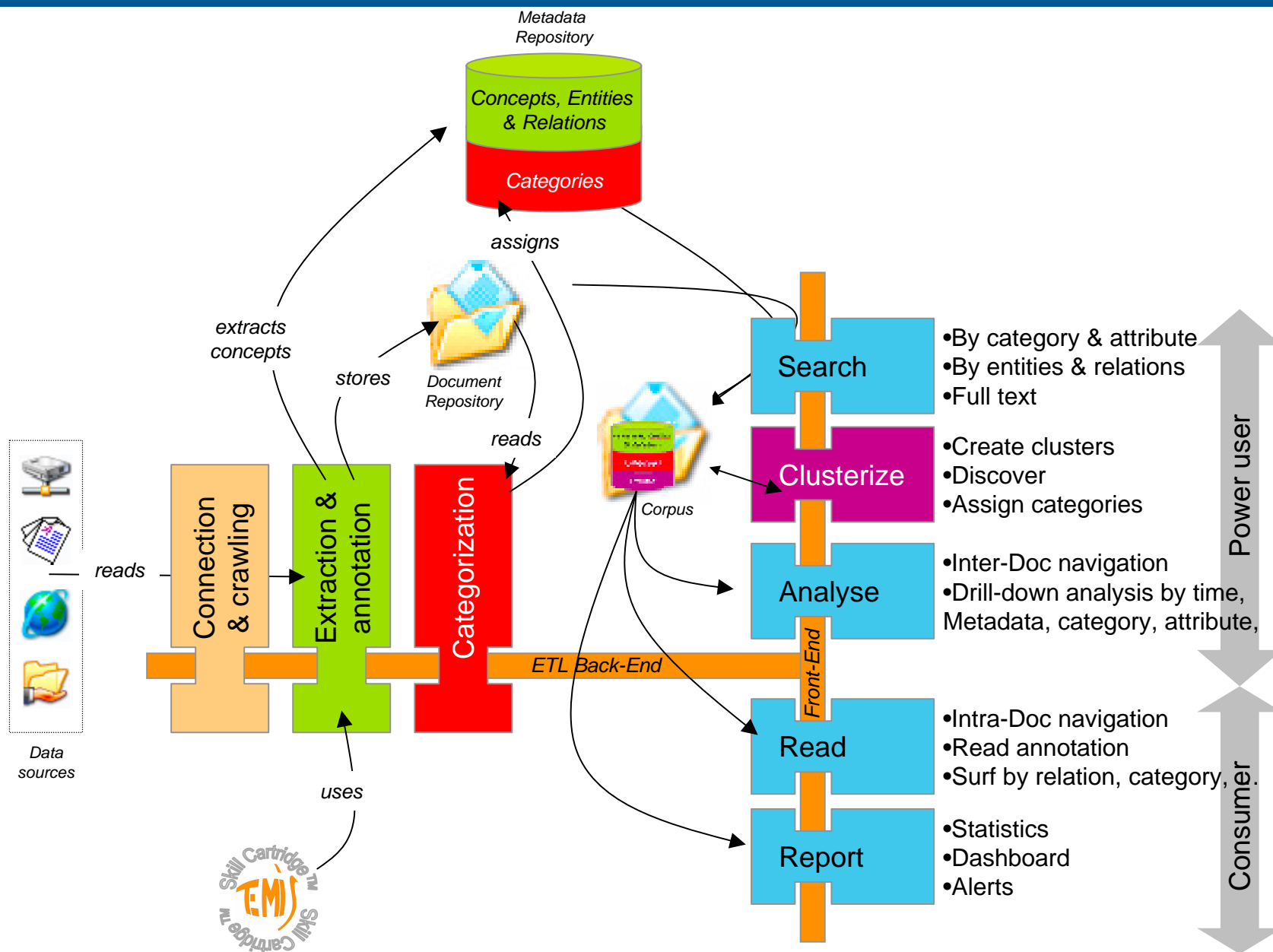
Summary - Our Value Proposition

- **Strong meaning & metadata extraction engines**
  - Extract relevant data out of any text
  - Go beyond simple term detection, leverage grammar!
  - Entities **and** relationship extractions
  - Generic but also domain specific

- **Strategic solutions used by blue-chip customers**
  - To support their "Information Intelligence" strategy
  - Leverage our clustering and categorization engines & our web application framework
  - Life Science (Pharmaceutical), Automotive, Media & publishing and Energy industries

- **Recognized Text Mining expertise**
  - Seasoned engineers & managers
  - Rewarded technology

Thank You!

Please contact us for demonstrations!

TEMIS France
193-197, rue de Bercy
F-75582 Paris Cedex
Tel. +33 1 40 04 46 70

TEMIS Italia
Via G.B. Amici, 29
I-41100 Modena
Tel. +39 59 237 634

TEMIS Deutschland
Kurfürstenanlage 3
D-69115 Heidelberg
Tel. +49 6221 13753-12

TEMIS UK
4, Berkshire Road
Bristol BS7 8EX, UK
Tel. +44 117 949 8646

TEMIS USA
200, Yoakum Parkway
Alexandria, VA 22304 USA
Tel. +1 703 823 5745