

# An Effective Visual Data Mining Environment

---

Stephen Kimani

Università di Roma "*La Sapienza*"

Supervisor: Tiziana Catarci

# An Effective Visual Data Mining Environment

---

## Motivation

### Major problems with existing DM systems

They are based on non-extensible frameworks

They provide a non-uniform mining environment - the user is presented with totally different interface(s) across implementations of different DM techniques

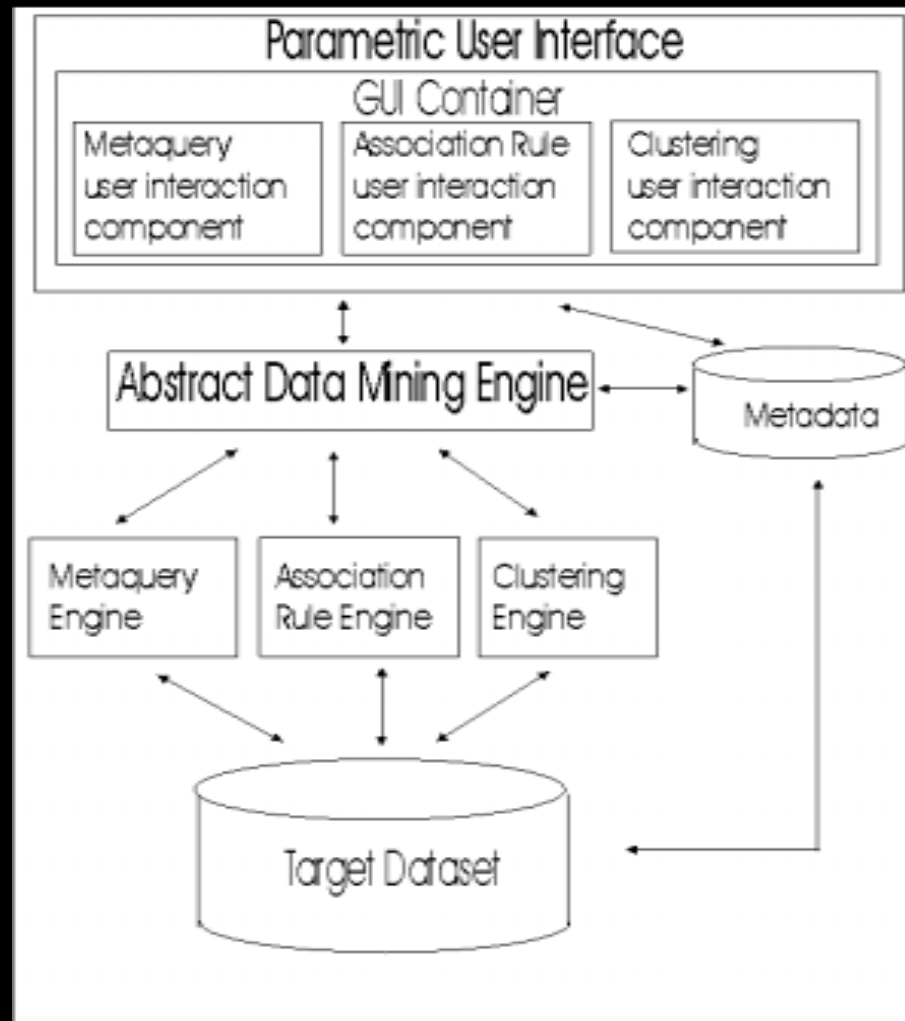
### Major needs

An **overall framework** that can support the entire Knowledge Discovery (KD) process (accommodate and integrate all KD phases seamlessly)

Placing the **user at the center** of the entire KD process/in the framework. In fact the corresponding system should provide a **consistent, uniform and flexible visual interaction environment** that supports the user throughout the entire discovery process

# An Effective Visual Data Mining Environment

## System Architecture



## Primary layers

- User layer
- Engine layer
- Data layer

## Features

- Component incorporate/modify
- Input of task  $y$  = Output of task  $x$
- User component can directly access virtually each of the other components; enable the user to: *process* data and knowledge, *drive* and *guide* the entire KD process

The proposed system supports, but is not limited to: metaqueries, association rules, and clustering

# An Effective Visual Data Mining Environment

## Visual Environment

A **consistent, uniform, flexible and intuitive GUI**, with support throughout the whole DM process. The principal focus is to support the user in:

**Planning:** E.g., advertising relevant prior knowledge

**Visual construction of the task relevant dataset:** The user directly interacts with data. For this task, there are two intuitive interaction spaces

**Visual construction of the mining query:** The user directly interacts with data and other parameters (e.g. threshold values) in making queries e.g., in the Metaquery Environment, the user can suggest patterns by linking attributes, while the Association Rule Environment offers 'visual baskets'

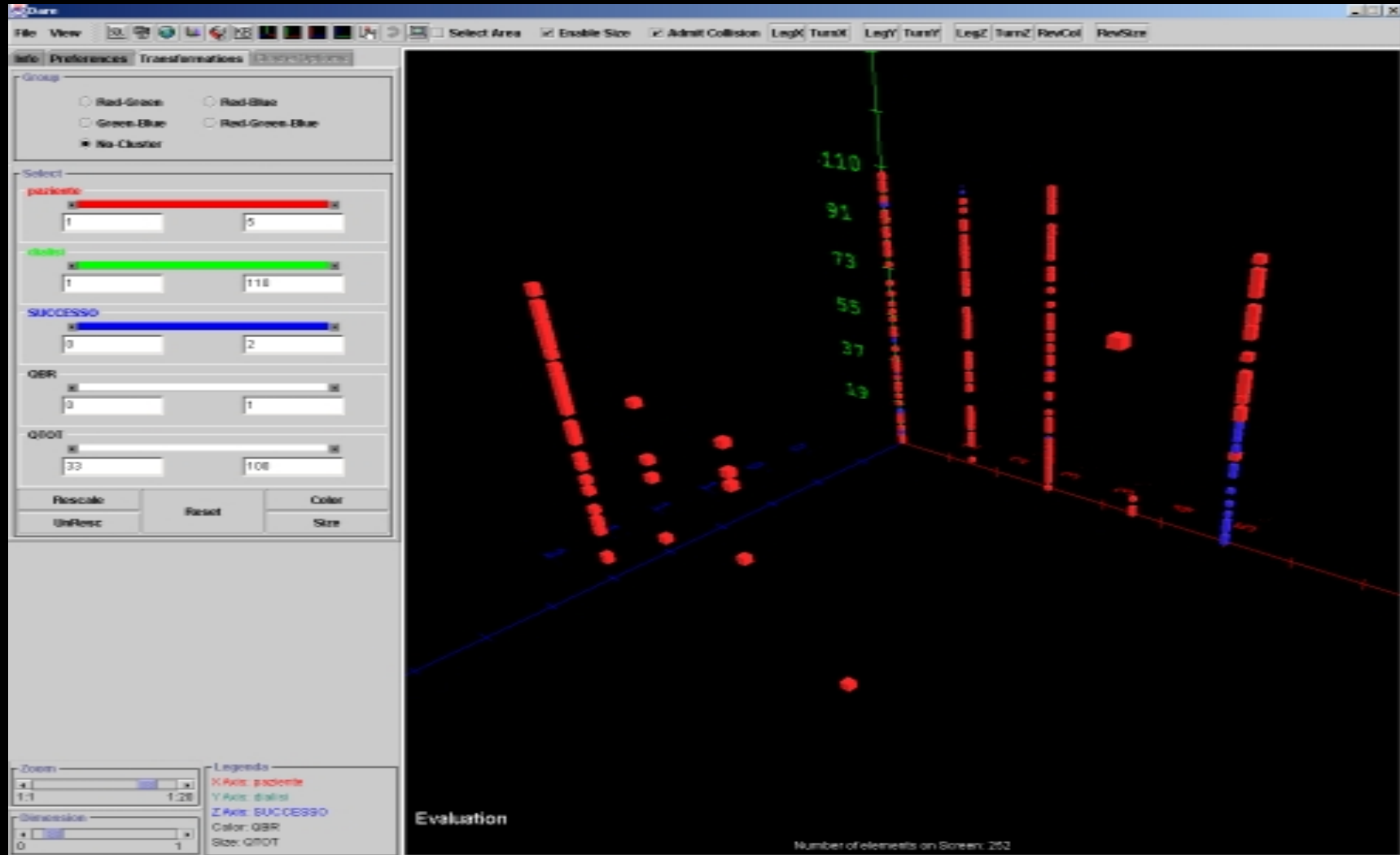
**Visual output presentation and interaction:** Exploiting relevant effective visualizations and where necessary, we have designed novel visualizations

**Handling the non-static nature of user's quest:** E.g., enabling user to adjust



# An Effective Visual Data Mining Environment

## Exploiting DARE Visualization System



# An Effective Visual Data Mining Environment

---

## Usability

**Usability heuristics:** This has been done, but regular reference to the same will go on till the project is over

**Mock-up tests:** These were done with DM experts. The experts gave an encouraging feedback and even suggestions on how to improve the interface

**Formal usability tests:** In the pipeline

At present, there is a partial prototype, complete implementation is underway

## Formalization

Defining a formal mapping /language between **DM Engine** and **Visual Interface**, an XML-based mapping between the abstract DM Engine components and the corresponding visual operators - the mapping definition will be given in terms of its syntax and semantics