

**MINI:
Mining Informative
Non-redundant Itemsets**

Arianna Gallo

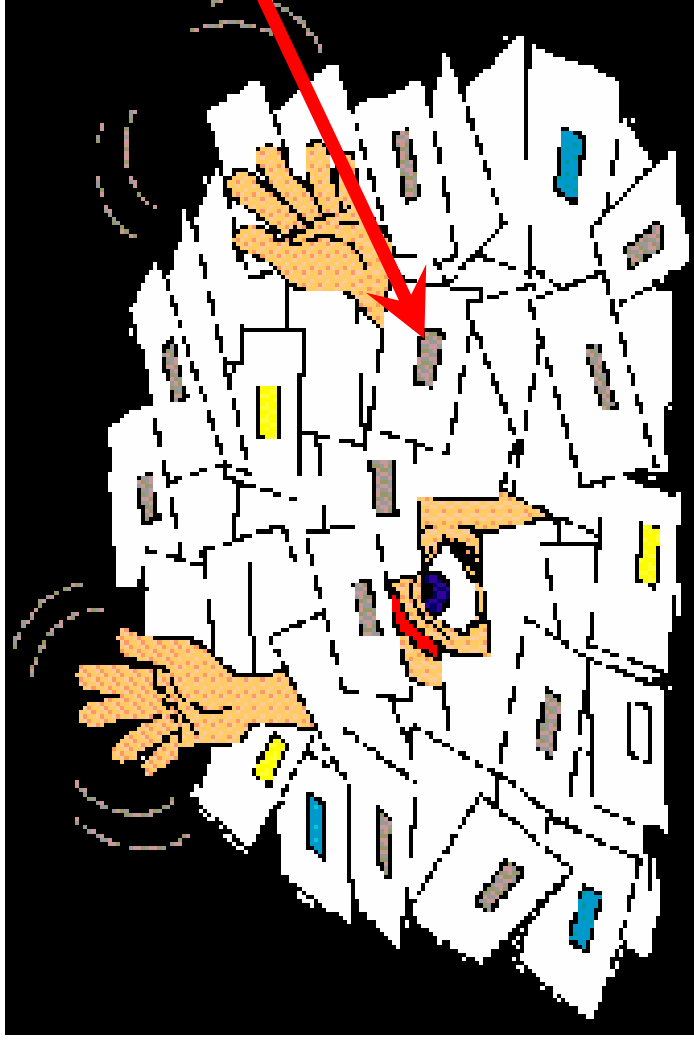
Tijl De Bie

Nello Cristianini

University of Bristol

We discuss *frequent itemset mining* ...

An impressive tool to reveal information from a database



Itemsets

- Many itemsets are 'obvious'
(if containing frequent items)
- Many are redundant
(largely overlapping itemsets)

"We're drowning in information and starving for knowledge."

Rutherford D. Rogers

Solutions?

- Maximal itemsets
- Closed itemsets
- ...

But...

- Insufficient reduction
- Do not reduce redundancy

Cause:

- Focus on *restoration error*

We will focus on a *statistical measure of informativeness* and *originality*

Result (PKDD abstracts):

1. Logic programming inductive
2. Vector machine classification support learning
3. Sequence pattern patterns sequences mining
4. Association rules mining paper
5. Art state
6. Trees decision tree
7. Itemset large itemsets frequent mining
8. Documents document