Campagne 2020 Contrats Doctoraux Instituts/Initiatives

Proposition de Projet de Recherche Doctoral (PRD)

Appel à projet HuMed - Humanités médicales 2020

Intitulé du Projet de Recherche Doctoral : Entrance Patterns and Epistemic Impact of Artificial Intelligence in Contemporary Medicine

Directeur de Thèse porteur du projet (titulaire d'une HDR) :

NOM : **Dubois** Prénom : **Michel**
Titre : Directeur de Recherche ou
e-mail : michel.dubois@cnrs.fr
Adresse professionnelle : 59 - 61 rue Pouchet, 75017, Paris
(site, adresse, bât., bureau)

Unité de Recherche :
Intitulé : Gemass
Code *(ex. UMR xxxx)* : UMR8598

Ecole Doctorale de rattachement de l'équipe & d'inscription du doctorant : **ED433-Concepts et langages**

Doctorants actuellement encadrés par le directeur de thèse (préciser le nombre de doctorants, leur année de 1ere inscription et la quotité d'encadrement) : 2 doctorants en co-direction, l'un inscrit en 2016, l'autre en 2019.

Co-encadrant :

NOM : **Gargiulo** Prénom : **Floriana**
Titre : Chargé de Recherche ou HDR ☐
e-mail :

Unité de Recherche :
Intitulé : Gemass
Code *(ex. UMR xxxx)* : UMR8598

Ecole Doctorale de rattachement : **Choisissez un élément :**
Ou si ED non Alliance SU :

Doctorants actuellement encadrés par le co-directeur de thèse (préciser le nombre de doctorants, leur année de 1ere inscription et la quotité d'encadrement) :

Cotutelle internationale : ☒ Non ☐ Oui, précisez Pays et Université :

Description du projet de recherche doctoral *(en français ou en anglais)*

3 pages maximum – interligne simple – Ce texte sera diffusé en ligne

Détailler le contexte, l'objectif scientifique, la justification de l'approche scientifique ainsi que l'adéquation à l'initiative/l'Institut.
Le cas échéant, préciser le rôle de chaque encadrant ainsi que les compétences scientifiques apportées. Indiquer les publications/productions des encadrants en lien avec le projet.
Préciser le profil d'étudiant(e) recherché.

Entrance Patterns and Epistemic Impact of Artificial Intelligence in Medicine

Context: In the history of science, several technical and/or methodological paradigms have been able to generate cascades of scientific innovations in apparently disconnected disciplines: it is the case, for example, of the introduction of calculus in mathematics, at the middle of the 17th century, that allowed to formulate the fundamental laws of mechanics or the developments in electromagnetism at the beginning of the 20th century. More recently, it is also the case for Artificial Intelligence (AI). AI has experienced an extremely rapid growth in terms of publications. AI research technology is getting outside its traditional academic framework, impacting other independent research fields.

Objectives: This project aims to analyze the penetration patterns of artificial intelligence, as a new epistemic approach, in the medical field. It focuses on three separated case studies: neurology, a discipline in which medical imaging plays a central role; cardiology that has an important tradition in time series analysis ; and finally epidemiology, a domain traditionally open to interdisciplinary "contaminations".

Methods: The project will be conducted with a computational social science approach [1] based on the interplay between quantitative scientometric analysis and qualitative field work in medical research unit. The project requires database analysis but also interviews with researchers, at different stages of their career, in three Sorbonne institutes or research units:

- Institut du Cerveau et de la Moelle épinière
- Unité de recherche sur les maladies cardiovasculaires, le métabolisme et la nutrition (UMRS 1166)
- Institut Pierre Louis D'épidemiologie Et Santé Publique (UMRS 1136)

Different publication metadata sources will be mobilized for the project. In particular the PubMed APIs will be used to extract the information on all the publications, in the last 10 years. A list of keywords related to AI will be created from the Wikipedia glossary pages, in order to identify the papers using AI tools. This list will be extended and validated trough experts' interviews. Further information on the citation patterns of the papers using AI will be retrieved using the Scopus APIs.

Project development: The first part of the project will be focused on the entrance patterns of AI in the selected case studies with a top-down approach going from a general observation of the propagation phenomenon to the in-dept analysis of the most important AI tools for each case study. The most important AI tools, for the disciplines, will be, on one side, extracted from the publication patterns, and on the other side, identified by a first contact survey distributed in the concerned laboratory. Within a general science study framework [2, 3, 4], we will first analyze the spreading mechanisms of AI-related innovations in neuroscience, cardiology and epidemiology, focusing on the role of collaborative structures [5] and of status-related mechanisms [6] on these dynamics. Secondly, with a sociologically informed approach, mixing text mining techniques and qualitative information from the field work, we will compare the different areas in terms of the internal

phylogenetic patterns [7,8] of AI, to understand which subdomains of AI have been mobilized in each domain. At the same time a first survey will be distributed among the researchers of the selected labs and institutes, to know more about their AI tools and which ones are mostly used by their collaborators. This survey will allow to validate and eventually re-interpret the scientometric analysis.

The second part of the project will focus on the impact of AI methods on the internal structure of the case study disciplines. Interviews will be conducted with researchers to understand how these new methods are challenging their preexisting practices. We will select researchers at different stages of their careers, to get possibly different feedbacks on the perception of the new methodologies. Going back to the metadata analysis driven by the expertise collected on the fieldwork, we will analyze if, thanks to AI methodologies, new research strands start to develop or restart from "sleeping beauties"[9] . We will analyze how these eventual strands enter in competition with the pre-existing ones [10] developing controversial scientific patterns. Analyzing, in parallel to the academic publications' repository, the patents delivered in the last years, we will evaluate the interplay between the technical and the scientific innovation, generated by AI based methodologies in the selected domains of medicine.

Institutional aspect and doctoral student profile: This project will be conducted under the supervision of Michel Dubois, CNRS Senior Research Fellow, director of the Gemass (UMR8598) and former Deputy director of the interdisciplinary research unit, Epigenetics, Data and Politics (EpiDaPo, UMI 2006). Michel Dubois is a sociologist of science and technology involved in multidisciplinary research projects. The doctoral student will also work closely with Floriana Gargiulo, Gemass researcher specialized in data mining and social network analysis [11, 12]. For the position, we will look for a doctoral student with a good computational background, above all concerning data extraction and large-scale network analysis. The PhD will be in sociology and therefore the candidate should be willing to be trained at the interface between qualitative and quantitative approaches.

References

[1] Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A. L., Brewer, D., ... & Jebara, T. (2009). Computational social science. Science, 323(5915), 721-723.

[2] Dubois M, La nouvelle sociologie des sciences, Paris, Presses Universitaires de France, 1999

[3] Dubois M., Introduction à la sociologie des sciences, Paris, Presses Universitaires de France, 1998

[4] Dubois M., Social Dynamics of Biomedical Research. An Actionist Perspective on the Sociology of Science, Oxford, The Bardwell Press, 2012, 216p.

[5] Newman, Mark EJ. "The structure of scientific collaboration networks." Proceedings of the national academy of sciences98.2 (2001): 404-409.

[6] Merton, Robert K. "The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property." isis 79.4 (1988): 606-623.

[7] Chavalarias, David, and Jean-Philippe Cointet. "Phylomemetic patterns in science evolution—the rise and fall of scientific fields." PloS one 8.2 (2013).

[8] Chinazzi, M., Gonçalves, B., Zhang, Q., & Vespignani, A. (2019). Mapping the physics research space: a machine learning approach. EPJ Data Science, 8(1), 33.

[9] Ke, Q., Ferrara, E., Radicchi, F., & Flammini, A. (2015). Defining and identifying sleeping beauties in science. Proceedings of the National Academy of Sciences, 112(24), 7426-7431.

[10] Aleta, A., Meloni, S., Perra, N., & Moreno, Y. (2019). Explore with caution: mapping the evolution of scientific interest in physics. EPJ Data Science, 8(1), 1-15.

[11] Floriana Gargiulo, Ilaria Bertazzi, and Sylvie Huet. Rise and fall of reputation in a web of trust : the bitcoin-otc market case. The Journal of Web Science, 7, 2019

[12] Floriana Gargiulo, Auguste Caen, Renaud Lambiotte, Timoteo Carletti. "The classical origin of modern mathematics". EPJ Data Science, 2016, Volume 5, Number1

Merci de nommer votre fichier pdf :
«ACRONYME de l'institut/initiative_2_NOM Porteur Projet_2020 »

à envoyer simultanément par e-mail à l'ED de rattachement et au programme :
cd_instituts_et_initiatives@listes.upmc.fr avant le 30 mars.