



Processing data streams by relational analysis

Ilhème Ghalamallah

Institut de Recherche en Informatique de Toulouse, IRIT-SIG



Plan

- Introduction
- Tetralogie
- Proposition
- X-Plor
- Conclusion

Introduction	Tétralogie	Proposition	X-Plor	Conclusion
<p>In the business intelligence (BI) context, the majority of the strategic information comes from relational sources and the relevance of extracted knowledge usually depends on considering data evolution and their interactions</p> <p>The multidimensional approach (nD) may bring forth a solution to identify and understand the underlying structures or strategies. But non-expert users get easily lost.</p> <p>We have already a platform called Tétralogie that is specialized for strategic scanning and another tool called Xplor which is dedicated to business intelligence. As a consequence, we provide a unified system to generate and manage relational data and extract implicit knowledge, whose content and format are adapted to decision-makers that are not experts in nD or BI.</p>				
1/19				

Introduction	Tétralogie	Proposition	X-Plor	Conclusion
<p><u>Tétralogie is a tool particularly adapted to the macroscopic analyses (Dousset, 2003), from a corpus of documents collected for a precise subject. It is able to detect:</u></p> <ul style="list-style-type: none"> ▪ Strong signals , ▪ Weak signals , ▪ Significant tendencies . <p><u>The elaborate information</u> results, represents a synthesis obtained by various methods of data analysis and diffused via graphic visualizations</p>				
2/19				

Introduction	Tétralogie	Proposition	X-Plor	Conclusion
--------------	-------------------	-------------	--------	------------

Tétralogie output visualisation

global strategic Aspect

3/19

Introduction	Tétralogie	Proposition	X-Plor	Conclusion
--------------	-------------------	-------------	--------	------------

System Xplor with perspective to **automate the on-line processing of relational information** and to propose analysis and navigation tools oriented to **business intelligence (BI)**.

- System provides strategic analyses on corpora of textual information resulting from the most various sources.

4/19

Introduction	Tétralogie	Proposition	X-Plor	Conclusion
--------------	-------------------	-------------	--------	------------

Information sources :

Invisible Web
Content from specialized web databases, not available through usual search engines (Patents, etc.)

Web Sites
Usual Internet Web Sites. Companies, universities, institutions websites, etc.

Web Pages
This source allow you to monitor web pages. You are alerted when a page is modified.

Newsgroups
Discussion groups using the NNTP protocol (not available through web browser).

Mailing lists
Internet users discussions (via e-mail).
BD métiers et serveurs prof.

Web News
News feeds available on specialized Internet sources.
Press releases, etc.

Flux d'informations (presse...)

5/19

Introduction	Tétralogie	Proposition	X-Plor	Conclusion
--------------	-------------------	-------------	--------	------------

Enables **decision makers** to perform their own **queries** and to interpret graphical output without the need for an **analyste**.

**Decision makers,
Experts**

6/19

Introduction	Tétralogie	Proposition	X-Plor	Conclusion
--------------	------------	--------------------	--------	------------

the BI is defined as the group of the coordinated actions for research, treatment, distribution and protection of information that are useful for the economic actors, taking in consideration their individual and collective strategies.

According to the report Martre

- **Business intelligence (BI) tools** enable organizations to understand their internal and external environment through the systemic acquisition, collection, analysis, interpretation and exploitation of informations (Chung, 2002).

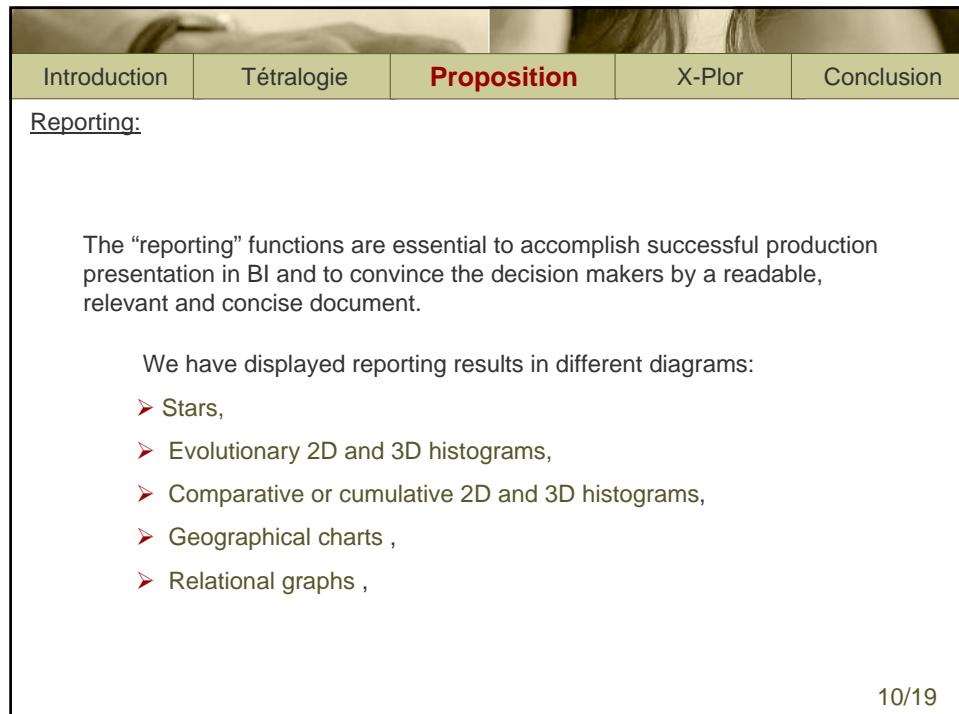
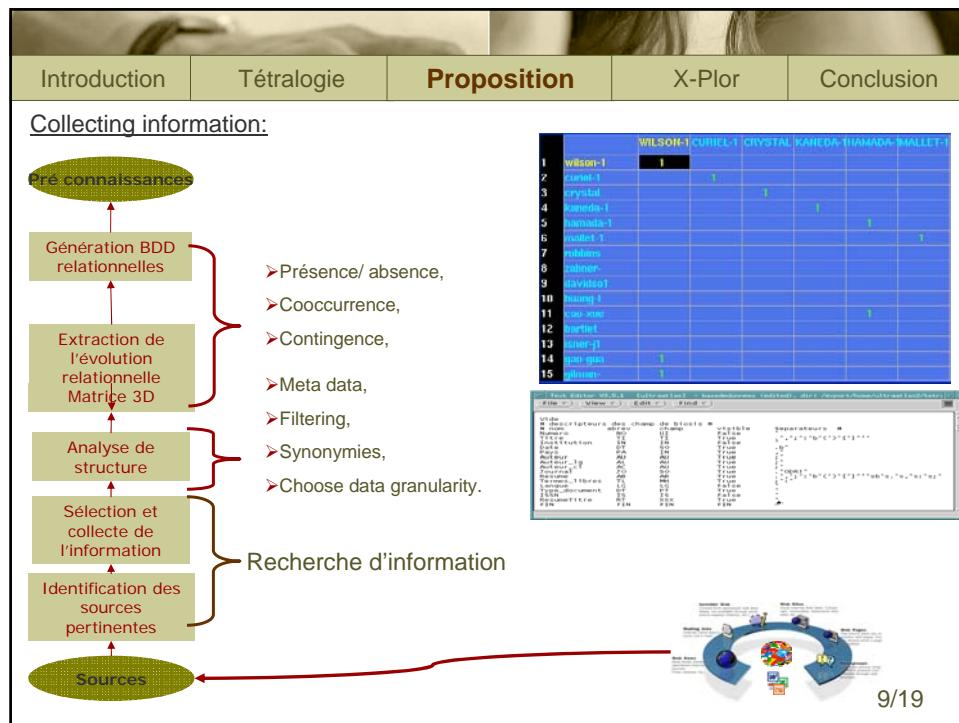
Exple: analyse de dépôts de brevets

7/19

Introduction	Tétralogie	Proposition	X-Plor	Conclusion
--------------	------------	--------------------	--------	------------

The diagram illustrates the BI Cycle as a continuous loop. It starts with 'Requirements formulation' (with an image of a person working at a computer), which feeds into 'Collecting information' (with an image of hands typing). This leads to 'Analysis' (with an image of a cylinder representing a DataWarehouse). 'Analysis' then feeds into 'Elaborated data' (with an image of a world map and charts). Finally, 'Elaborated data' leads back to 'Reporting' (with an image of a person working at a computer), which completes the cycle. Arrows indicate the flow from one stage to the next, and double-headed arrows connect 'Requirements formulation' and 'Collecting information' to the central cycle.

8/19



Introduction	Tétralogie	Proposition	X-Plor	Conclusion
--------------	------------	-------------	--------	------------

Experimentations:

The star represents the analysis theme (authors) that is extracted from a relational database, and shows all the elements related to them in function of date (publications, authors, journal, country, affiliation).

11/19

Introduction	Tétralogie	Proposition	X-Plor	Conclusion
--------------	------------	-------------	--------	------------

Expérimentations:

Les Auteur_c

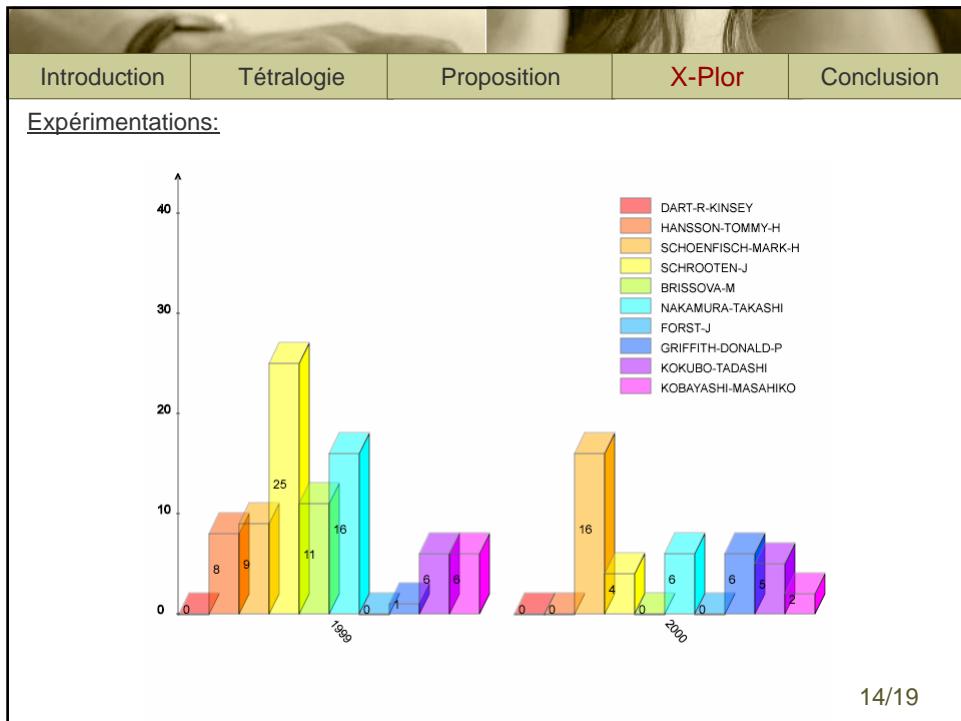
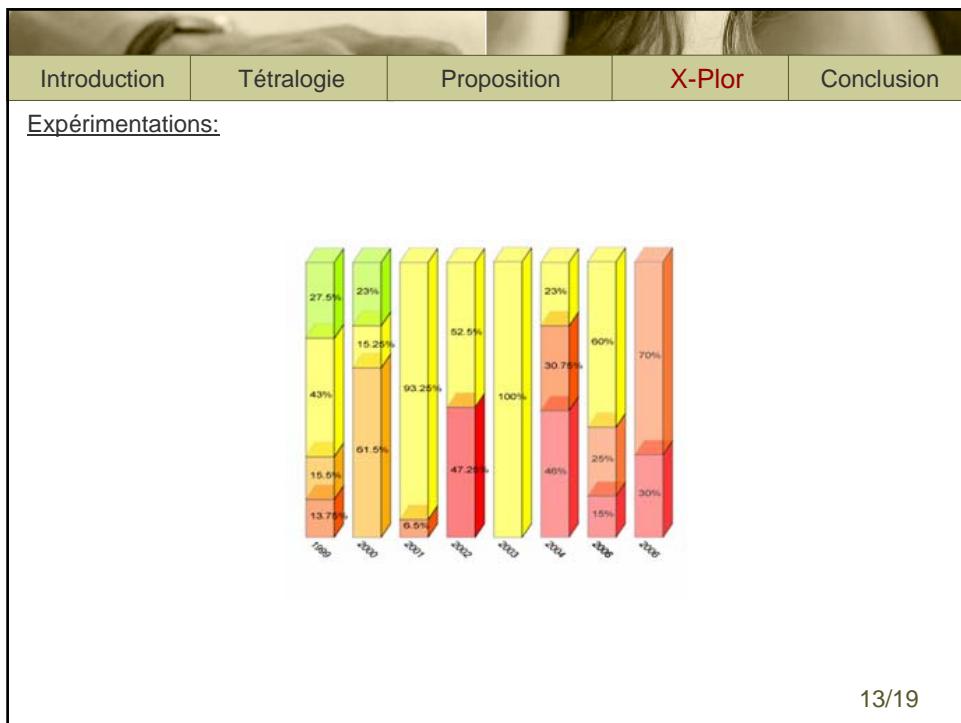
Legend:

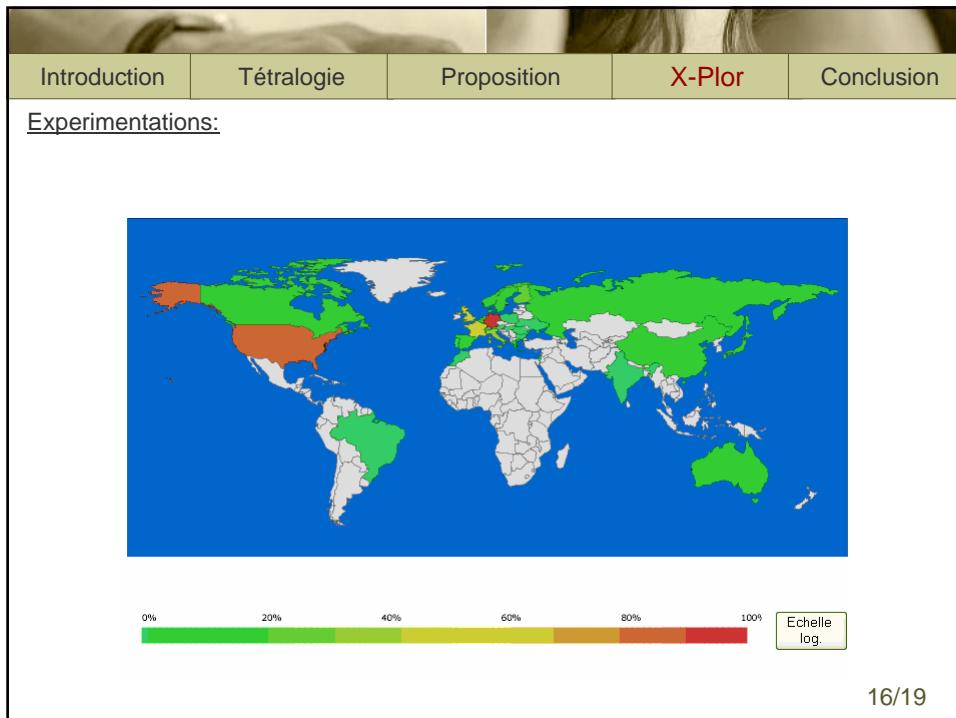
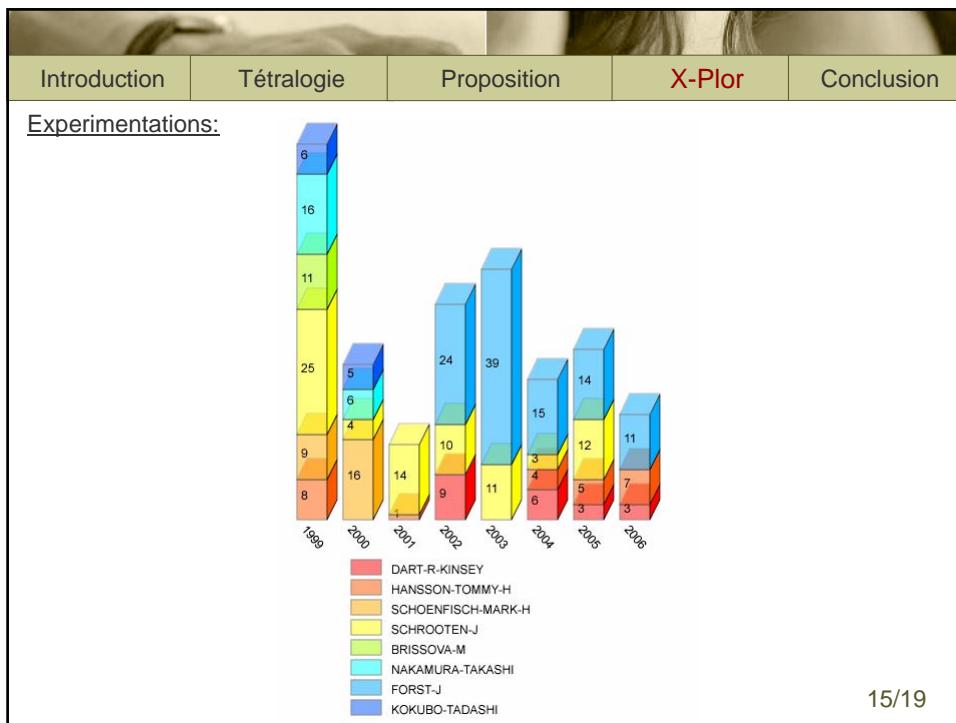
- Pourc du max 32%, avec 88 occ
- Pourc du min 2%, avec 7 occ
- la moyenne : 26.7

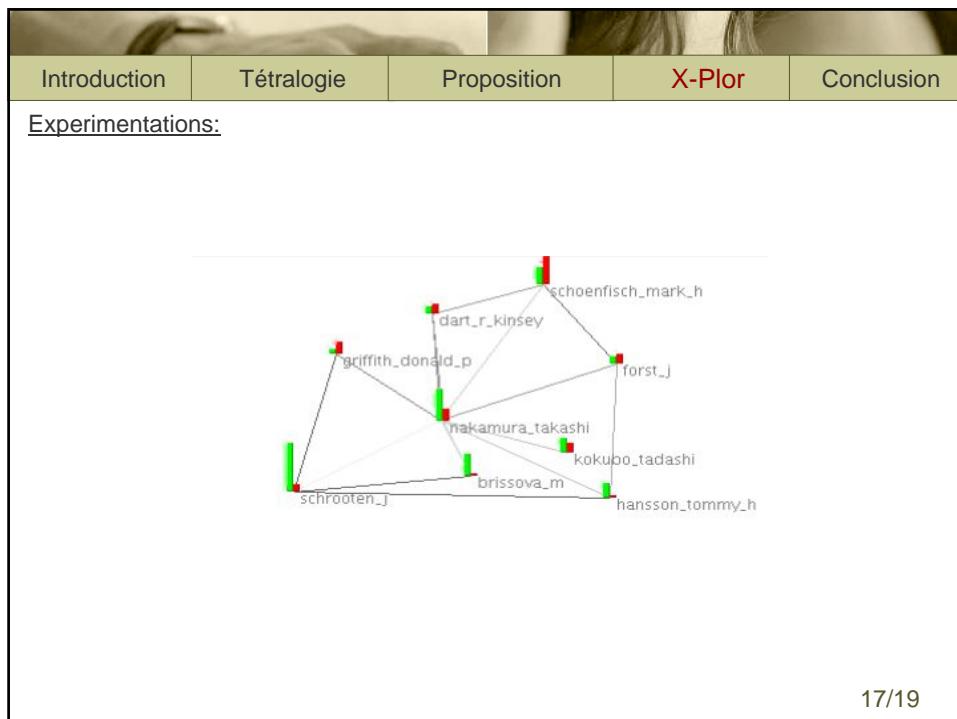
Author distribution percentages:

Author	Percentage
FORST-J	33%
SCHROOTEN-J	30%
SCHOENFISCH-MARK-H	8%
NAKAMURA-TAKASHI	4%
KOKUBO-TADASHI	4%
BRISSOVA-M	3%
HANSSON-TOMMY-H	3%
KOBAYASHI-MASAHIKO	3%
GRIFFITH-DONALD-P	3%
SAWADA-MASANORI	2%

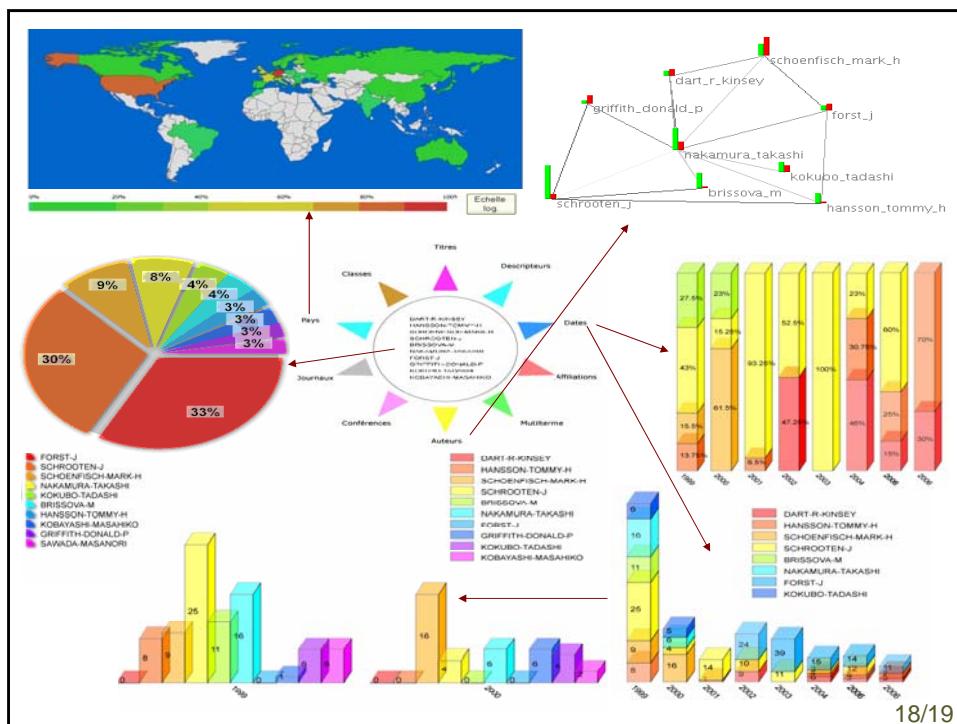
12/19







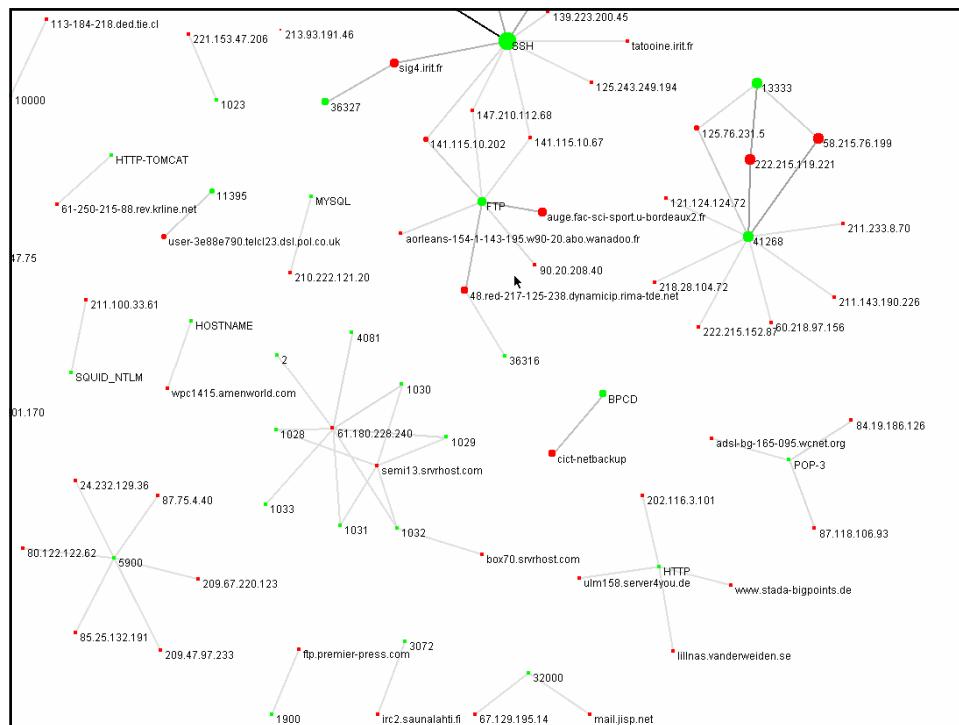
17/19



18/19

Introduction	Tétralogie	Proposition	X-Plor	Conclusion
<u>Experimentations:</u>				
Num: 1 Number: 9328 Date: 5Feb2007 Nomjour: lundi NumJ: 5 Mois: Feb Annee: 2007 Time: 2:17:39 Heure: 2 Minute: 17 Seconde: 39 Source: AOrleans-154-1-143-195.w90-20.abo.wanadoo.fr Destination: atlas-dmz Service: ftp Action: Accept				
17/19				

Introduction	Tétralogie	Proposition	X-Plor	Conclusion
<u>Experimentations:</u>				
Num: 8 IP: 88.121.182.114 Jour: Mardi J: 30 Mois: Jan Annee: 2007 Heure: 01 Minutes: 13 Secondes: 01 Fichier: /IMAGES/firework/FLACCUEIL1.gif ER: 200 2536 Tps: 1				
17/19				

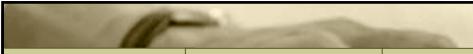


Introduction	Tétralogie	Proposition	X-Plor	Conclusion
--------------	------------	-------------	--------	-------------------

Perspectives:

- continue experimenting,
- Formalize data stream model for our system.

19/19



Introduction	Tétralogie	Proposition	X-Plor	Conclusion

Choo, C. W. (1998). "The Knowing Organization". Oxford: Oxford University Press.

Dousset, B. (2003). "Intégration des méthodes interactives de découverte de connaissances pour la veille stratégique", Mémoire d'habilitation à diriger les recherches, Université Paul Sabatier, Toulouse.



Thank You