

Postdoctoral research topic

- Title of the proposed topic: Argument-based quality assessment of clinical trial articles
 - Research axis of the 3iA:
 - Axes 2: AI FOR INTEGRATIVE COMPUTATIONAL MEDICINE
 - Axes 1: CORE ELEMENTS OF AI
 - **Supervisor (name, affiliation, email): Serena Villata (Université Côte d'Azur, CNRS, Inria, I3S), email: serena.villata@univ-cotedazur.fr**
 - Potential co-supervisor (name, affiliation): Elena Cabrio (Université Côte d'Azur, CNRS, Inria, I3S), email: elena.cabrio@univ-cotedazur.fr
 - The laboratory and/or research group: WIMMICS (<http://wimmics.inria.fr/>) is a research team of Université Côte d'Azur (UCA), Inria, CNRS. The research fields of the team are graph-oriented knowledge representation, reasoning and operationalization to model and support actors, actions and interactions in web-based epistemic communities.
-

Apply by sending an email directly to the supervisor.

The application will include:

- Letter of recommendation of the supervisors indicated above
 - Curriculum vitæ including the list of the scientific publications
 - Motivation letter
 - Letter of recommendation of the thesis supervisor
-

- Description of the topic:

Argument(at)ion mining [3], the new and rapidly growing area of Natural Language Processing (NLP) and computational models of argument, aims at the automatic recognition of argument structures in large resources of natural language texts. In the clinical domain, argument mining has been proved to be beneficial in providing us with methods to automatically detect in text the argumentative structures that are at the basis of Evidence-Based Medicine (EBM), which is the “conscientious, explicit, and judicious use of current best evidence” to guide clinical decision-making with scientific information from systematic reviews [2].

The **goal** of this post-doc position is to define and evaluate new argument mining methods to automatically assess the quality of the arguments put forward in clinical trial articles. Relying on one of the main features of artificial argumentation, i.e., the ability to capture

inconsistencies among arguments, we aim at mining argument components both from the full text of the clinical trial articles and from the abstract. As it has been noticed in the literature [1], sometimes abstracts contain a more positive reporting of the main findings of the article than what stated in the full text. Employing argumentation mining methods to automatically identify these instances of misrepresentation and distortion of the results in clinical trial articles is a challenging and crucial research line for healthcare intelligent applications.

The **main objectives** of the post-doc program therefore are:

1. Identify argumentative components in the full text of clinical trial articles.
2. Assess the quality of argumentative components in the full text and the abstract of the articles. This objective will include the definition of quality guidelines for clinical trial argumentative structures. An annotated linguistic resource will be built.
3. Check for the consistency of the claims and premises in the abstract and the full text of the clinical trial articles. This step is in charge of identifying automatically those instances of misrepresentation and distortion of the results in analysed clinical trial articles.

References

[1] I. Boutron, P. Ravaud. Misrepresentation and distortion of research in biomedical literature, *Proceedings of the National Academy of Sciences* 115 (11) (2018) 2613–2619.

[2] Tobias Mayer, Elena Cabrio, Serena Villata. Transformer-Based Argument Mining for Healthcare Applications. *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI 2020)*, pages 2108-2115, 2020.

[3] Elena Cabrio, Serena Villata. Five Years of Argument Mining: a Data-driven Analysis. *Proceedings of Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI 2018)*, pages 5427-5433.

Keywords:

Natural Language Processing, Argument Mining, Clinical Text, Quality Assessment, Consistency Checking

Skills and profile:

- Master degree in Data Science, Computer Science or Computational Linguistics is required.
- Programming skills are required.
- Knowledge of Natural Language Processing and Machine Learning is preferred.
- Fluent English required, both oral and written. French is appreciated but not mandatory.