JOBIM 2024 - Mini-symposium

AI IN THE LIFE SCIENCES – IMPLICATIONS FOR OUR PROFESSIONS

Thursday, June 27, 2024, 2:00 PM - 4:00 PM

Organized by the CATI BIOSBiol:

Gabryelle Agoutin (INRAE, Toulouse), Lucas Auer (INRAE, Nancy), Philippe Bardou (INRAE, Toulouse), Cédric Cabau (INRAE, Toulouse), Patrice Déhais (INRAE Toulouse), Cervin Guyomar (INRAE, Toulouse), Sarah Maman (INRAE, Toulouse), and Géraldine Pascal (INRAE, Toulouse).

AI, and in particular generative AI, is undoubtedly one of the major events that is disrupting and will continue to disrupt our professions. Widespread access to a variety of tools such as ChatGPT, Copilot or Gemini is fundamentally changing our daily lives. In many aspects of our activities, from software development to writing support and bibliography, these tools have become indispensable aids. However, these changes are not without challenges. Ethical and environmental considerations, transparency of algorithms, and the need for close collaboration between researchers and developers are essential issues to be addressed to ensure responsible and ethical use of AI in research.

In this symposium, we propose two sessions. The first aims to explore the functioning of Large Language Models (LLMs), address the general limitations of machine learning approaches, and

examine the uses and applications of these LLMs, as well as the risks involved. In the second session, we propose to focus on the use of generative AI in scientific publication by exposing the misconduct and fraudulent practices that pollute the scientific literature.

PROGRAM: AI IN THE LIFE SCIENCES – IMPLICATIONS FOR OUR PROFESSIONS

2:00 PM: Introduction to the mini-symposium by the organisers.

2:05 PM - 2:50 PM: **VINCENT GUIGUE** (AgroParisTech Paris-Saclay), Professor of Computer Science at AgroParisTech Paris-Saclay. His work, focused on machine learning architectures, is structured around three axes: time series analysis, since his thesis; natural language processing and information extraction tasks; profile learning for recommendation systems. Over the last twenty years, machine learning has been a very dynamic field of research. But the speed at which approaches have evolved in the last decade is astonishing: such a breakthrough has probably never been observed in the history of science. These advances have implications for society as a whole and pose many challenges. Vincent Guigue's research is highly applied and relies mainly on representation learning, which allows the decomposition and analysis of different types of inputs while imposing business constraints.

> Vincent Guigue will present
>
> - How LLMs work and the specificities of chatGPT
> - The general (technical) limitations of machine learning approaches
> - The uses and applications of LLMs and technical developments around other generative AIs
> - The risks involved

2:50 PM - 3:05 PM: Q&A session

3:05 PM - 3:50 PM: **GUILLAUME CABANAC** (IRIT, University of Toulouse is Professor of Computer Science at the University of Toulouse III - Paul Sabatier and holds a research chair at the Institut universitaire de France (IUF) entitled "Cleaning up the scientific literature". His work aims at identifying unreliable publications through text mining, notably within the ERC Synergy project 'Nanobubbles', which questions the self-correction process of science. He is developing the 'Problematic Paper Screener', which detects unreliable publications, often published and sold by leading publishers. On a daily basis, his reports of problematic papers,

particularly those with convoluted wording, have led to hundreds of retractions. His research has been recognised in the "Nature's 10" ranking of 3:50 PM - 4:05 PM: Q&A session and closing remarks "ten people who have contributed to shaping science in 2021" according to the journal Nature.

Guillaume Cabanac will present:

- Cases of misuse of ChatGPT text and data generation in scientific articles
    - Publication of fictitious data
    - Impersonation of authors,
    - The phenomenon of paper mills.
- Problematic Paper Screener: a tool to detect suspicious publications
- Elements of good practice

3:50 PM - 4:05 PM: Q&A session and closing remarks

Vincent Guigue

Guillaume Cabanac