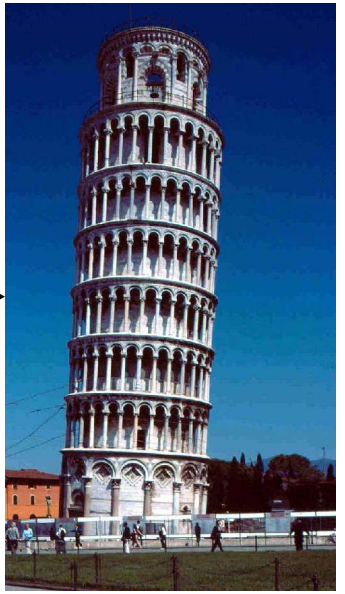
From Nancy, France *to* Pisa, Italia



Ontology-guided Data Preparation for Discovering Genotype-Phenotype Relationships

Adrien Coulet, Malika Smaïl-Tabbone, Pascale Benlian, Amedeo Napoli and Marie-Dominique Devignes

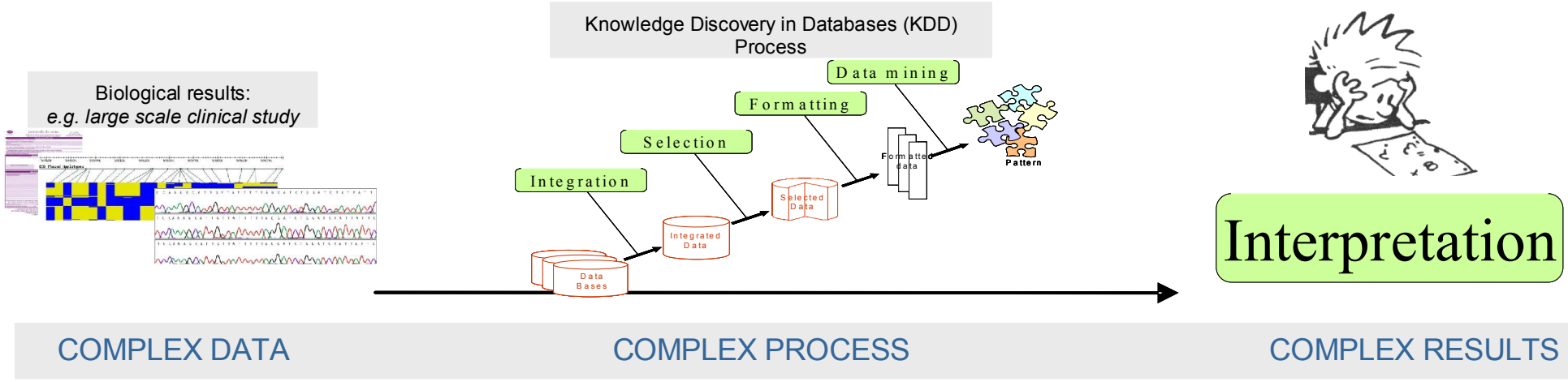
Laboratoire Lorrain de Recherche en Informatique et ses Applications (CNRS, INRIA, University of Nancy), Nancy, France

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

INRIA

Nancy-Université

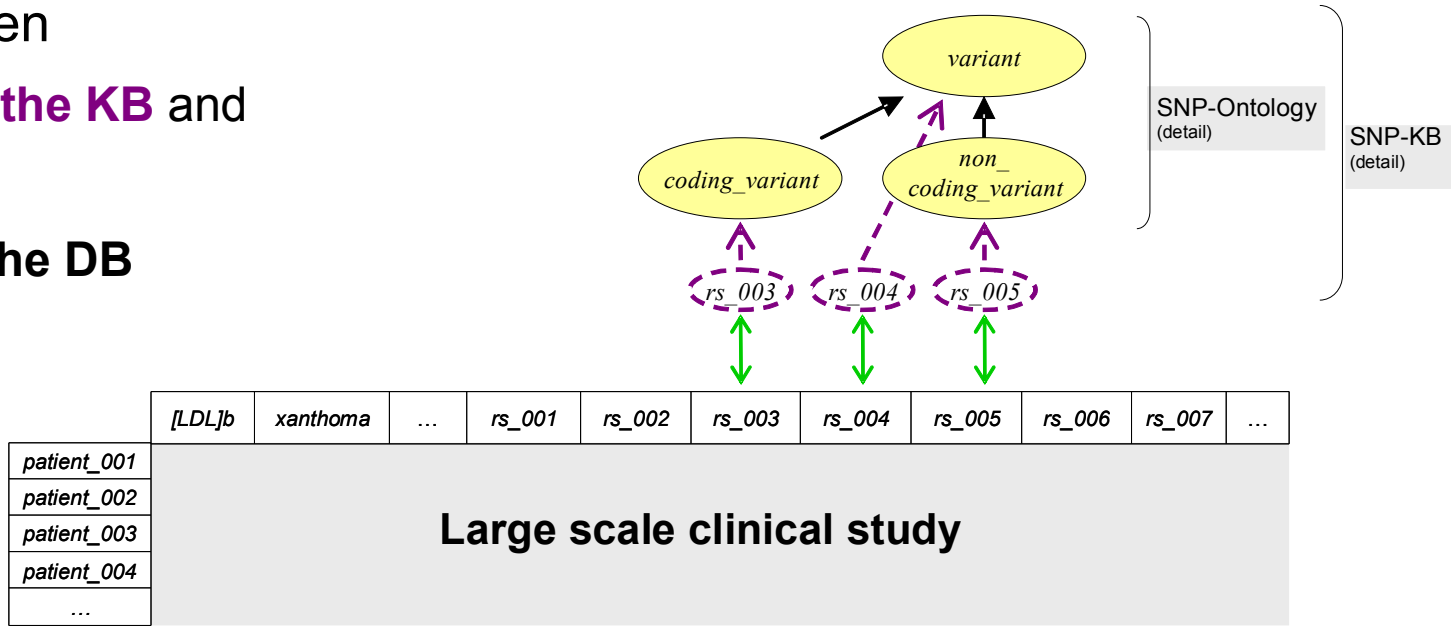
The Problem: Limits to KDD in life sciences



- Results of KDD in biology are complex

Proposition: Use ontologies for guiding the KDD

- 1) Build bridges between data and knowledge
  - **Mapping** between
    - **assertions of the KB** and
    - **attributes of the DB**
  - Example:



- 2) Use knowledge in order to reduce the size of the data set
  - Thanks to *subsumptions*, *object properties*, *class definitions*, etc.
  - In order to simplify the interpretation step of KDD process

For more details …

- …see you around the poster
- Poster n°7
- Contact: adrien.coulet@loria.fr