ergo
doc

Analysing activity traces in order to assess digital resources.

By Thibaud Hulin.

Teaching assisant at Laboratoire de Sémiotique, Linguistique, Didactique et Informatique (LASELDI), Université de Franche-Comté.

Credits: Laboratoire d'Informatique en Image et Systèmes d'Information (LIRIS), INSA / Université de Lyon I.

PLAN

- Issue of assess
- Definitions
- Presentation of the methodology
- Conclusion

To assess digital ressources

- Analysis of datas from system logs (*web log analysis*)
- Informations from **server-side**
- Advantage of **quantitative** data processing [Boudroux 08]
- Problem of **standardization** of analysis [Boukacem 05]

Analysing activity retrieval

- Studies about uses of electronic journals:
  - point of view of **readers**
  - point of view of **writers**
- Problems of standardization of interfaces [Lomprez 07]
- Studies on modeling of information retrieval [Tricot 98]

Methodology for analysis of activity traces

- Combining the **flexibility** of the statistical analysis
- With the **precision** of activity analysis
- Get information on the **client side**
- Role of the **user as actor** analysis

Issues

- **Decentralization** of resources assessment (top / down)
- Enhanced **autonomy** in libraries from data provided by publishers
- Better **knowledge** of the actual activity of retrieval
- Reduction or **cost** optimization

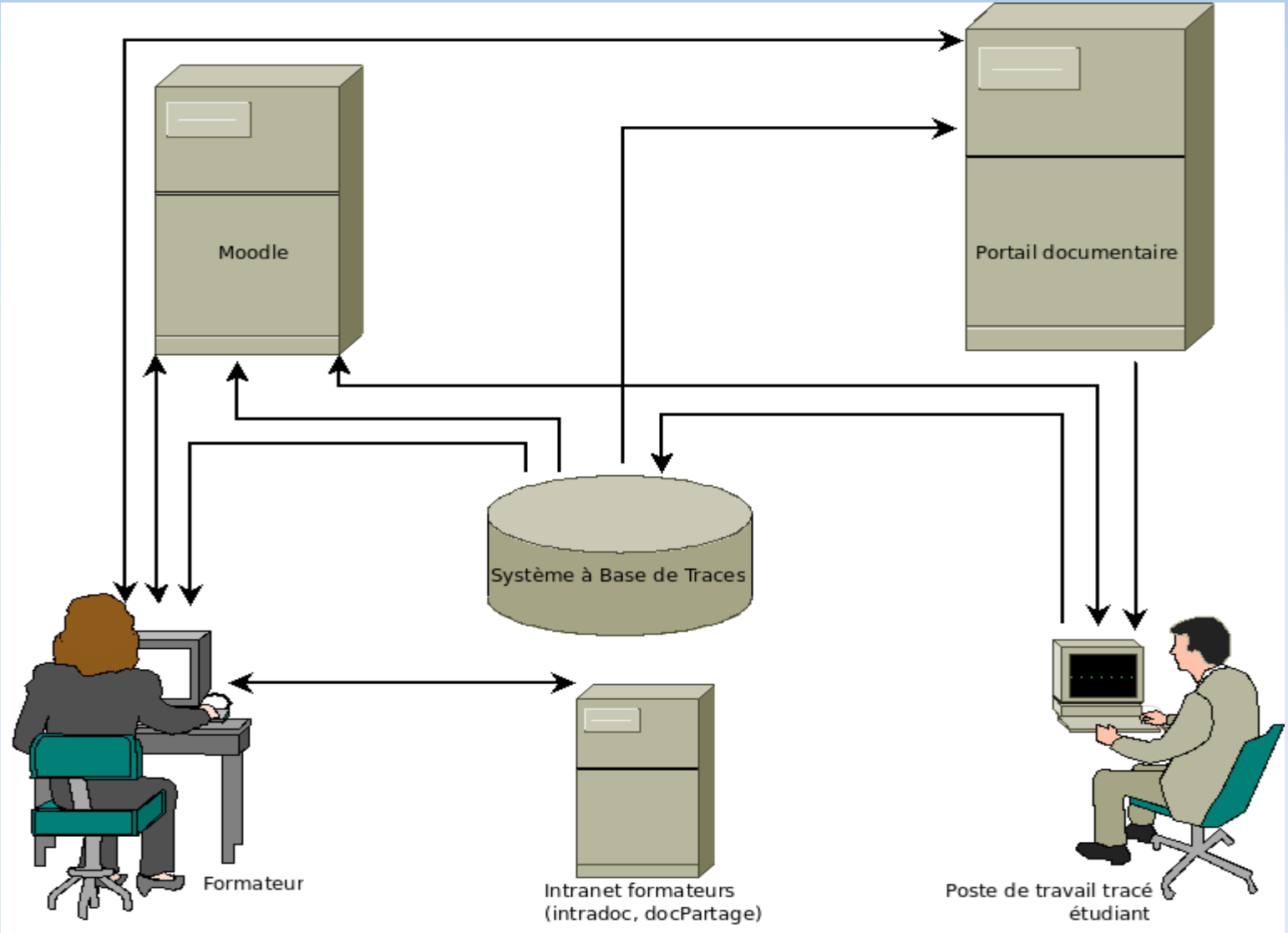
Definitions

- Management system based on traces
- Observed
- Collection
- Trace interaction
- Documentary sequences
- Transformation
- Visualization

Trace database management system TBMS

- **Computer system**, with or without memory,
- managing a database of traces modeled through a **system observed**
- and providing processing **services**, application, sharing ...

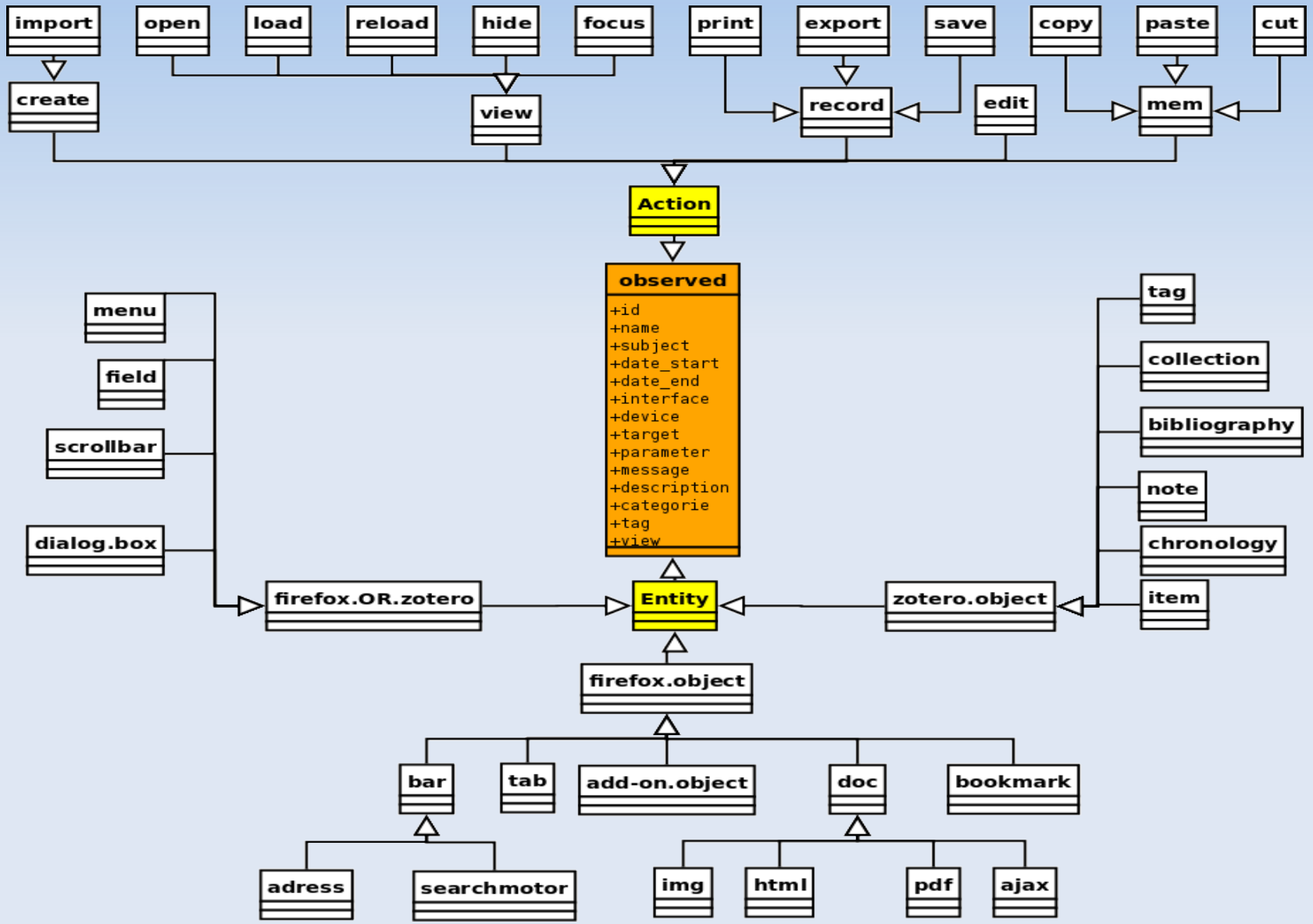
Draft of a TBMS in library



Observed

- This is something **temporally located** in a trace modeled
- Has a defined **type**
- Has a **subject**
- May be **related** to other observed

Example de modeling observed



Collection

- Collect source: probes, system logs, audio and video record...
- Collect: creates a trace material in a database management traces

Interaction modeled traces

- **Modeled trace**: composition observed, organized according to a model trace (trace model): types, relationships, transformations ...
- **First trace**: trace directly after collection, unprocessed
- **Computed trace**: trace associated with automatic processing, related to views

Documentarizing sequences trace

- Creating a document **exchanged** outside the BCMS,
- is derived from the **native** or empirical observation (video, audio ...),
- is from a documentary sequences **of observed** (text, tables, chronology ...).

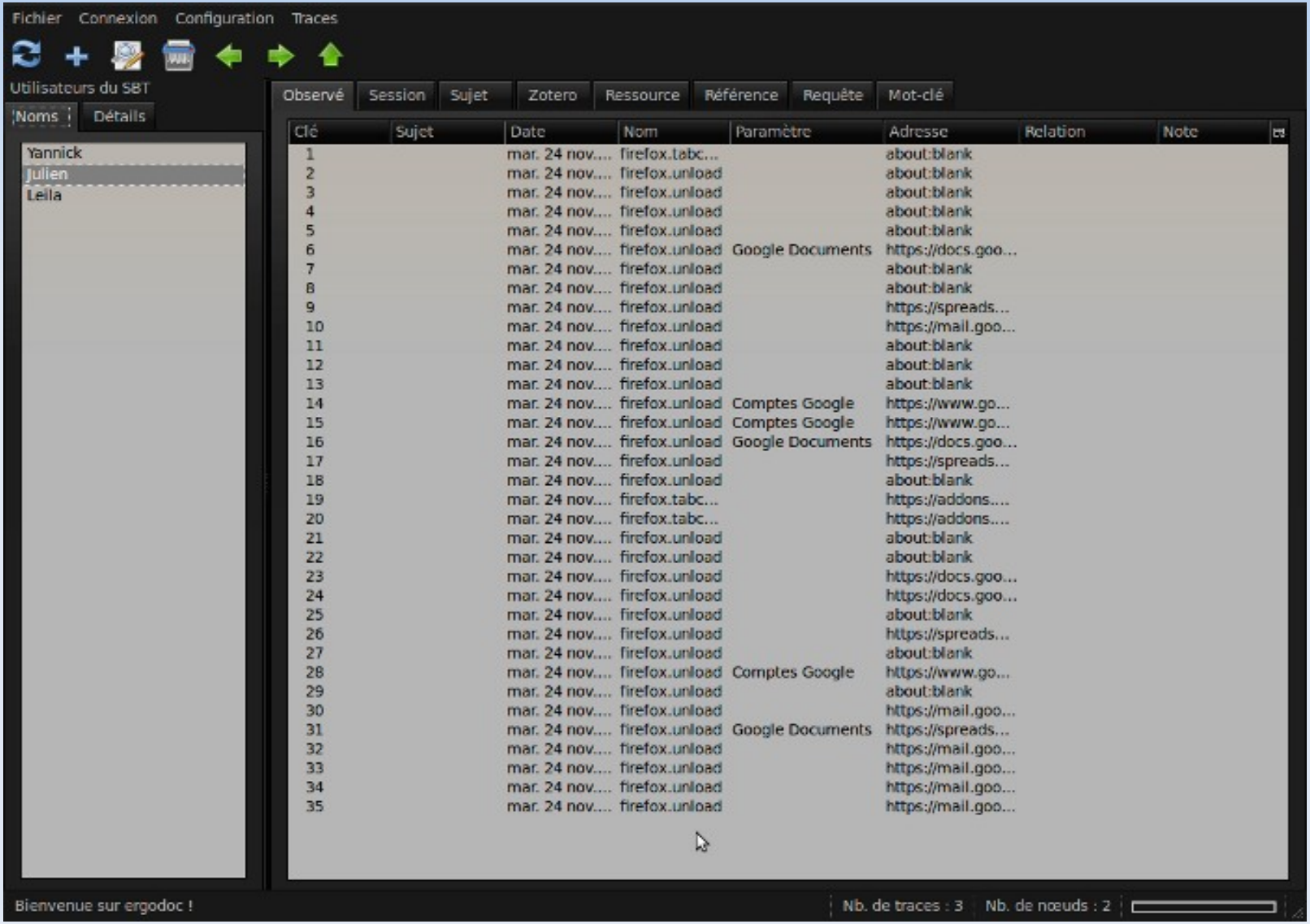
Traces transformation

- **Operation** on one or more traces modeled first or compound
- in order to **produce** a new trace modeled;
- if it is described by a model transformation, it is called **automatic**
- if it is the result of a specific intervention, it is called **manual**.

Interactive trace visualization

- **Consultation** by a user with a visual representation of a trace
- according to predefined **rules**.
- **Views** can be directly associated with different traces calculated.

Example of presentation (ergodoc)



Activity model retrieval

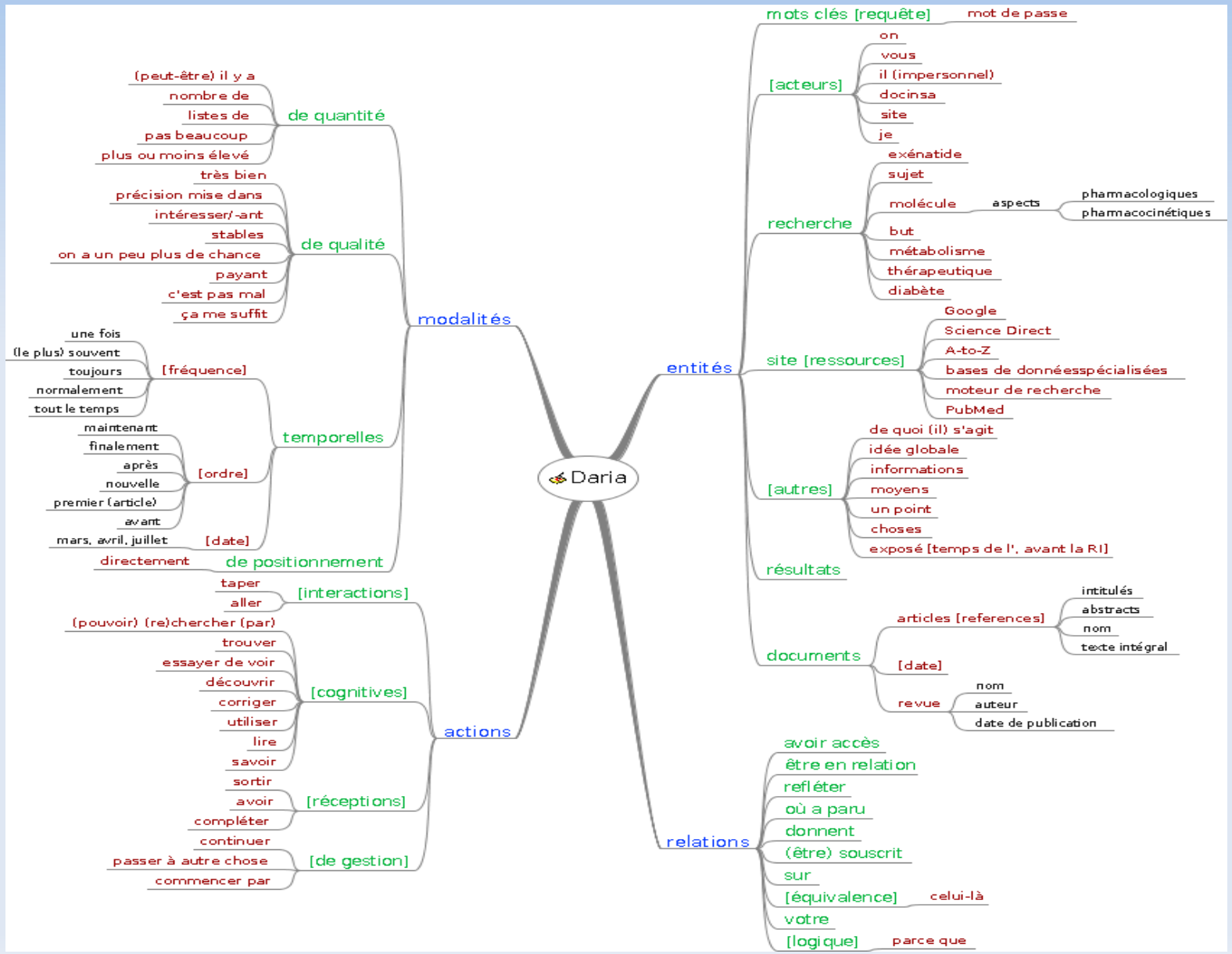
The discovery of a system of reflexivity [Jermann 01] involves several steps:

- Model collection
- Construction of a model trace
- Defining a profile evaluative
- Representative traces
- Creating a calculation engine traces
- Classification of resources

1. Data collection

- Study the **ergonomics** of the interfaces used
- Studies of the **semantics** of the retrieval activity
- Identify **user actions** relevant
- Choose **site implementations** in the code of sytems observed

Example analysis of the semantics of the user



2. Construction of a modeled trace

- Defining a model seen as an object with **attributes**: id, description, dates ...
- Building a **hierarchy** of actions that users can describe observed, giving a record first
- Define **transformations** from the first model track to make it usable.

Example of transformations

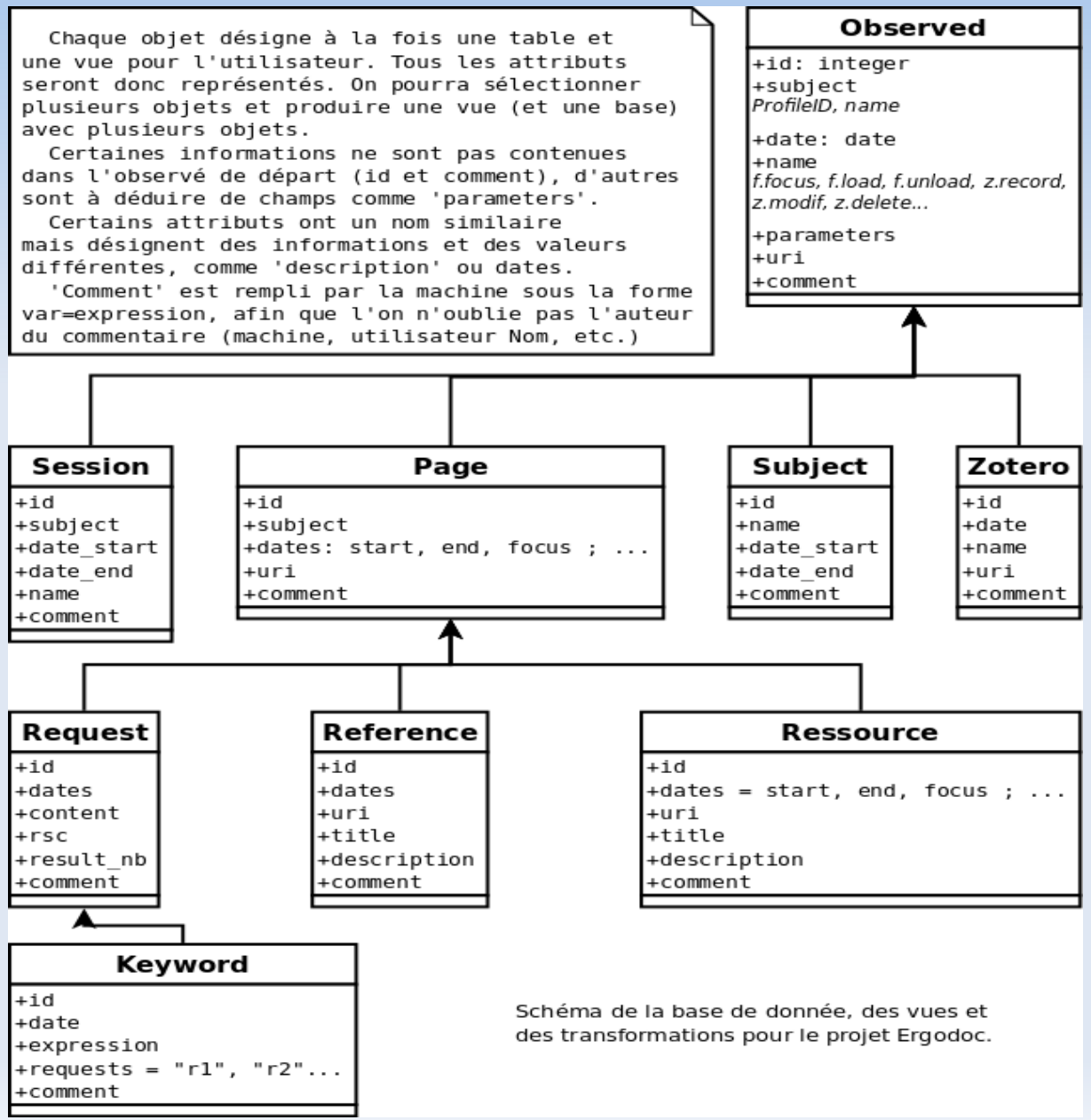
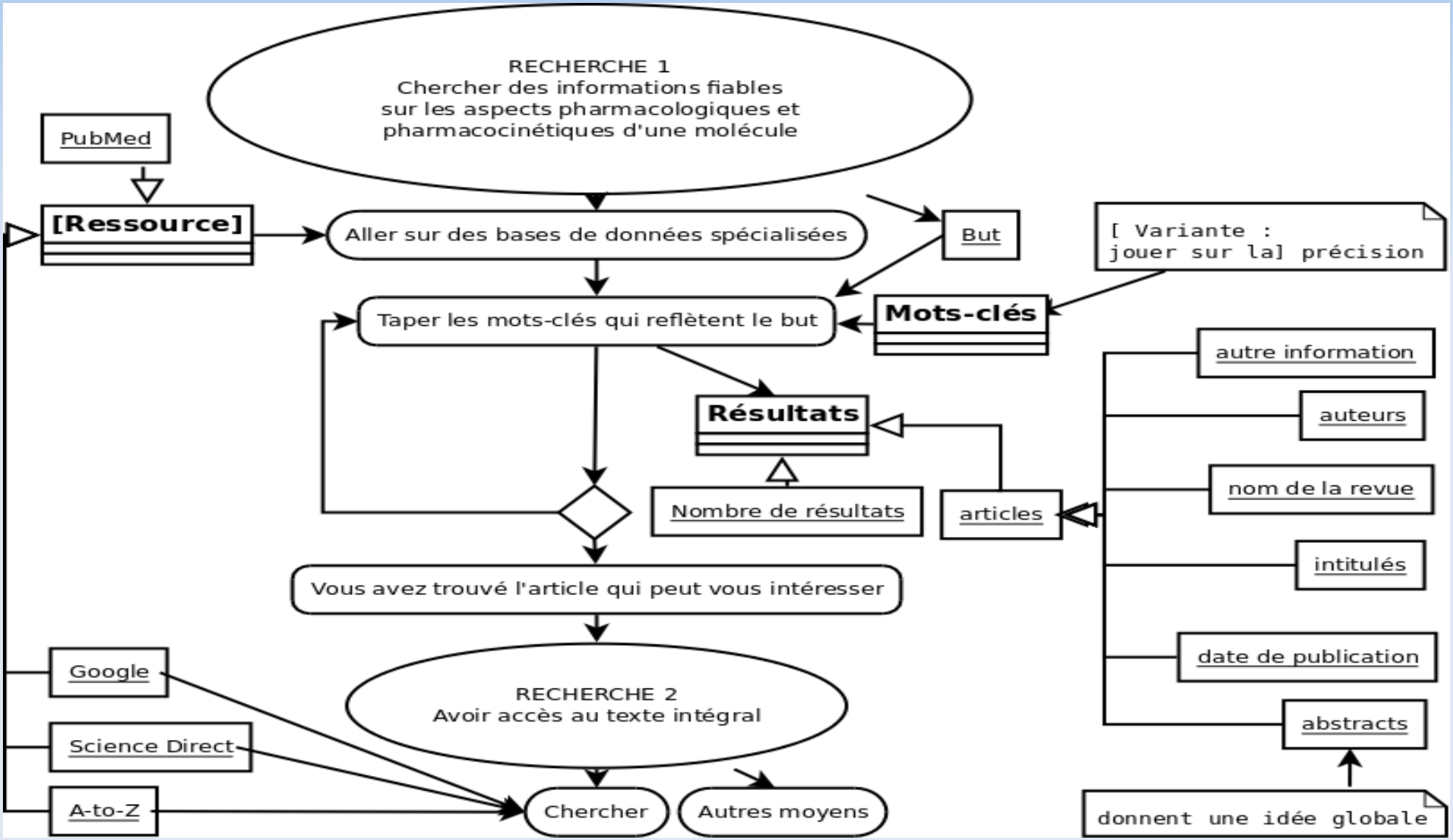


Schéma de la base de donnée, des vues et des transformations pour le projet Ergodoc.

3. Defining a profile evaluative

- To evaluate a resource, the profile concept refers to a type of use and user type.
- These types represent a model of transformation.
- This avoids the simplicity of the model action = one point for the resource.
- Example: saving a document in Zotero + display for at least 2 'document.

Example of search model



4. Visualization of traces

- Choice of **modes of representation**:
  - tabular form;
  - under formpe chronology.
- Opportunity to **comment** on the tracks (attribute "notes").
- The analysis of these comments may lead to further evaluation using a **semantic formalism** appropriate, eg. "Eval = too Wait" (subjective?).

5. Creating a calculation engine traces for evaluation

- We define a standard activity model or **ideal type** (eg presence of steps with expected activity...)
- We **compare** this model with that described by the collector for real activity.
- It **regulates** and adjusts on the business model based on user profiles, observed constants, etc..

6. Classification of resources

- It is based on the traces calculated by comparing the traces on the actual activity with **expected patterns**, as **automatons** .
- It is to **reduce** these traces evaluation indicators, which means to interpret the activity with reservations, which is a work of **humans**.

Conclusion

- The proposed methodology tries to **combine** the advantages of qualitative analysis (ergonomic) and quantitative analysis (data mining).
- It doesn't avoid the difficulties of comparing actual activity with the activity for early or statistical interpretation.
- It offers a rich from a semantic activity.
- Other advantages: ownership by the user, sharing, collaboration, etc..

Projet ergodoc
https://addons.mozilla.org/fr/firefox/addon/51326/

Thank you for your attention!