



Towards an Ontology for Tertiary Bioinformatics Research Process

Sara Pidò, Pietro Crovari, and Franca Garzotto

November, 4th 2020

POLITECNICO
MILANO 1863

sara.pido@polimi.it

Motivations

Bioinformatics analysis:

Primary analysis: it identifies and evaluates raw data, focusing on generating readable sequencing reads and ranking the basis's consistency.

Secondary analysis: it consists of aligning the reads against the human genome and variant calling. It transforms raw data in processed data.

Tertiary analysis: it studies the processed data for sense making.

Motivations

Major issue of bioinformatics tools:

- Usability ⁽¹⁾
 - Not fully exploitation of these tools
 - Lose time to understand their usage

How can we tackle this problem?

(1) Bolchini, D., Finkelstein, A., Perrone, V., Nagl, S.: Better bioinformatics through usability analysis. *Bioinformatics* 25(3), 406 - 412 (2009)

Our Research

Goals:

- Think from a researcher's perspective: definition of the **user requirements**
- Build an **ontology** that represents the complete tertiary analysis process

Empirical Study

Population : 8 bioinformaticians

Protocol : Semi-structured individual interviews composed of 3 phases:

- describe the steps
- classify the process elements
- create a hierarchical task tree

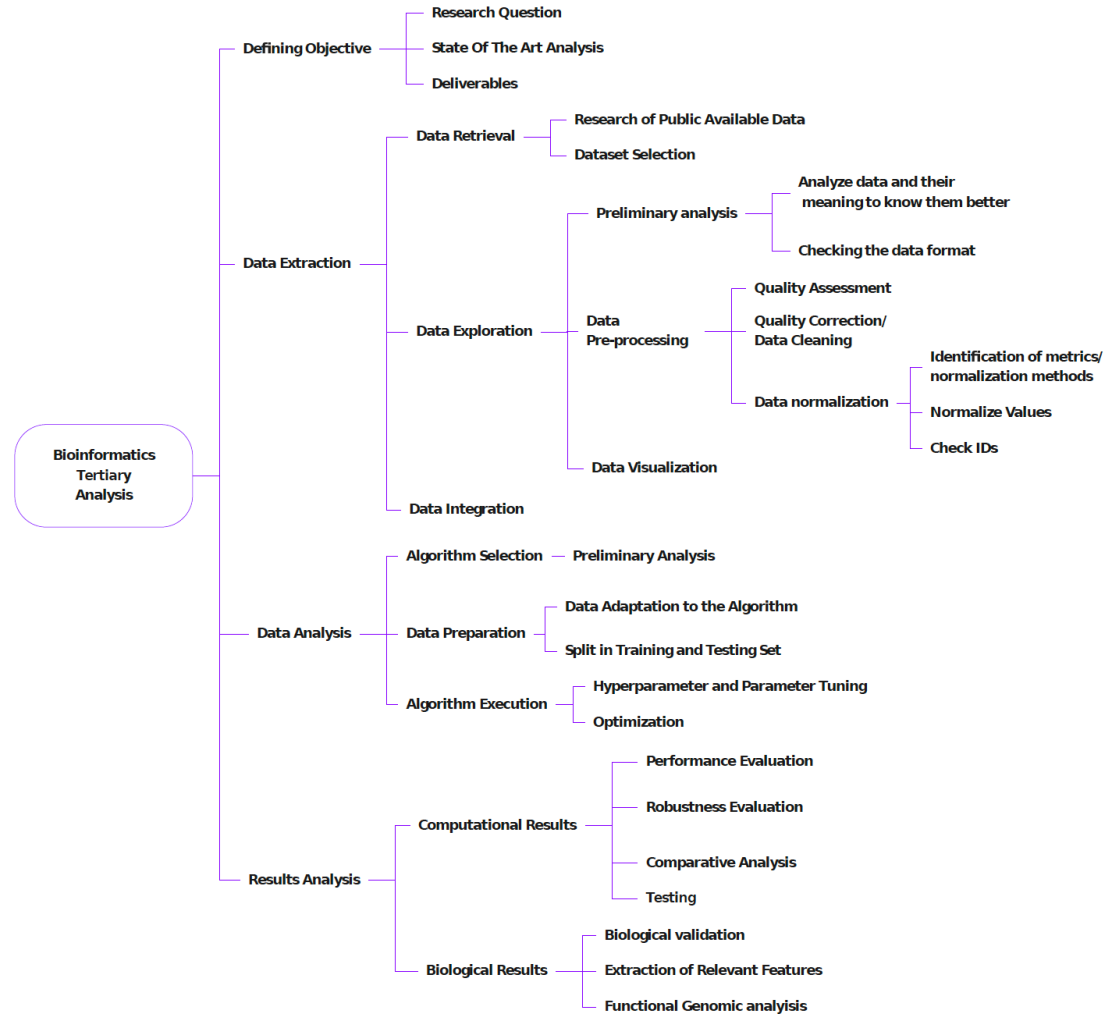
Result: From the interviews emerged:

- Several abstraction levels
- Process skeleton shared among all the interviews
- Each subject focused on different aspects

Ontology

- From the 8 hierarchical task trees,
 - we retrieved 1 hierarchical task tree:
 - it represents the **ontology** of the bioinformatics tertiary research process

Ontology



Conclusions and Future Works

Main contribution:

- Definition of an **ontology** of the bioinformatics tertiary research process to help the development of usable bioinformatics tools.

Future works:

- Going in deep in the last phases of the process to enlarge the ontology
- Validate our ontology through a bottom-up analysis
- Formulate the ontology using standard notations.



Towards an Ontology for Tertiary Bioinformatics Research Process

POLITECNICO
MILANO 1863

Sara Pidò, Pietro Crovari, and Franca Garzotto

Thank you for the attention!

sara.pido@polimi.it