

SpagoBI 5

Turn your data into knowledge

Technical session





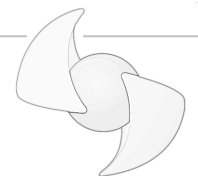
Agenda

- Introduction
- Engineering and Open Source
- SpagoBI: a unique solution
- SpagoBI overview
 - Analytical features
 - Live demo
- Hot topics
 - Big data
 - What-if
 - Mash-up & federation
 - Cloud
 - Data mining
 - Social listening
- Case studies, with real user experience
- Roadmap and next steps
- Make your own solution





Engineering & Open Source





Engineering Group



(Wilmington)

USA

(São Paulo / Rio de Janeiro / Recife Belo Horizonte / Curitiba)

BRAZIL

31 branches in **ITALY**



(Buenos Aires)

ARGENTINA

A global player

Business integration
Consulting
Outsourcing
Products and solutions

(Brussels)

BELGIUM

(Belgrade)

REPUBLIC OF SERBIA

about 7,300

Professionals

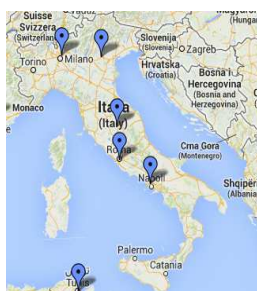
1,000

Large accounts

7.2%

Italian market

822.8 mn€



Italy: Bologna,
Milano, Napoli,
Roma, Padova,
Palermo e Torino



**New York, NY,
USA**



**São Paulo, SP,
Brazil**



**Brussels,
Belgium**

OSS Competency Center & SpagoBI Labs



www.spagoworld.org

New! SpagoBI Lab in Belgrade, Serbia





Open Source Landscape

Economically efficient

Technically efficient



Strategically efficient



Socially efficient





Users' expectations

Users expect market-ready offerings (i.e.: code complemented by: packaging services, training, maintenance, support, etc.).

Users want a full business proposal, not just bare code.

Code is only a fraction of the software value-chain that delivers market-ready offerings.





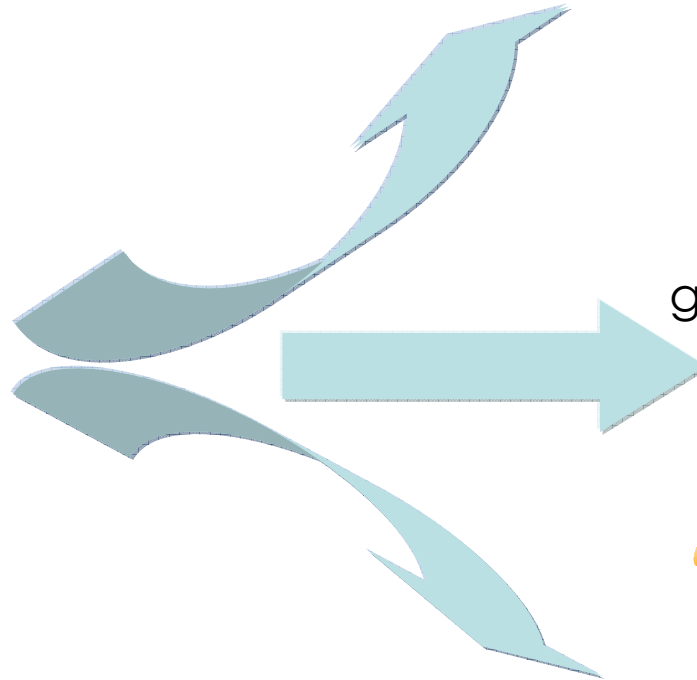
OpenSource@Engineering

INTEGRATOR

open source adoption



Competitive lever



PURE PLAYER

global communities

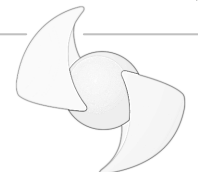


INNOVATOR

R&D collaborative projects



**DIGITAL AGENDA
FOR EUROPE**





SpagoBI : a unique solution





A unique solution

- The ONLY entirely open source Business Intelligence solution **for enterprises**
- The ONLY BI solution without any mandatory cost
- The ONLY BI product that can be embedded into others without licence problems or additional costs (IVS/OEM model)
- UNIQUE open source answers for:
 - Visual **query builder** for different source type (rdbms, file, WS)
 - User friendly **OLAP** client
 - **Big data** analysis
 - **What-if** analysis
 - Self-service **in-memory cockpit** with data mash-up
 - **Location Intelligence**
 - **Social listening**, sentiment analysis, campaigns evaluation





Open source for enterprise

- Technical points:
 - User profiling, visibility rules
 - Functionalities for administrator
 - Scalable
 - Portable
 - Multitenant
 - BI lifecycle
 - Cache management
- Other:
 - Full-services guaranteed
 - Bug tracker
 - Trouble ticketing
 - Engineering Group
 - OW2 consortium





Analytical features





Analytical documents' scope





Analytical documents' scope

- **Ready-to-use**

- Report
- Chart
- Interactive cockpit
- KPI
- Mobile

- **Build-by-yourself:**

- Ad-hoc reporting
- Self-service BI
- In-memory cockpit

- **Explorative**

- OLAP cube
- Visual Inquiring

- **Advanced Analytics:**

- Location Intelligence
- Data mining
- RT console
- Network analysis
- What-if
- Social listening






Live demo










User Name:


Password:

Login →


Business User
biuser/biuser


Showcase User
bidemo/bidemo


Administrator
biadmin/biadmin





Hot topics



Hot topics

- Hot topics
 - Big data
 - What-if
 - Mash-up & federation
 - Data mining
 - Social listening
 - Cloud





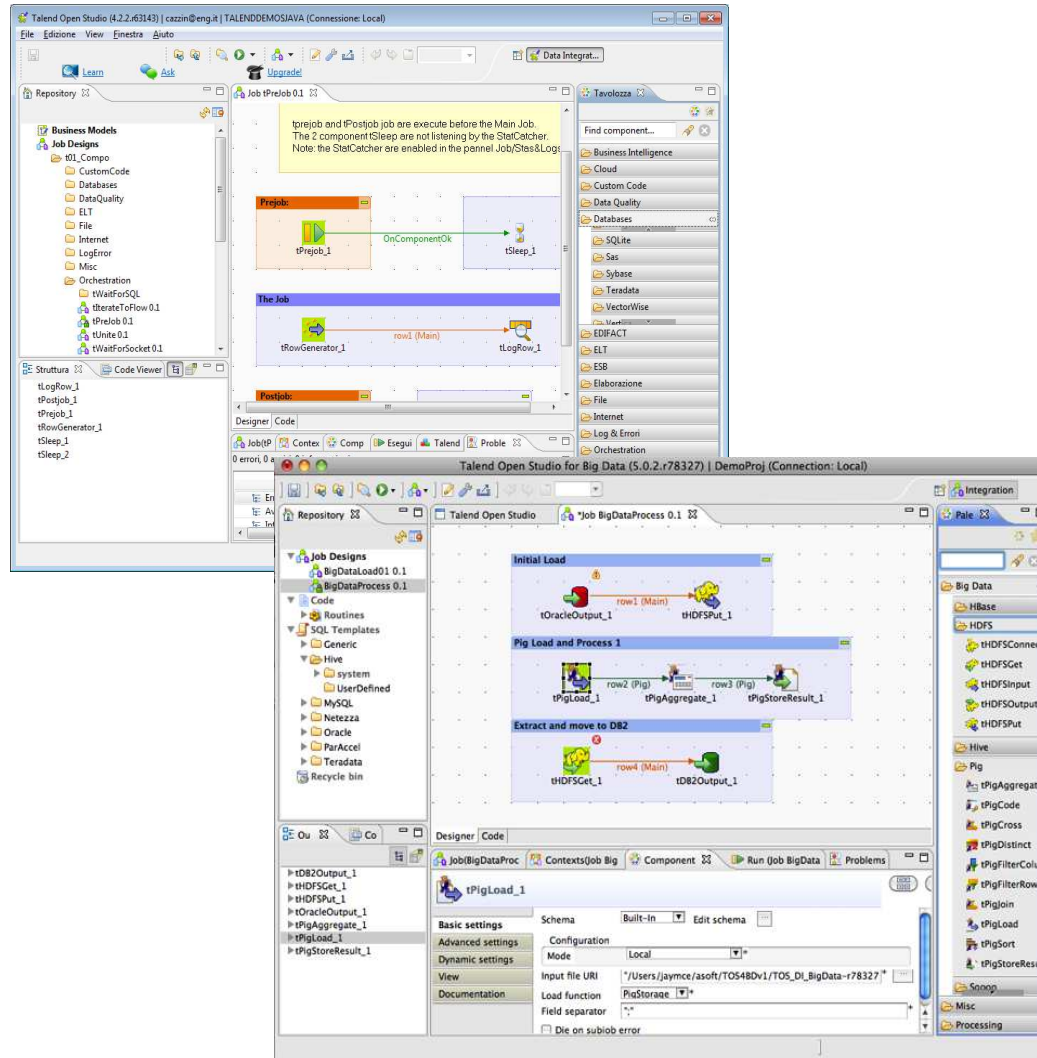
Big data & Business Intelligence

- Tasks:
 - Manage big-data (ETL) → Talend
 - Read, interpret and show big-data (BI) → SpagoBI
 - Big-data and real-time (BI) → SpagoBI





ETL: Analytical databases & appliances



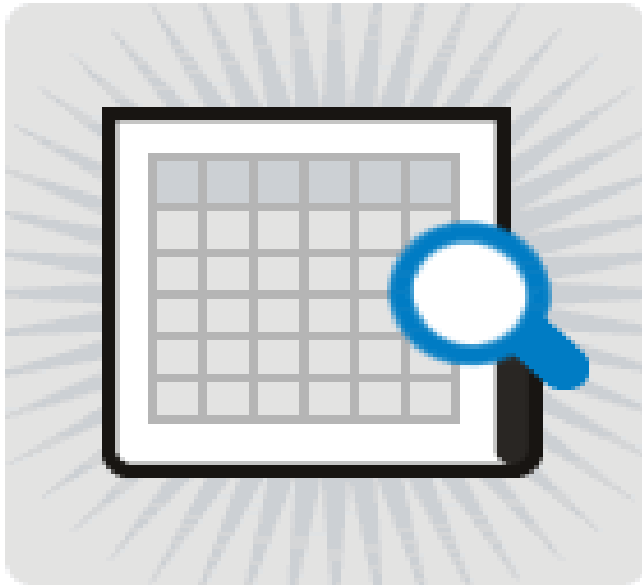
Connectors from/to:

- Greenplum
- Netezza
- Sybase
- Teradata
- VectorWise
- Vertica
- HDFS
- HBase
- Hive
- Cassandra
- MongoDB





SpagoBI - load



Certified appliances:

- Teradata
- VectorWise

Connectors from:

- Cassandra
- HBase
- Hive
- Impala
- Hadoop

.....
RT with:

- Storm
- WS02

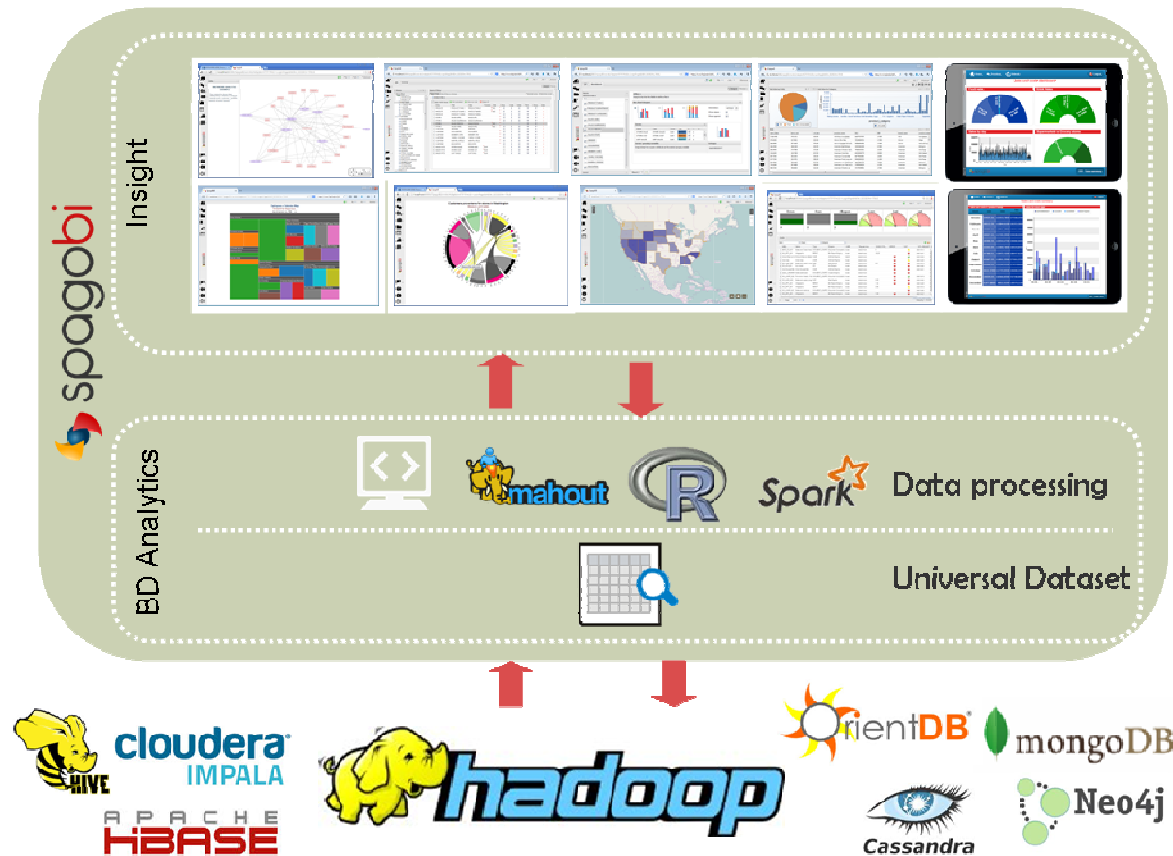
More:

- Scheduled data-set
- In-memory data set





Big data analysis

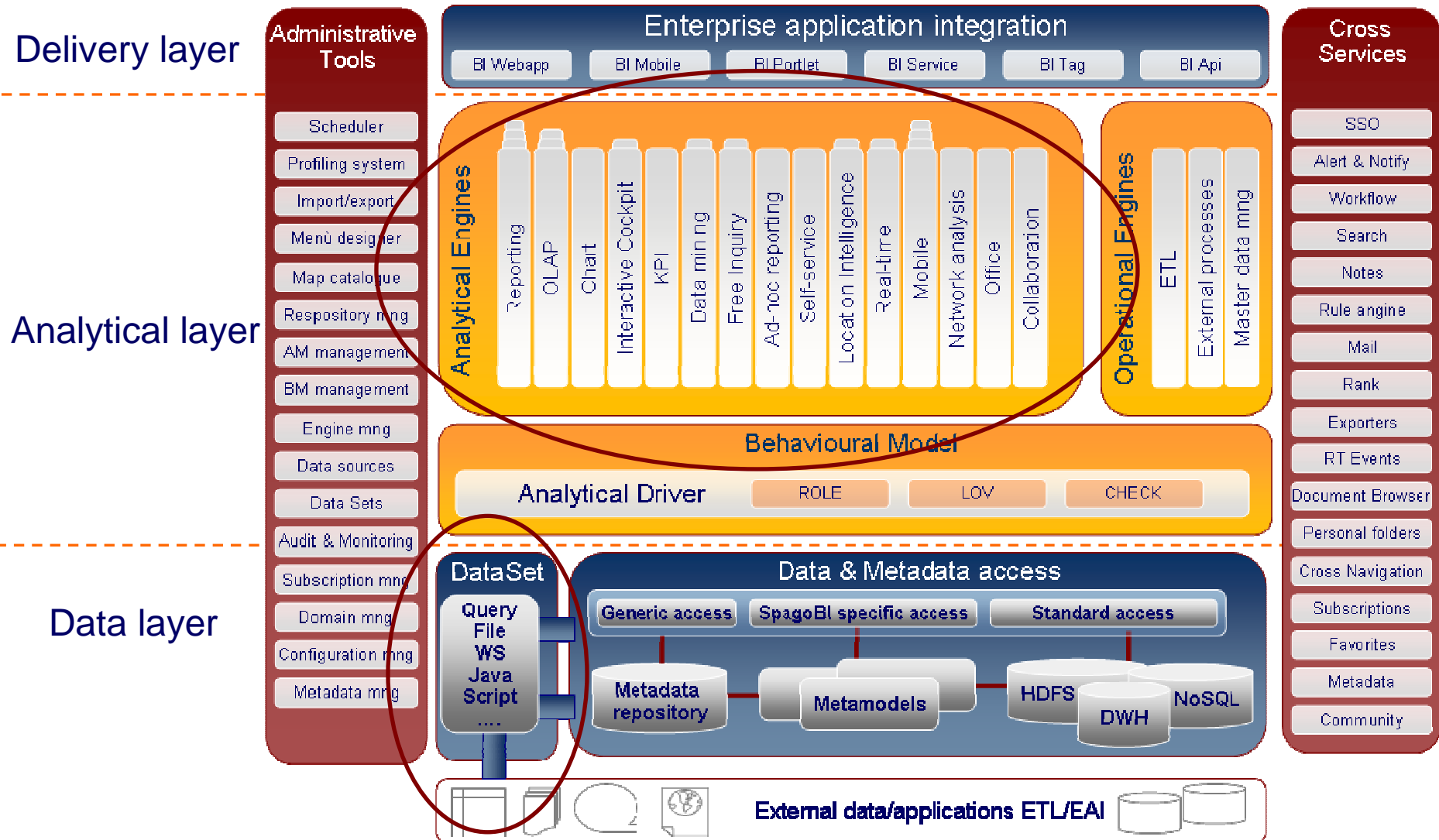


- DataSet component drives big data as data towards analytical engines
- DataSet can be processed by advanced techniques (mining)
- DataSet immediately opens new doors
- New data source → a new implementation for the dataset → every SpagoBI solution immediately works on it
- Next steps:
 - Data federation
 - RT
 - sandbox





SpagoBI Server Architecture





Intuitive what-if solution

The screenshot shows the SpagoBI web interface. The main data table displays sales metrics for various products. Two callout boxes highlight the 'what-if' functionality by showing how a user can modify a value in the table and see the resulting changes propagate through the data hierarchy.

Callout Box 1 (Top Right):

| Store Cost | Store Sales | Unit Sales |
|------------|-----------------|------------|
| 311,931 | 155240.74 + 30% | 365,852 |
| 12,786.96 | 6,755.24 | 14,979 |

Callout Box 2 (Bottom Left):

| Store Cost | Store Sales | Unit Sales |
|------------|-----------------|------------|
| 311,931 | Time.2011 + 20% | 365,852 |
| 12,786.96 | 6,755.24 | 14,979 |

Main Data Table (Partial):

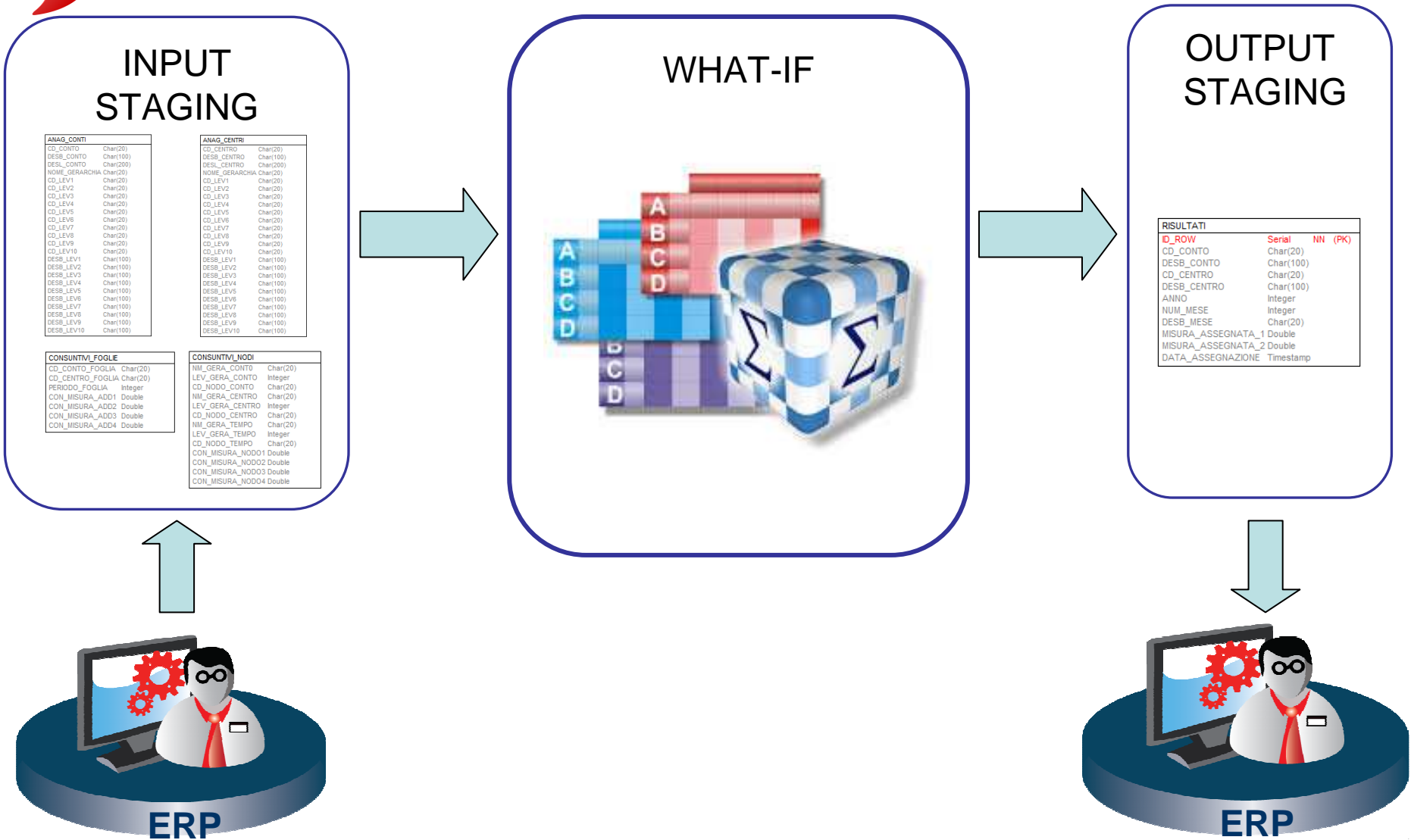
| Product | Sales Count | Store Cost | Store Sales | Unit Sales |
|-----------------|-------------|------------|-------------|------------|
| Food | 118,122 | 311,931.70 | 155,240.74 | 365,852 |
| Baked Goods | 4,838 | 12,786.96 | 6,755.24 | 14,979 |
| Baking Goods | 12,533 | 29,993.39 | 14,729.26 | 38,937 |
| Breakfast Foods | 2,138 | 5,416.26 | 2,756.01 | 6,640 |
| Canned Foods | 11,626 | 30,010.30 | 14,845.87 | 35,881 |
| Canned Products | 1,088 | 2,494.43 | 1,247.53 | 3,440 |
| Dairy | 7,876 | 22,766.78 | 11,434.42 | 24,282 |
| Deli | 7,378 | 19,352.60 | 9,727.49 | 22,992 |
| Snacks | 4,276 | 11,094.47 | 5,503.50 | 13,212 |
| Starchy Foods | 3,140 | 8,851.79 | 4,241.93 | 9,810 |

- Based on a OLAP model
- Write-back
- Meta-language to set new values
- Propagation up to the leafs and root
- Managment of versions and scenarios



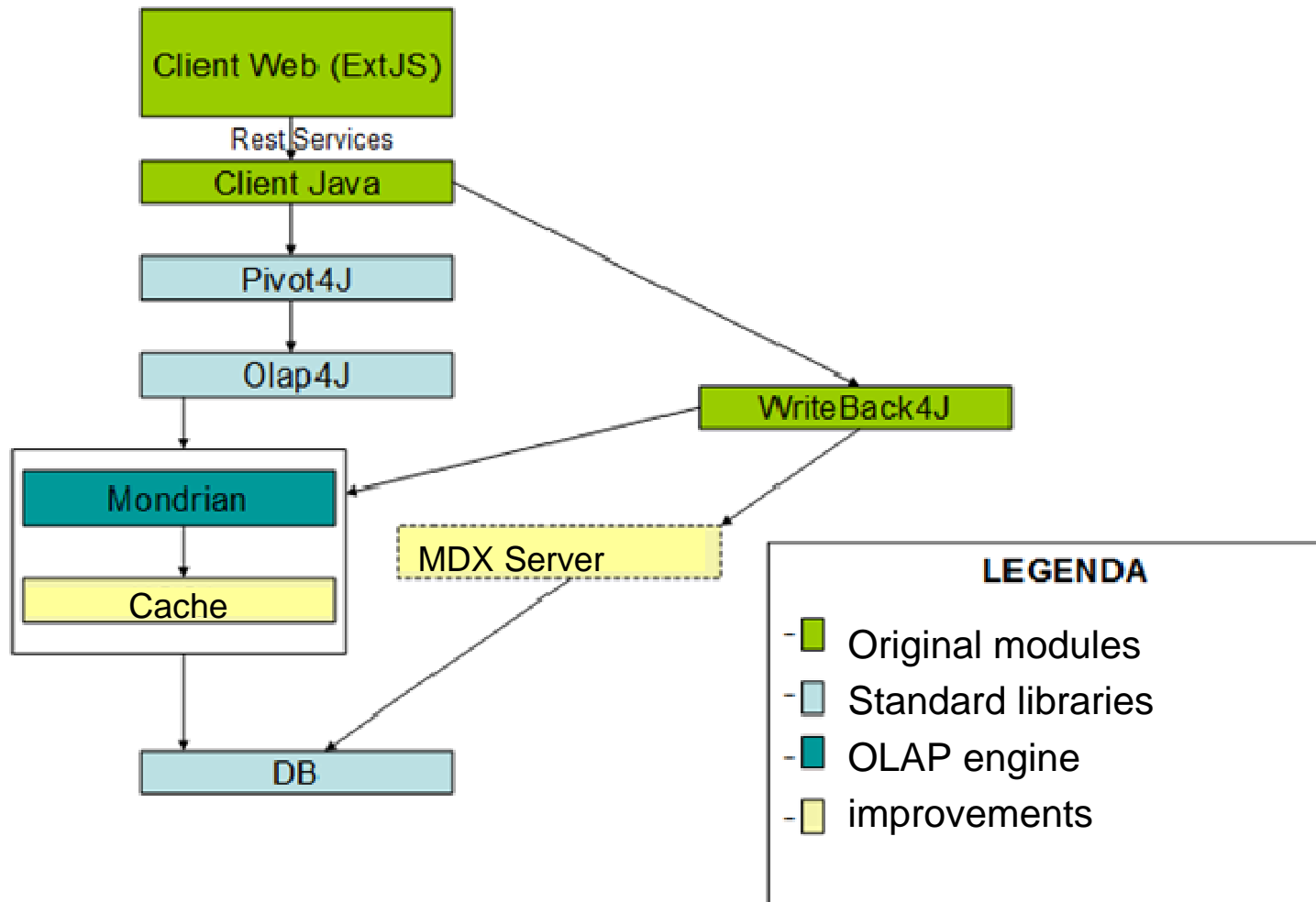


Architecture



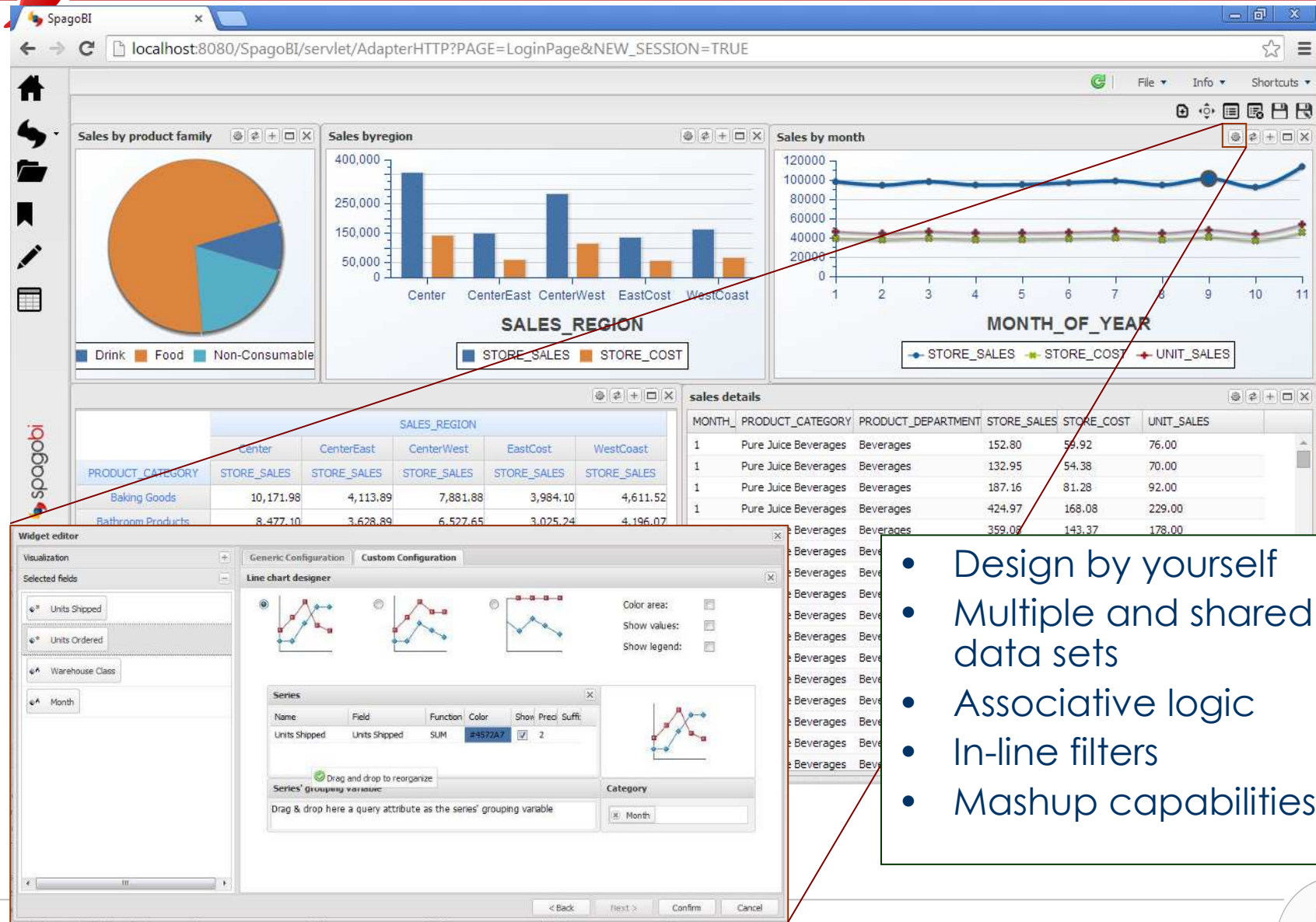


Architecture





Self-service in-memory cockpit

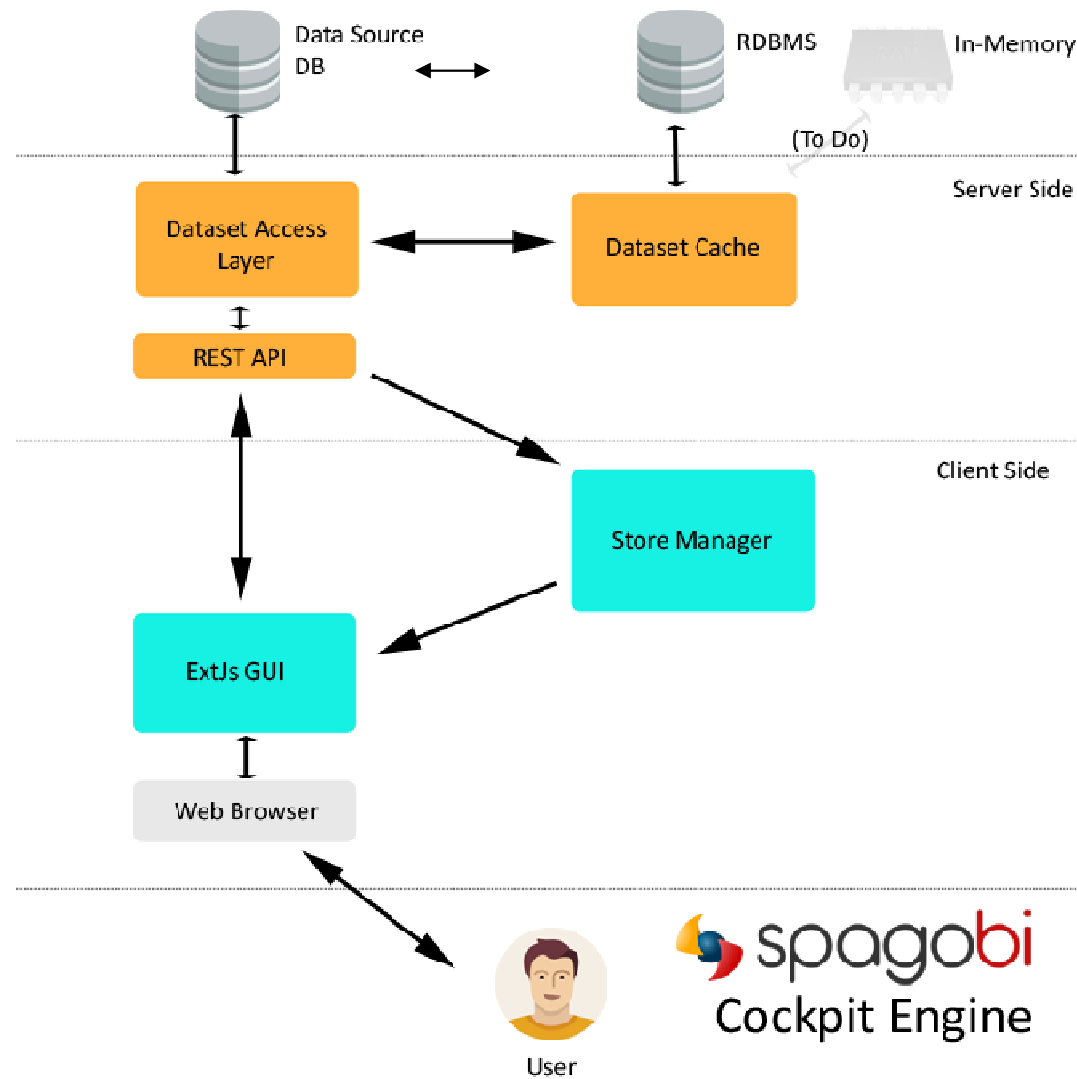


- Design by yourself
- Multiple and shared data sets
- Associative logic
- In-line filters
- Mashup capabilities





Dashboard – logical architecture





The only OS product for social listening

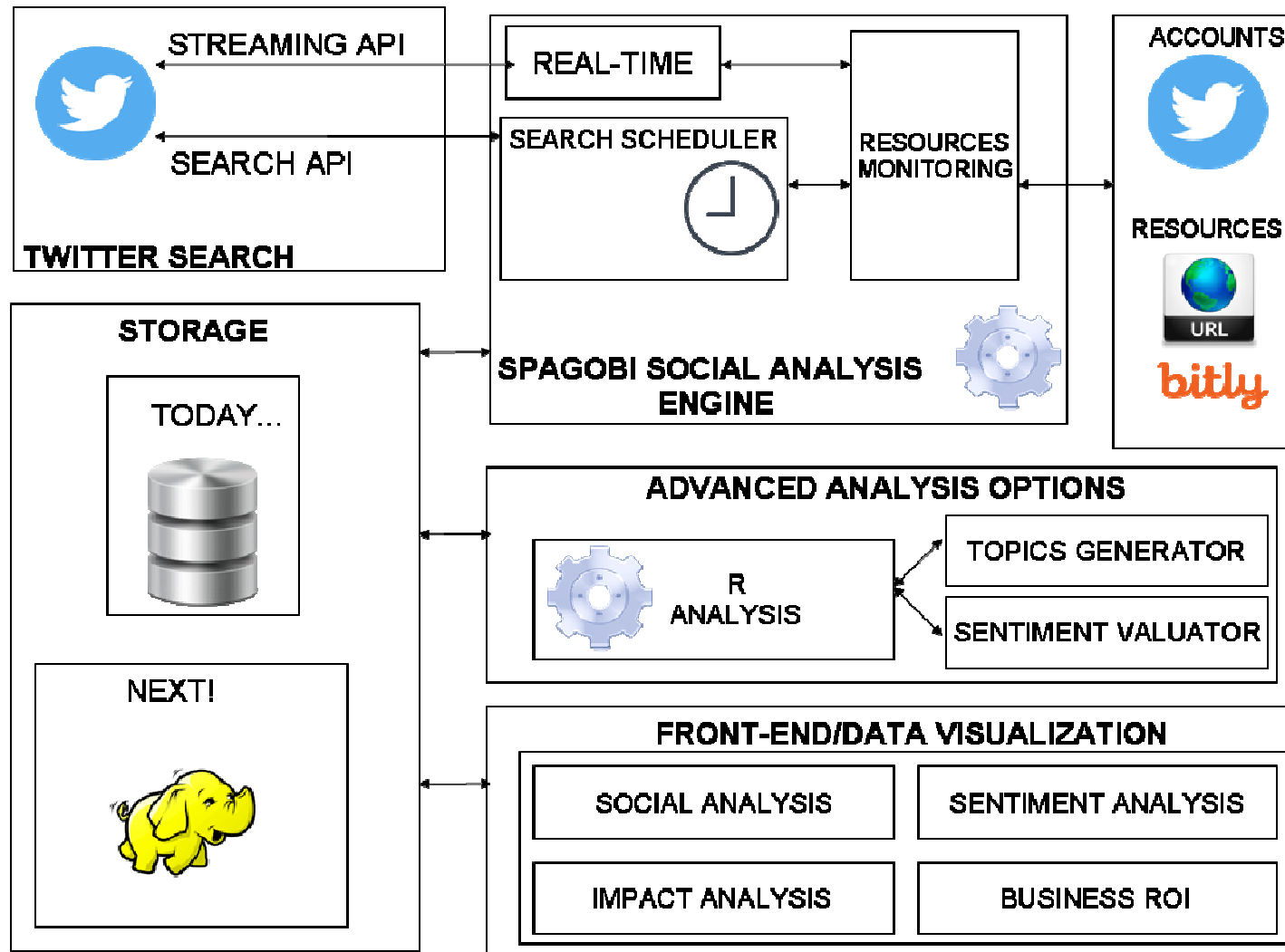


- Social listening
- RT monitoring & historical data
- Sentiment evaluation
- Evaluation of impacts on:
 - Visibility and interest
 - Business
- Multilanguage
- Text and data mining
- Gets insights
- Multiple perspectives





Social listening

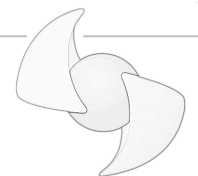






Targets

- Use R scripting language
- Execute it on R
- Interactive/automatic execution
- Display multiple outputs
- Execute multiple scripts
- Use multiple datasets
- Use R powerful charts
- Use SpagoBI AD
- Use SpagoBI datasets
- R per user workspace





Features

- **JRI** libraries (RForge): Java/R Interface, which allows to run R inside Java applications as a single thread.

• R environment installed on the same machine of spagoBI server (for beta release)

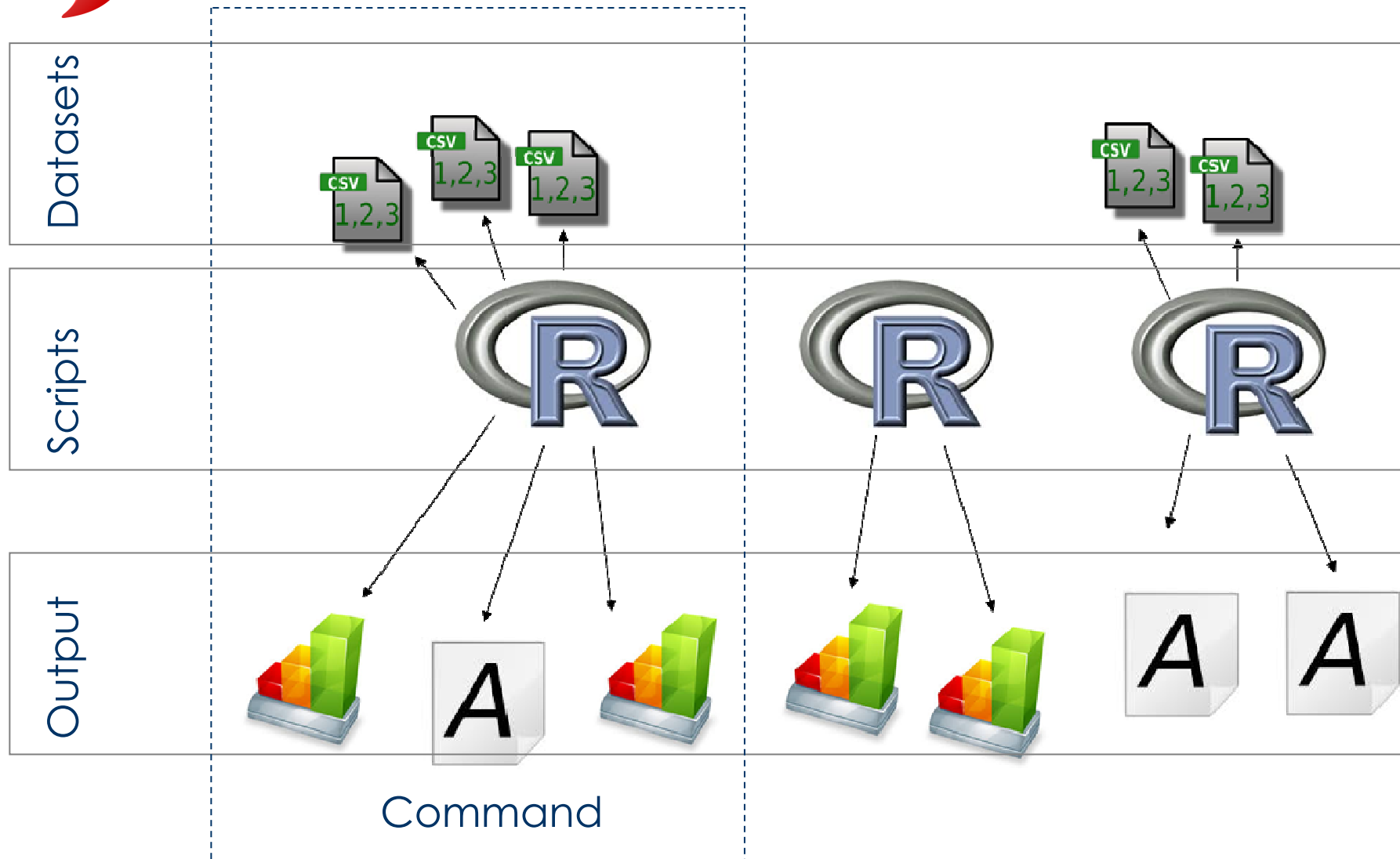
Rserve libraries (RForge): TCP/IP server which allows other programs to use facilities of R

• R remote installation (Rserve)





The SpagoBI data mining document





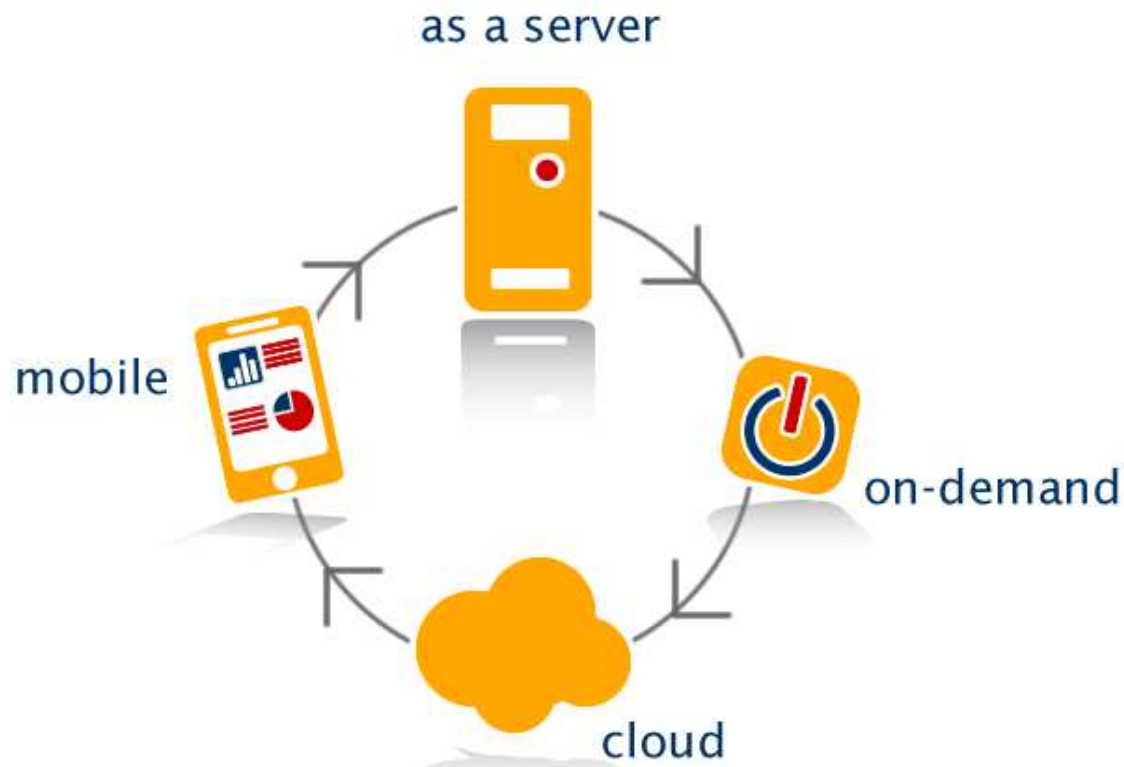
Template

```
<DATA_MINING>
  <PARAMETERS>
    <PARAMETER name="par1" alias="Param1"/>
  </PARAMETERS>
  <DATASETS>
    <DATASET name="fileDS" readType="table" type="file" label="label Data set 1">
      <![CDATA[ ...read_options... ]]>
    </DATASET>
    <DATASET name="spagobiDS" spagobiLabel="datasetQQQ" type="spagobi_ds" label="label Data set 2"/>
  </DATASETS>
  <SCRIPTS>
    <SCRIPT name="scriptAAA" mode="auto" datasets="fileDs,spagobiDS" label="label Script1">
      <![CDATA[....x,y... ]]>
    </SCRIPT>
    <SCRIPT name="scriptBBB" mode="manual" datasets="fileDs" label="label Script2">
      <![CDATA[...z... ]]>
    </SCRIPT>
  </SCRIPTS>
  <COMMANDS>
    <COMMAND name="command1" scriptName="scriptAAA" label="label Command 1" mode="auto" action="function1(x)">
      <OUTPUTS>
        <OUTPUT type="image" name="a" value="x" function="plot" mode="auto" label="label Output 1"/>
        <OUTPUT type="image" name="c" value="z,k" function="biplot" mode="manual" label="label Output 2"/>
        <OUTPUT type="text" name="b" value="y" mode="manual" label="label Output 4"/>
      </OUTPUTS>
    </COMMAND>
    <COMMAND name="command2" scriptName="scriptBBB" label="label Command 2" mode="manual">
      <OUTPUTS>
        <OUTPUT type="text" name="c" value="z" function="function2(y,z)" mode="manual" label="label Output 1"/>
      </OUTPUTS>
    </COMMAND>
  </COMMANDS>
</DATA_MINING>
```





Hot topics: Cloud / SaaS

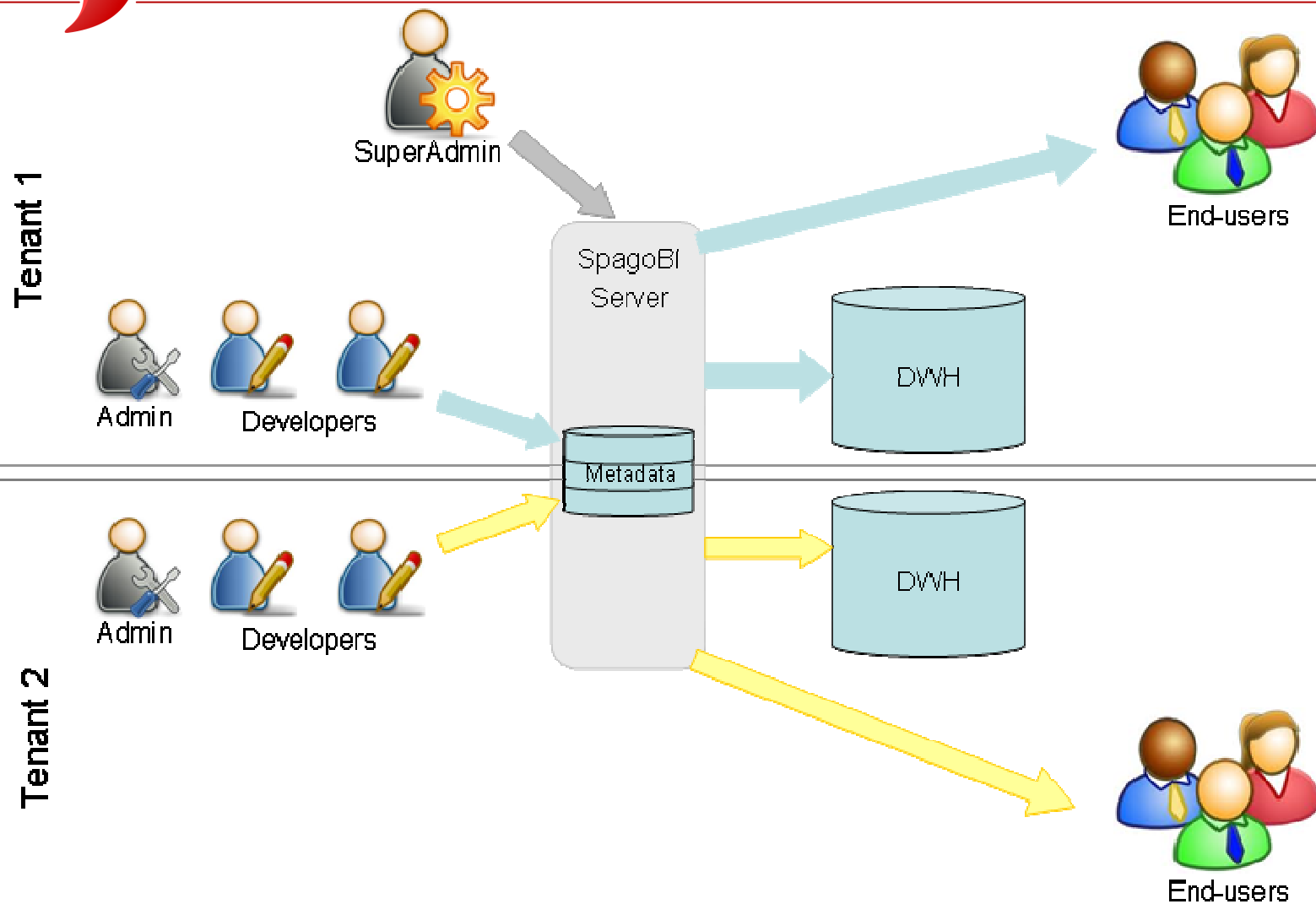


- In three words:
 - Client=browser
 - Multitenant
 - Self-service capabilities



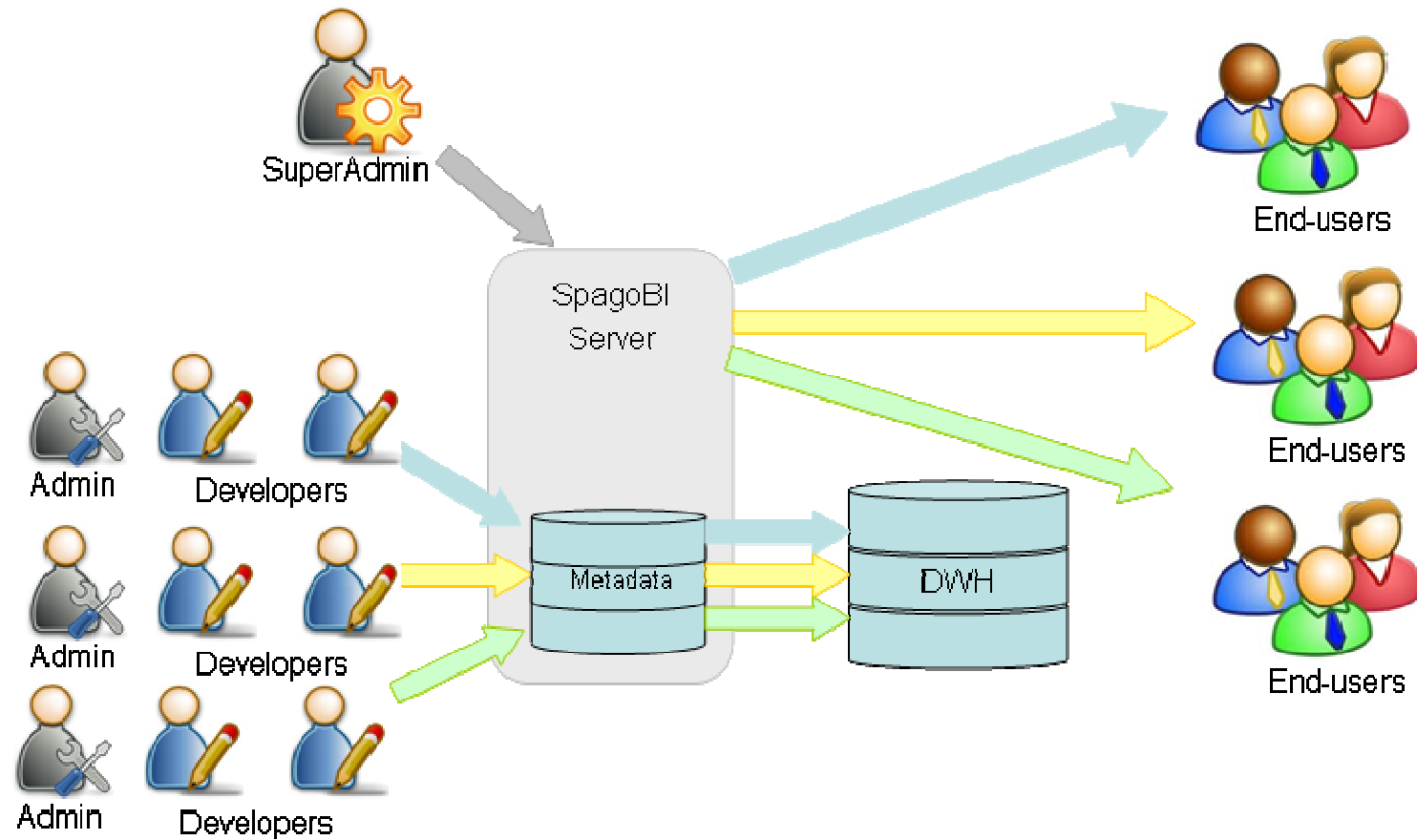


Multitenant : case #1





Multitenant : case #2



Roadmap and next steps



Roadmap

- Architecture review (REST compliant) → embedded BI
 - Analytical Engines
 - OLAP (export, calculated field, drill-through, print, search)
 - What-if (more algorithms, multi-user, I/O)
 - Cockpit (association discovery, new widgets, multiple sorting)
 - Social listening: more sources
 - New chart engine
 - Parametric data mining processes with R
 - Qbe: data federation
 - Improvements on mobile engine
 - Ad-hoc & self-service
 - My Analysis from smart filter
 - My Data : data set preview
 - SpagoBI Meta: metamodel lifecycle
 - New topics:
 - Open data (ckan integration)
 - Linked Open Data (GUI for semantic inquiry)
 - Big data
 - RT, HDFS in R/W mode
 - Advanced analysis and visualization
 - Offering: Cloud
- Association discovery on cockpit
 - Social listening
 - **Data federation**
 - **Metamodel lifecycle**
 - Open data
 - **Linked OD / semantic inquiring**
 - (big) Data visualization



Make your own solution



How to

- Enterprise BI project:
 - External user authentication and profile settings
 - Internal visibility rules based on received profile
 - SSO
 - L&F thematization
 - Multiple data sources
- Analytical application
 - Data model
 - ETL processes
 - Ready to use documents
 - Import/export
- Embedded BI
 - SDK usage
- Services based on BI
 - Multitenant
 - Cloud

