

Unconventional tools for analysis and synthesis of information

Yves-Michel MARTI - EGIDERIA

171 rue Saint Honoré, 75001 Paris, France

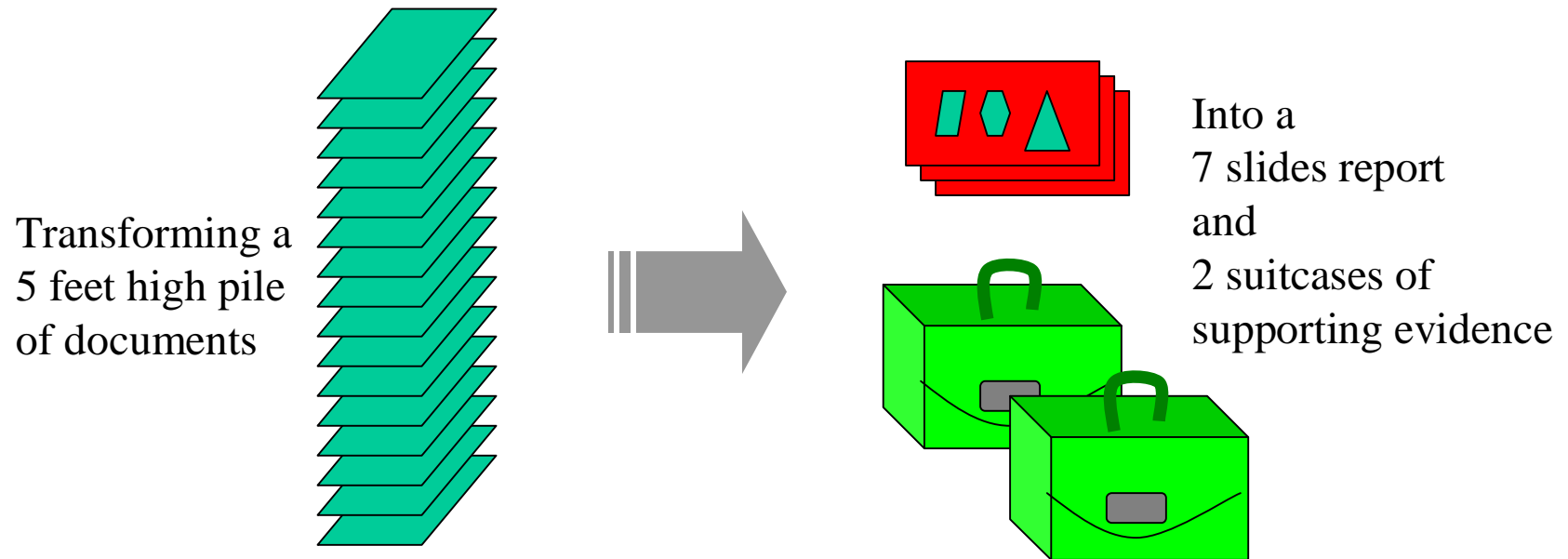
Tel : 00 33 1 40 20 15 50

Fax : 00 33 1 40 20 15 60

Email : mail@egideria.com

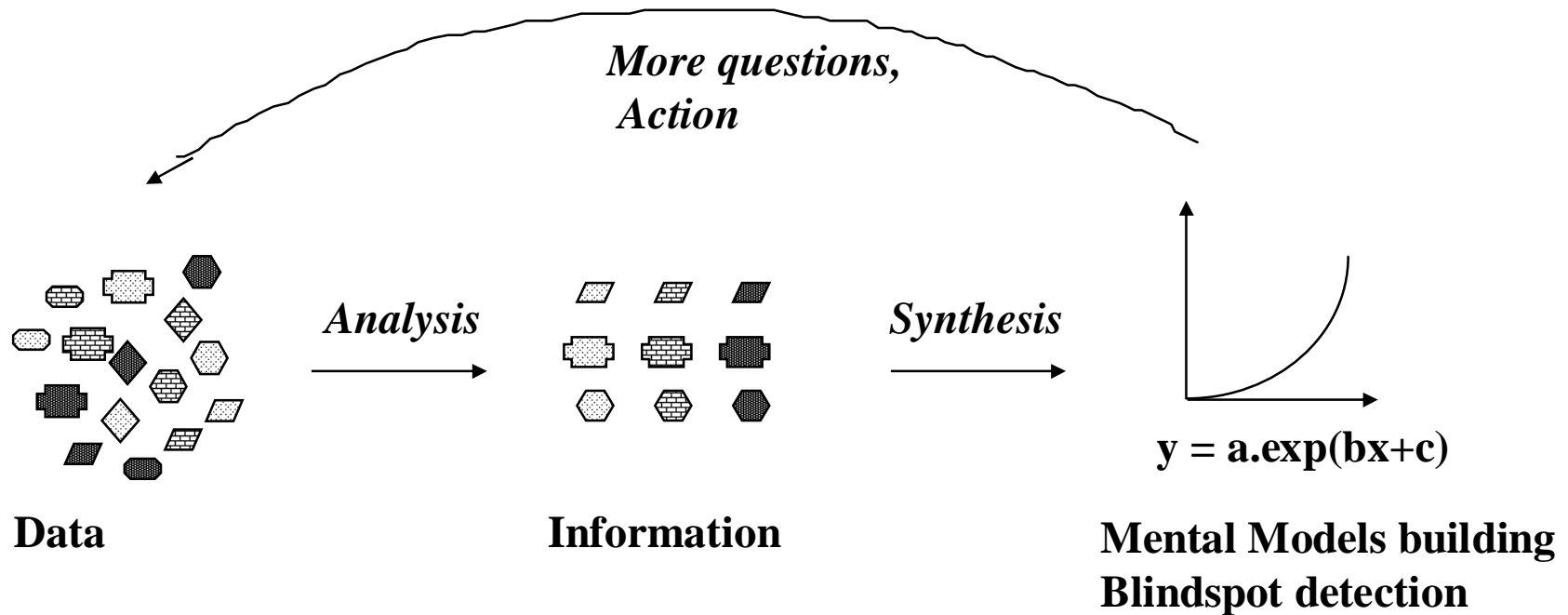
www.egideria.com

The challenge

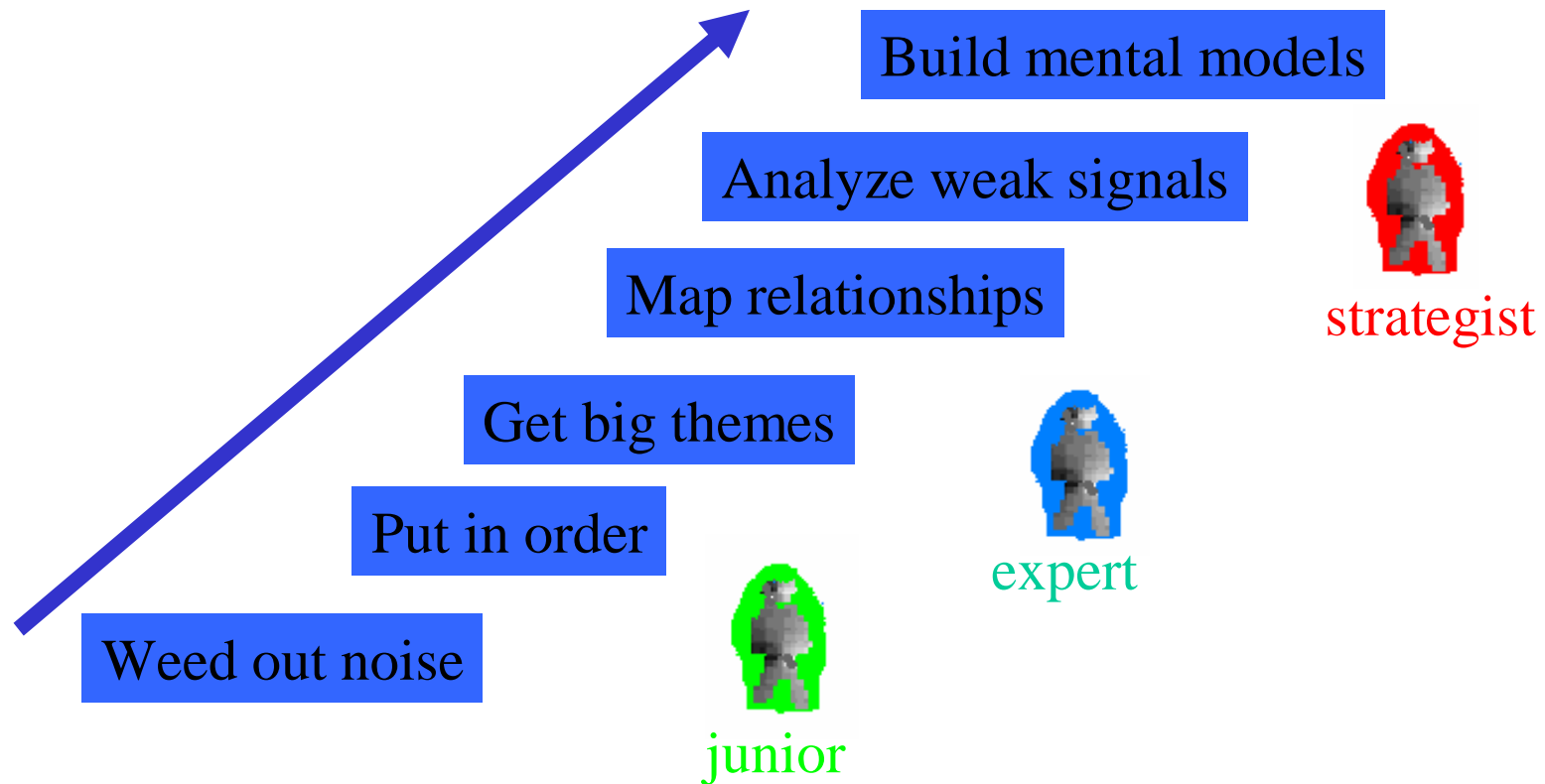


Information fatigue brings bias (tunnel syndrome)

The process of Analysis and Synthesis



The process of Analysis and Synthesis

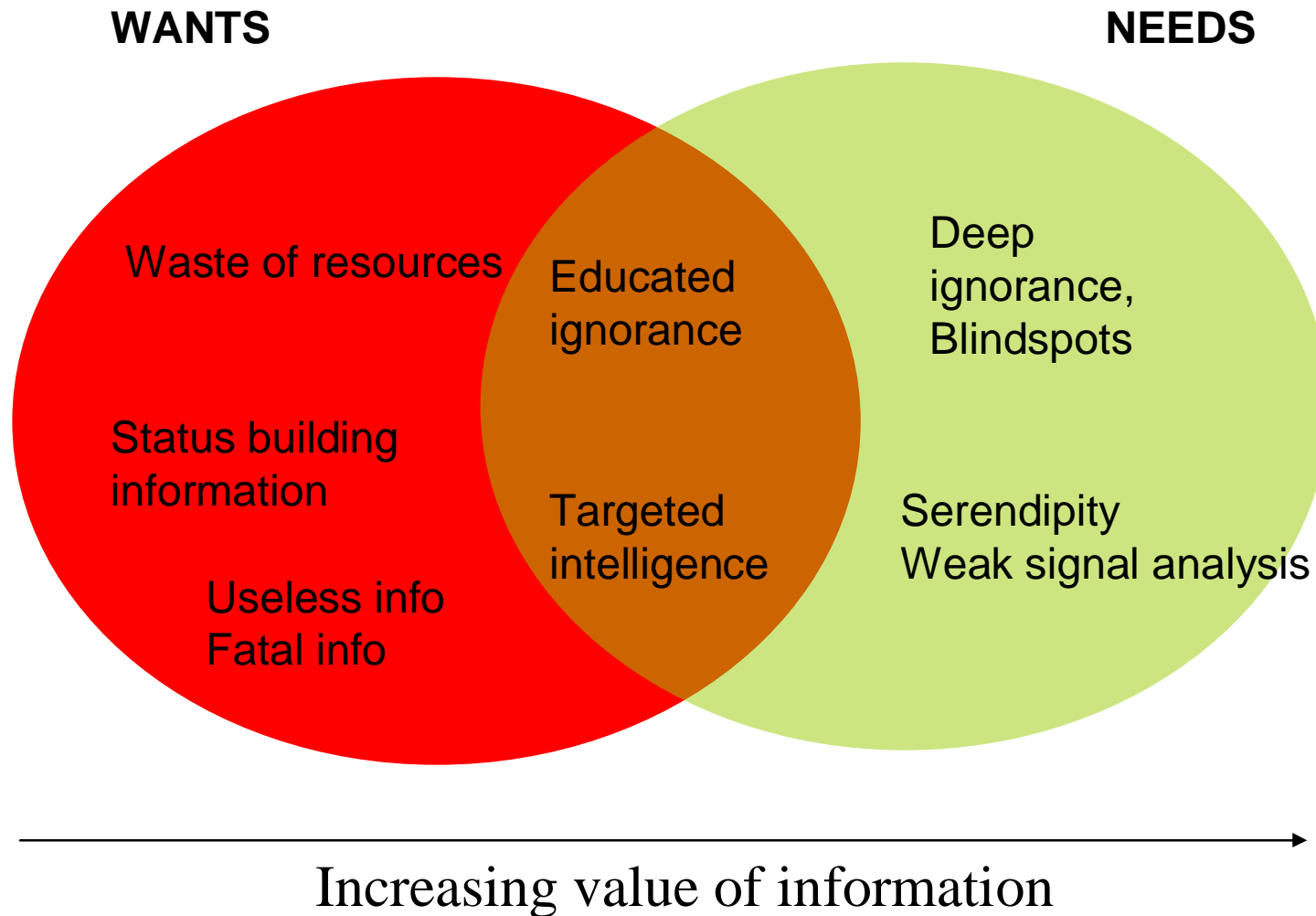


Good synthesis is more an Art than a Science

- Really understanding the decision maker 's needs in order to estimate the value of each piece of information, and of its links to other pieces of information.
- Courage to see the truth and to tell it
- Contrarian spirit in order to question assumptions and to seek weak signals
- The ability to see the world from many points of view

*Fully automated analysis is a dream,
But proper tools can influence our thinking*

Information needs



Weak signal analysis tools

- Primary sources
- Historical analysis of source (databases)
- Cross validation with independent sources
- Intent analysis of source (psy, socio and political analysis)
- Content analysis (textual analysis tools)
- Different viewpoints on same data (information mapping tools)
- Event dynamics analysis (rumour analysis tools)



Analyzing information is like analyzing music

*Every word is like
a musical note.
Looking for
harmony and
dissonance in
textual documents.*

- Quality of recording :
 - primary source
- Tone (synthetic tones) :
 - source analysis, historical record
- Harmony :
 - (dominant note analysis) :cross validation with independent sources
 - (chords) : content analysis
- Melody :
 - event dynamics analysis

Different views on same data

Asian politician :
I may be obliged to
make apologies

Asian worker :
what a loss of face if
I lose my job !

Western worker :
Who cares of a few
quick asian millionaires
lose their BMW 's ?

Wall Street Analyst :
Chaebols should demerger
and clarify their
accounting practices

Asian financial
Crisis 1997

Western politician :
Don 't worry, we
won 't be impacted

Western bank loan officer :
Will Peregrine default
on its loan payments ?

The software tool : UMAP

- Developed by a french company.
- Works from a dataset of documents (corpus), originating from Internet, online databases, scanned documents, etc.
- Analyses dominant keywords within each document, and amongst the document set.
- Creates an information map according to the strength of keywords, and their relationship in documents and in document set.
- Parameters can be tuned for producing many different views on same data
- Allows for exploration and manipulation of dataset (corpus) and keyword set (thesaurus)

Method : Analyzing a competitor



<u>DATA</u>	<u>INTERPRETATION</u>
• Website	Official image
• Online database data	Media image
• Patents	Technological platform
• Web search	Image in industry
• Newsgroups	Image with end user

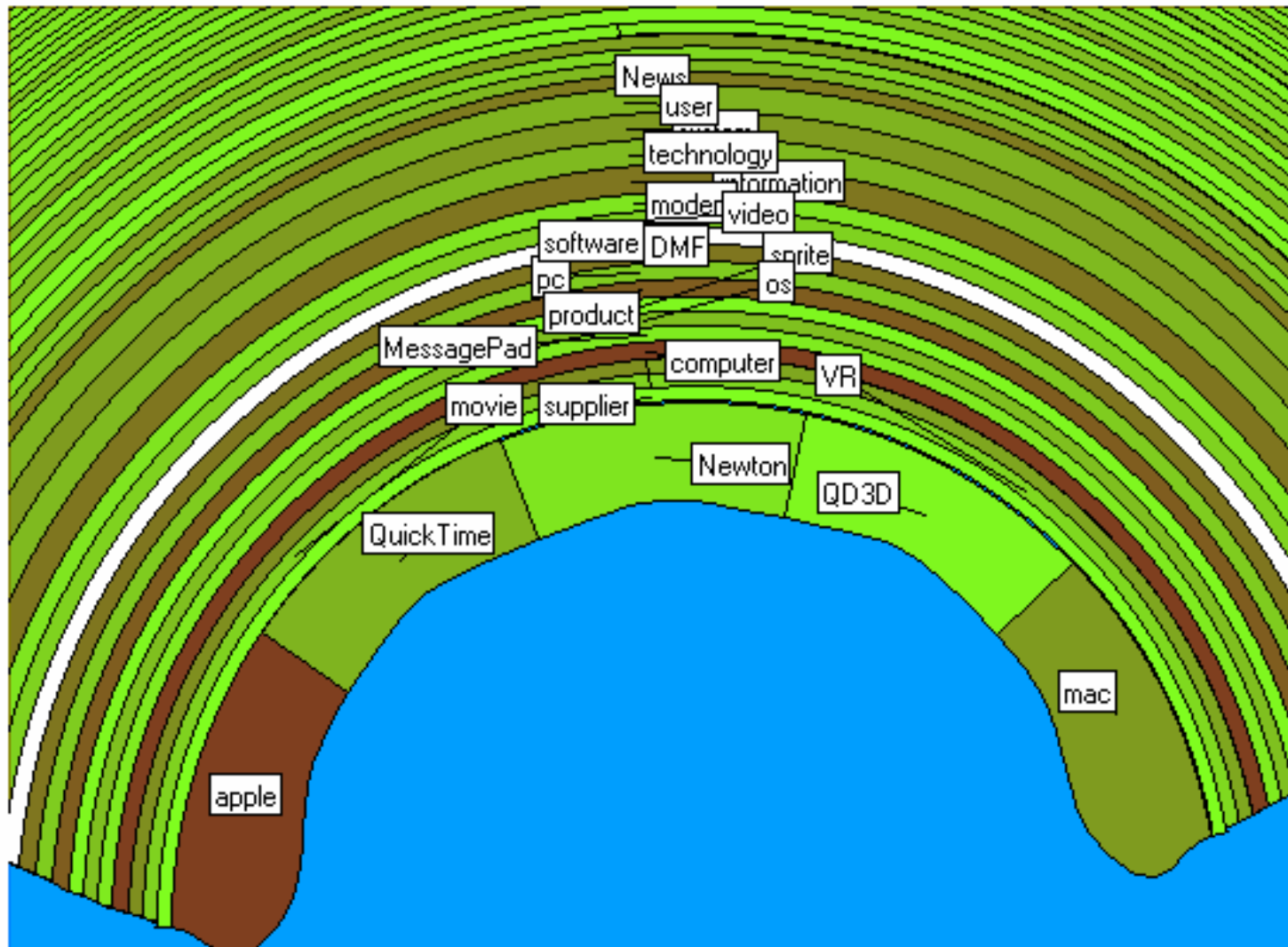
Case : company analysis

The dataset

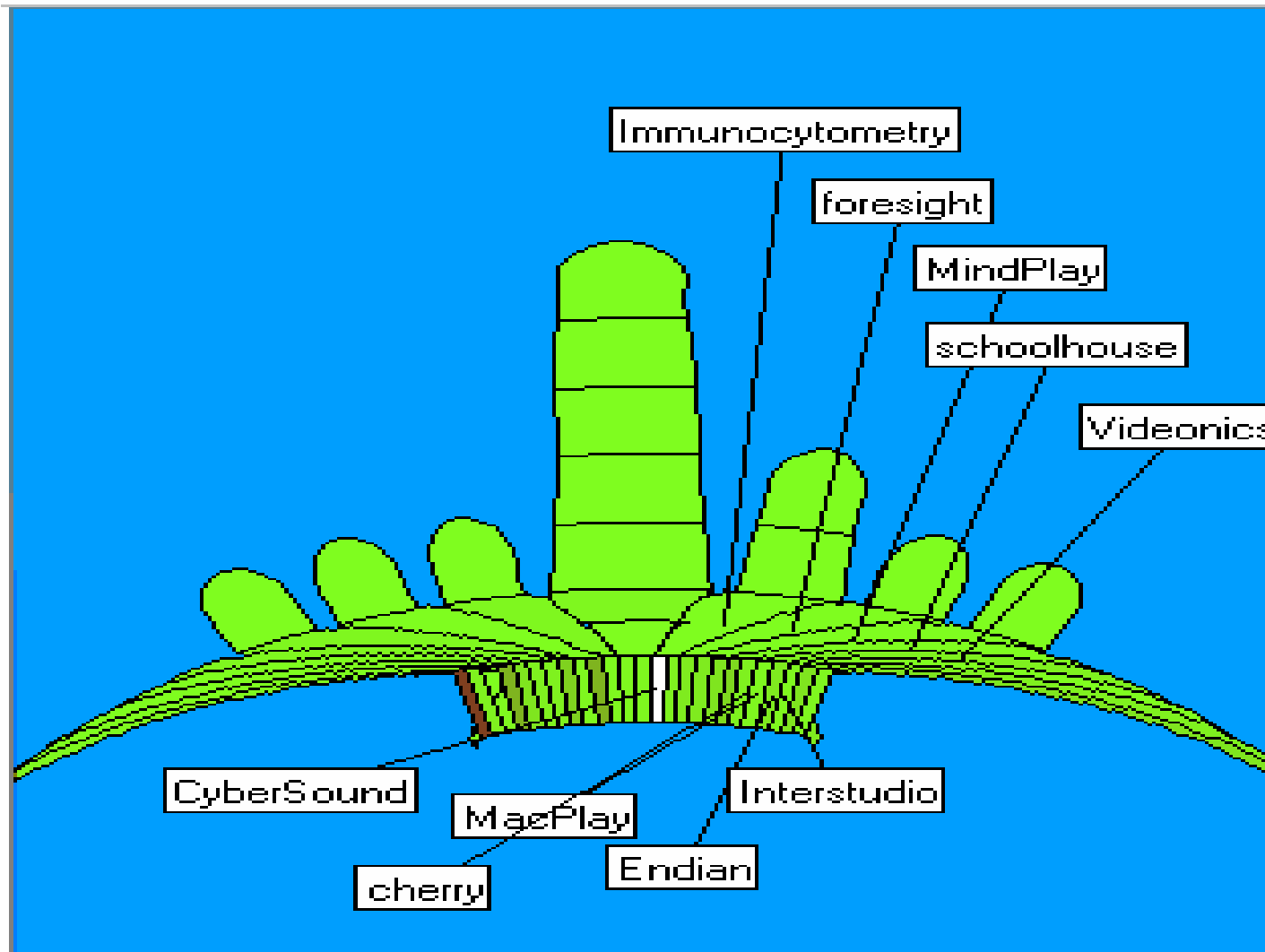
<u>Source Type</u>	<u>Size</u>
Competitor Website	20 Mo
General Press (online databases)	12 Mo
Specialized Press (scanned)	2 Mo
Web search engine	20 Mo
Newsgroup Archives	1 Mo
Internal Memos	10 Mo

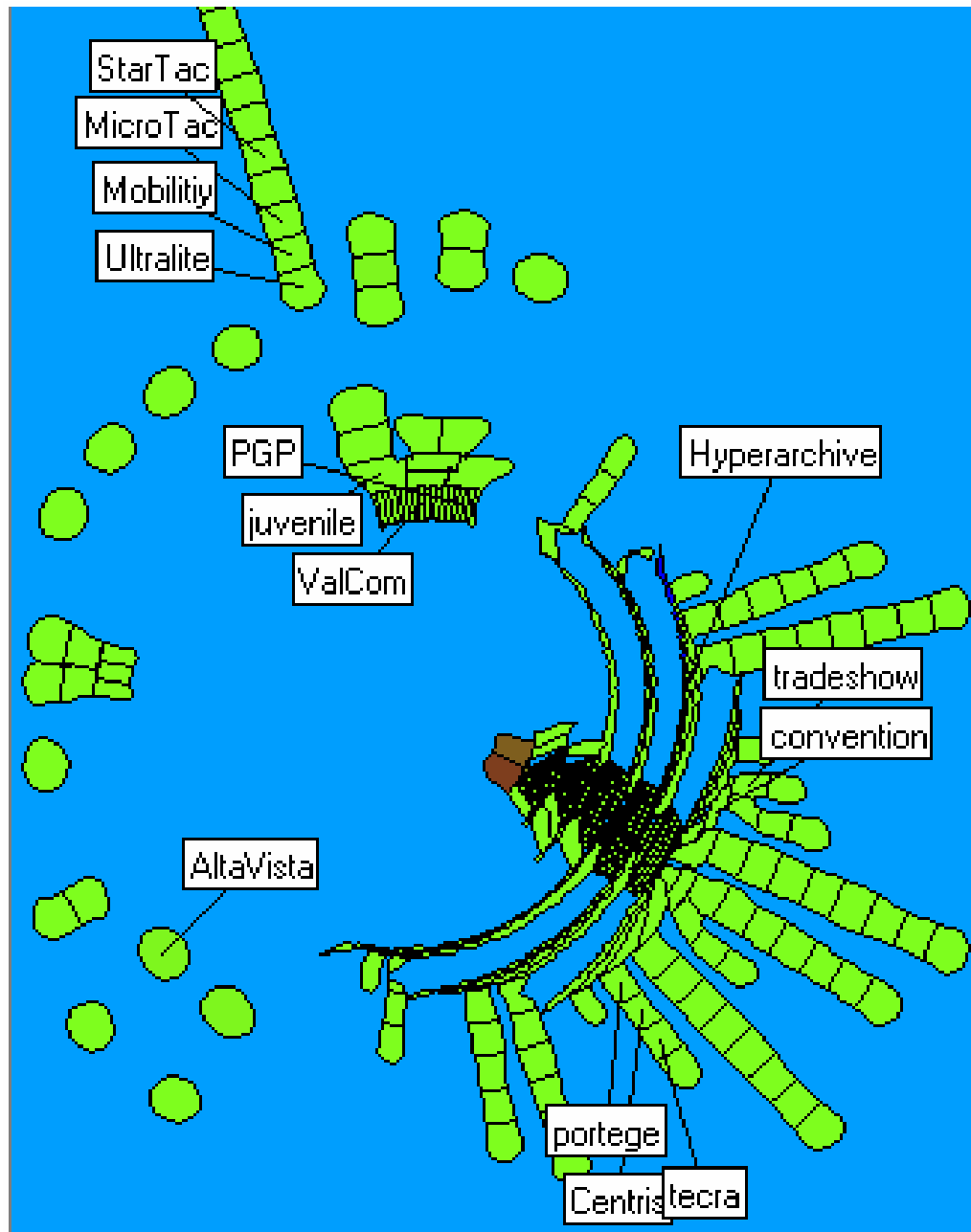
Note : size of text only, stripped of pictures, java applets, etc.
10 's of thousands of pages

Website data, Novice mode, Noise cleaned out, Global view



Websearch data, Expert view, Main theme analysis, Semi-local view





Websearch data,
Semi-expert mode,
theme segmentation #2

Results obtained

- Retail food business : Immediate identification of a weak signal of high significance (customer lawsuit to a competitor)
- Hotel industry : Debunking corporate talk (much talk about quality initiatives and human assets, but clearly not linked to operational activities)
- Media industry : Grasp of acquisition patterns and relationships.
- Manufacturing : Understanding the link between patents filed and development strategy.

The potential

- Company public image assessment
- New product perception analysis.
- Competitor product perception analysis.
- Market and technology monitoring.
- Decision makers ' speech analysis.
- Etc.

Lessons learned

- Not an intuitive representation. It puts many people off.
 - A plus for discovery and innovation
 - A minus : requires training and good conceptual skills.
- There are thousands of visualizations and interpretation that one can make from the same dataset. There is no one unique truth.
- The interpretation must be done by people who understand the business. The more expert the better. Unsuitable for juniors.
- Massaging the data is work. « Will users have the courage and patience to build their own freedom of interpretation rather than operating their browsers as they watch their TV sets ? » (Gartner Group review).
- Emerging trend : business information guerilla