

Creation of a BI environment from scratch with open source software

A practical case

EUNIS 2012 - Umeå, Sweden

Thierry DAUTCOURT

Inria's Research Centres



Inria PARIS - Rocquencourt



Inria LILLE
Nord Europe



Inria NANCY
Grand Est



Inria SACLAY
Ile-de-France



Inria GRENOBLE
Rhône-Alpes



Inria SOPHIA ANTIPOLIS
Méditerranée



+200 Research's Teams

+3500 Researchers

About 4500 « employees»

Inria RENNES
Bretagne
Atlantique



Inria BORDEAUX
Sud-Ouest

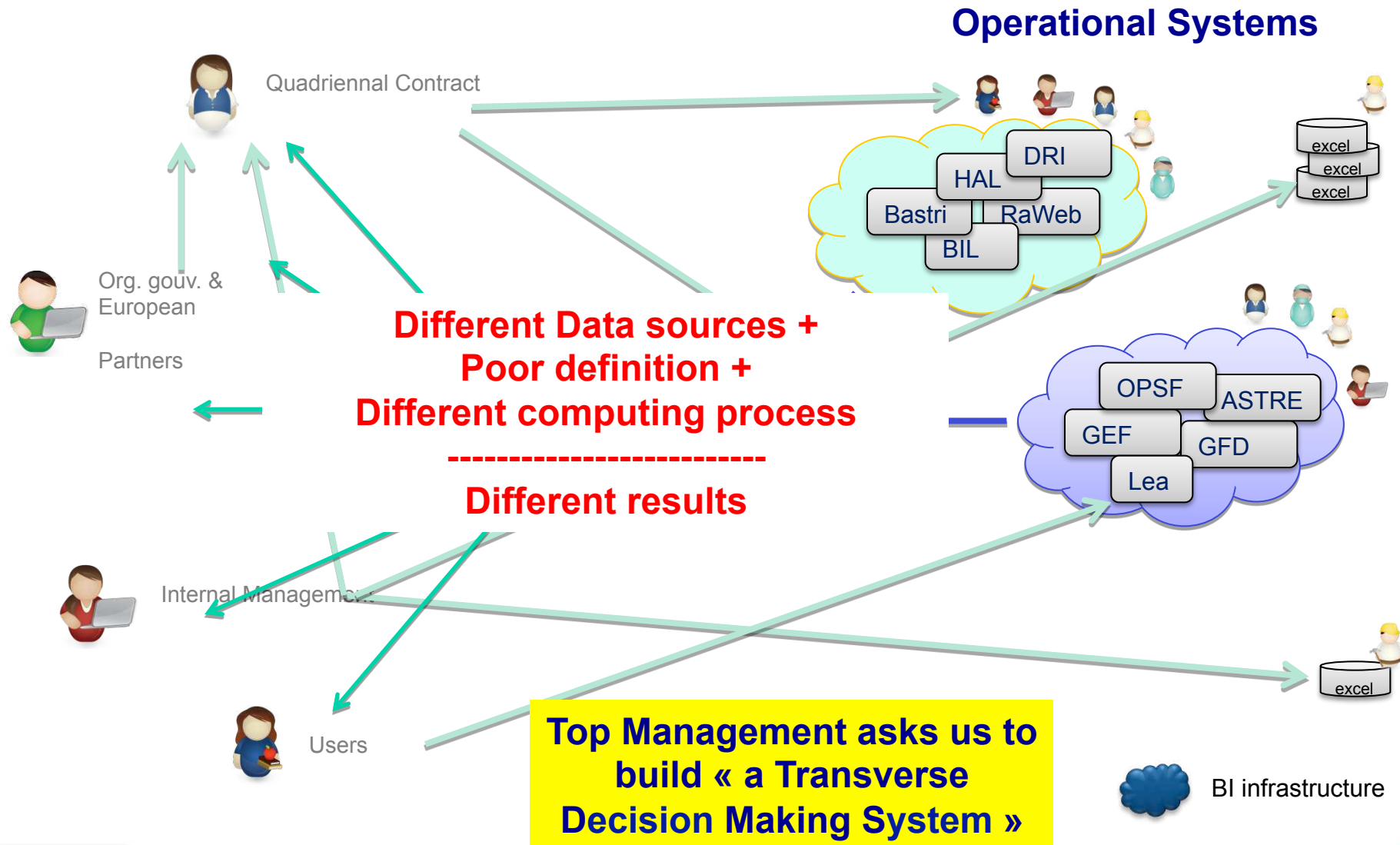


Creation of a BI environment from scratch with open source software

Agenda

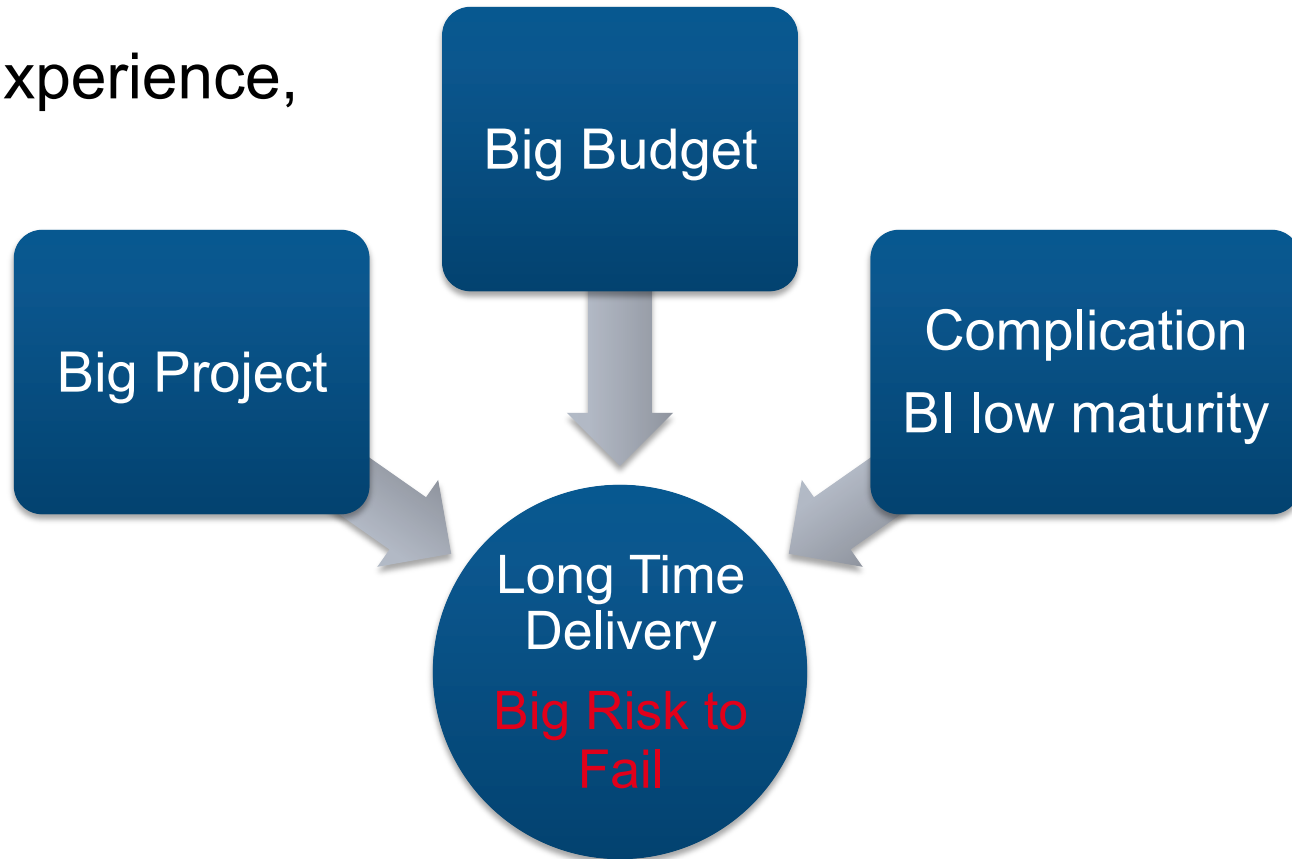
- 1- Context and request
- 2- Strategy adopted
- 3- Selection of Open Source Software
- 4- Results and Future

1- Context and Request



1- Context and Request

By Experience,



How to proceed ?

Choose one key indicator

Find people to work with

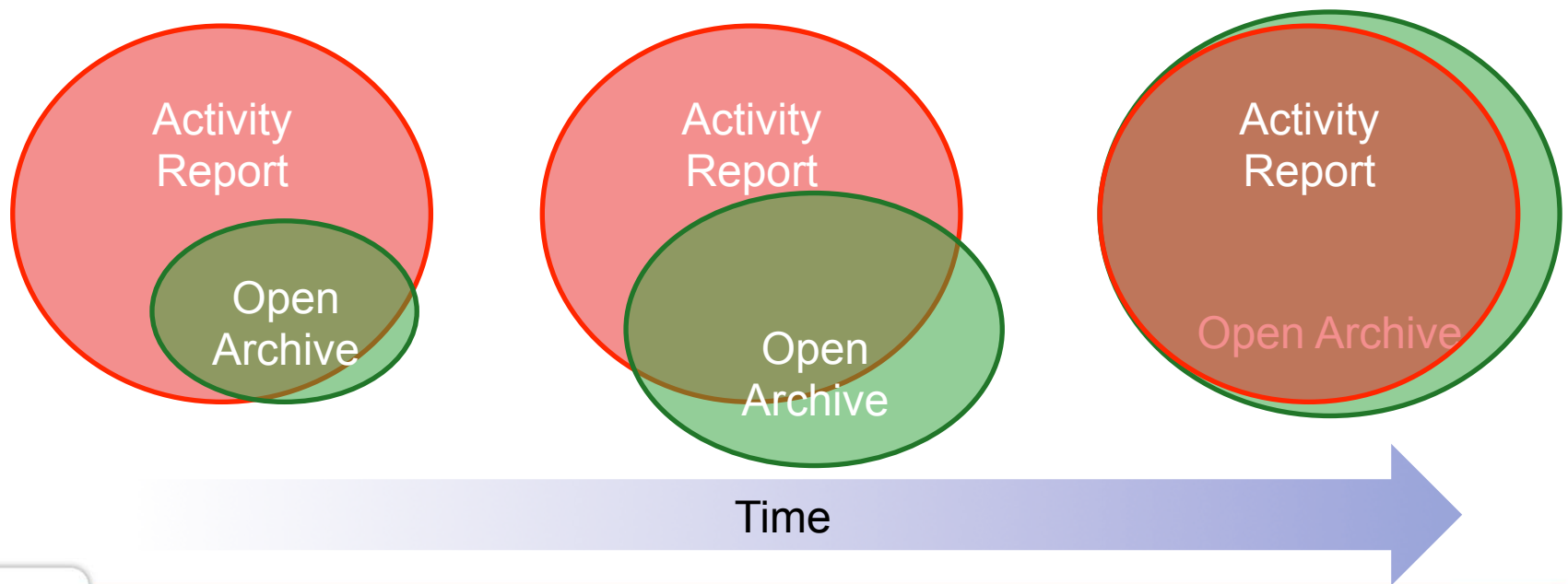
Choose and build a technical BI environment

Strategy adopted

2.1- Focus on one indicator

- A key indicator not previously computed,
- A good understanding of the Organism (and strategy)

- Berlin Declaration Involvement
- Measure the publications available in the open archive -> goal 100%

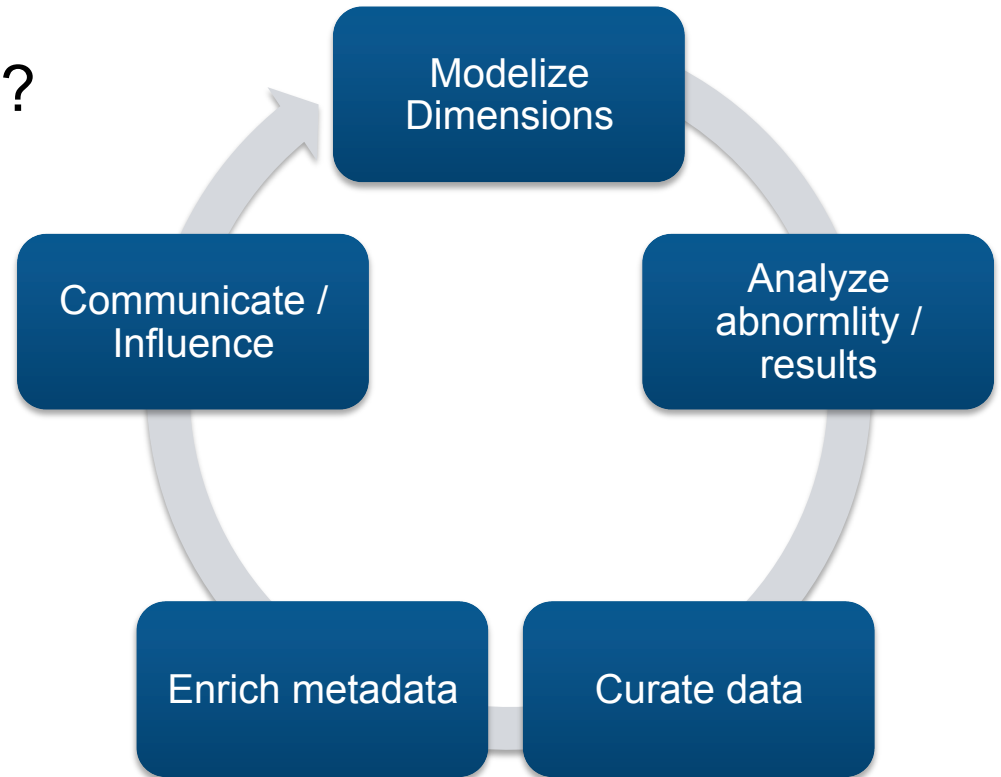


Strategy adopted

2.2- Find people to work with

Subject is publication
Librarians are a relevant choice

What are their tasks ?



Strategy adopted

2.3- Need a technical frame as support

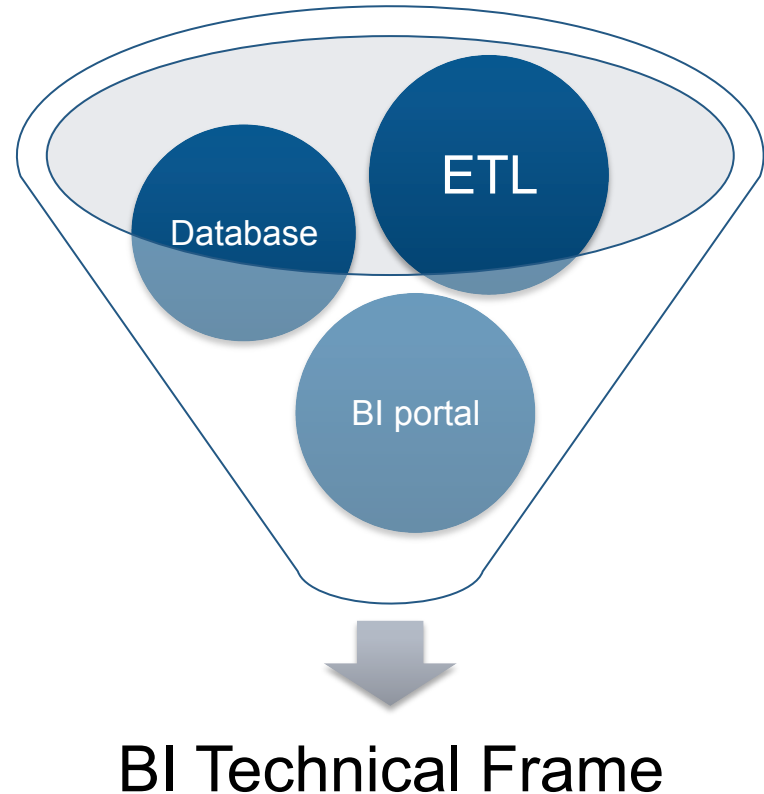
2 main Functions

- Create and Manage DW,
- Offer a Front End Application

Questions

Proprietary or Open Source Software

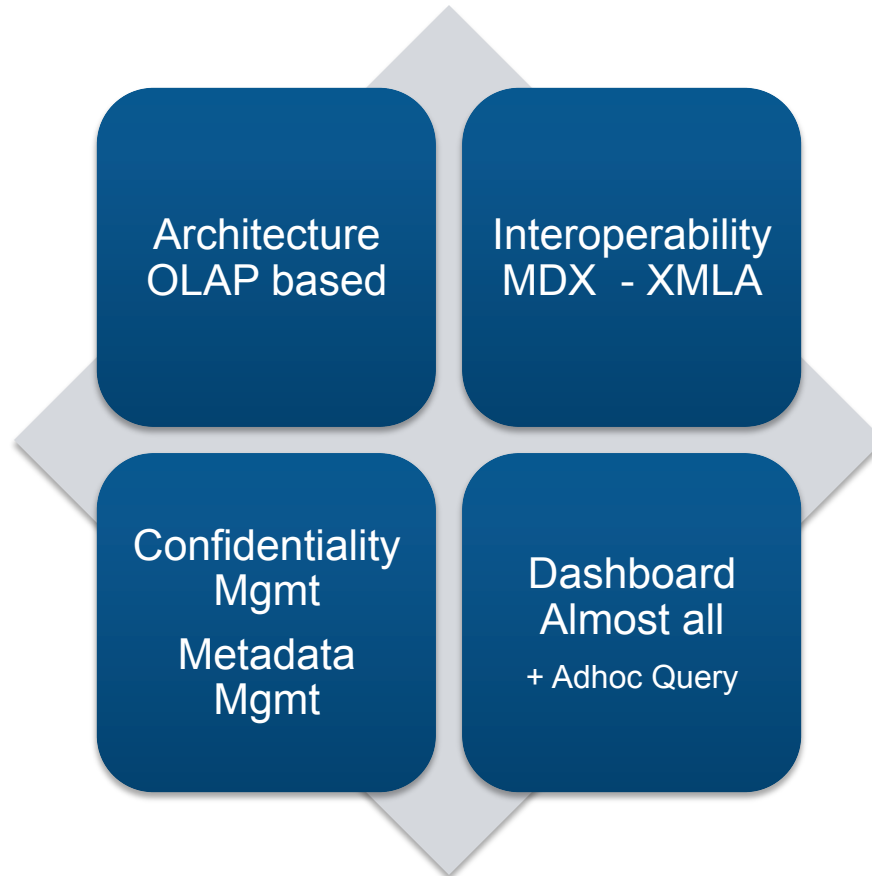
- What are the **risks** ?
- Where is the **value** ?



Focus BI tools : 2010 excerpt of comparison

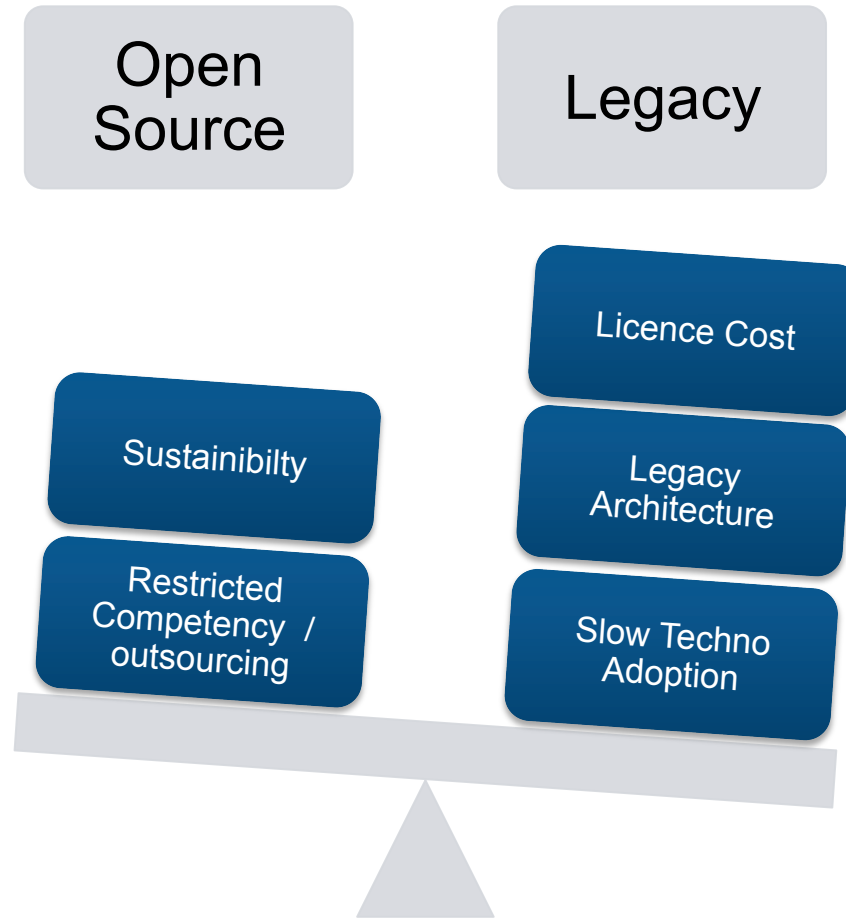
	DB	OLAP	Analytics & Retrieval		Dashboard
SAP-BO	-	InfoCube (ROLAP)	BO voyager	BO, BEX	BO Xcelsius Dashboard Builder
ORACLE	11G	Oracle OLAP Hyperion Essbase	Essbase via excel	OBI EE	Interactive Dashboard
IBM	DB2	Cognos 8 Infosphere WareHouse, DB2 OLAP server	Cognos8	Cognos8	Cognos8
Microsoft	Sql Server	SSAS	Performance Point server	SSRS	Performance Point Service
SAS	Legacy	Legacy SAS	SAS WRS	SAS WRS	Sas Performance Management
Pentaho	MySql, Postgres, Oracle...	OLAP Mondrian	Jpivot	Pentaho Ad hoc Query	Pentaho Dashboard
SpagoBI	Mysql, Postgres,..	OLAP Mondrian, SSAS	Jpivot, jpalo	QBE	Dashboard interactive and KPI
Jasper BI suite	Mysql, Postgres,..	?	Jasper Analysis	Jasper Report	Jasper Report Dashboard

BI functions : short summary



BI tools : Open Source vs Proprietary software

What are the Risks ?



BI environment : Open Source vs Proprietary

And what are the Long Term Value ?



Results

Indicateurs	annee			
	2010	2011	2012	2013
	Mesures	Mesures	Mesures	Mesures
	pourcentage	pourcentage	pourcentage	pourcentage
– Publications de référence	81 %	83 %	86 %	100 %
Actes de colloques	82 %	83 %	85 %	100 %
Livres	63 %	66 %	83 %	100 %
Revue et chapitres de livres	80 %	83 %	87 %	100 %

**In 2013
Goal hit**

Slicer: [source=RaWeb]

On publications, more than 15 analytic dimensions are available

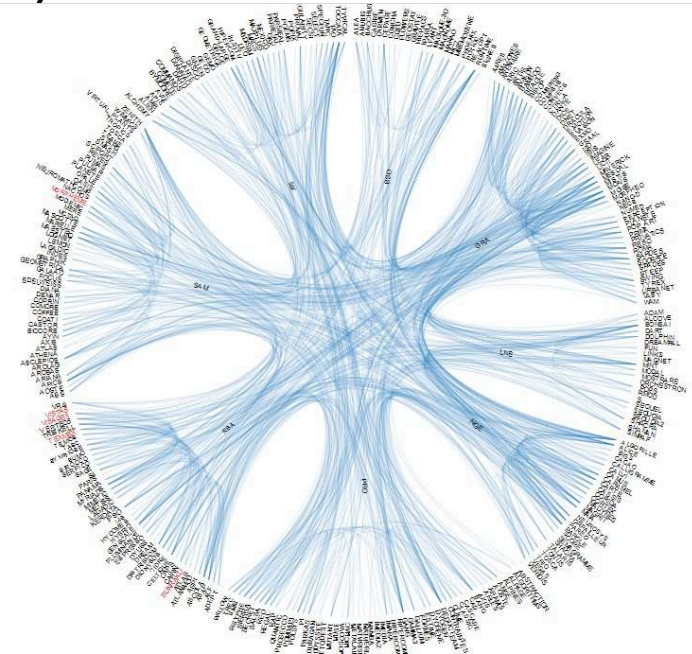
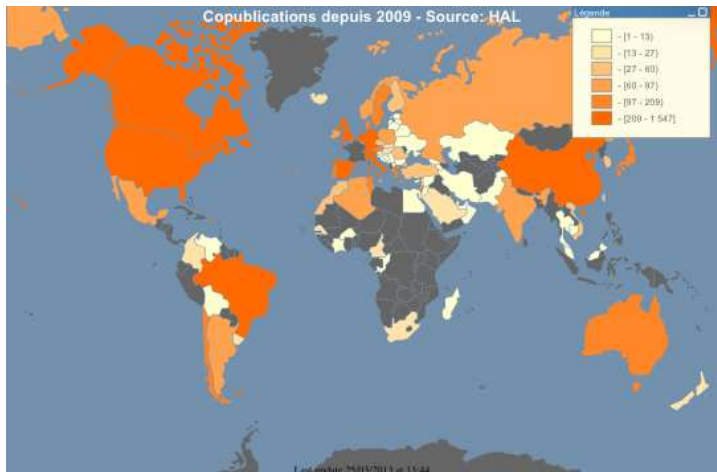
- Time dimension
- Organisational dimension
- Document type (congress, article, thesis, ...)
- Geographical (by affiliation of authors)
- ...

Results

From 2010 until now

More than 10 subjects by incremental building

- ✓ Publication (type, author, affiliation, ...)
- ✓ Funding on research programs :
International, European, National, ...
- ✓ Mobility Research Exchange (in/out)
- ✓ Software, ...



And create different graphic representations with analytic data

And now, where to progress

Find strong(er) Sponsorship resources, new roles defined,

Improve communication and large diffusion

Capitalize on services,

**Datawarehouse is not only a media for indicators,
it also models IT, as provider of digital services.**

**The effort made on the quality of data has to
benefit all the users.**

Thanks for your listening

Comments and/or Questions ?