Recruitment of a postdoctoral researcher with a fixed-term contract
(12 months with a possibility of extension)

AI, data intelligence and lifestyles prediction

Predicting the evolution of lifestyles in megalopolis with AI and ML methods

Keywords:

Greater Paris Region, lifestyles in megalopolis, location data, vehicle data, enrichment and refinement of synthetic population, artificial intelligence, deep learning, machine learning, prediction

Scientific Context

The post-doctoral researcher is participating in the research project INDIUM *INtelligent Data orIented sUstainable Metropolis*, under the HORIZON research programme, funded by CY Initiative of Excellence.

The ultimate objective of the project INDIUM is to propose sustainable planning strategies, based on data intelligence techniques, in order to support land management strategies and reduce emissions. In our framework, we consider a metropolis/region as a complex system where decision entities' interactions generate movements. In simple words, activities induce mobility. Therefore, the unsustainability of mobility is mostly generated by mindless spatial organisation of activities, creating complex polluting patterns. So, to better understand the driving forces behind those behaviours, we need to describe the complex system that is our territory, by representing territorial morphology and different lifestyles through synthetic populations; then, show by modifying some lifestyles, we could reach much more sustainable consumption and production mechanisms. Literature already showed how the current sprawled morphology is unsustainable for the planet and unhealthy for the human being.

Research questions

Among environmental problems, the structural organization of our territory has the most lasting effects. The spatial dispersal of economic activities, besides the irreversibility of artificialization, makes trips long and polluting. Moreover, this man-made dispersal is intensifying since remote work is becoming a new norm. Families are choosing more remote suburbs for their residential locations, and daily travels generally take place around their residences. If pollutions driven by the use of personal vehicles are not in city centre, they might now more frequently occur in far suburbs. At the same time, the 2050 Zero Land Take and zero emission objectives are more and more difficult to reach. Disruptive evolutions must be undertaken.

To react immediately, the first step of the project is to complete *and* refine the description of current lifestyles, in a systematic manner, by adding spatial-environmental attributes to the population. This helps to give responsibility to everyone, to each of the lifestyles, at right place *and* at the right time. It also means that we could reproduce precisely where and when each represented agent consumes space and generates emissions at a very local scale, called an enriched synthetic population (**ESP**). It is synthetic because we don't trace everyone: we represent equivalent lifestyle by a same agent. What can be conducted, in parallel, is to automate the generation of **ESP**. Indeed, automation could be realised if only longitudinal data series are collected and treated to be in a uniform format. In a longer term, the automatic formulation of **ESP** could be a great contribution to feed agent-based transport models continuously, as well as to allow considerable acceleration of scenarios' creation.

The second step aims at generating **ESP** in a long-term future. We would be using learning and/or Artificial Intelligence methods to develop data intelligence in prediction. We would train data and let data propose possible trend lifestyles, or suggest, for example, “low land consuming” and “low emission” lifestyles, at different geographical scales.

Job description

The recruited person will have following tasks:

1) *Data collection*: you are expected to collect specific data that contributes to enrich the spatial-environmental attributes of existed synthetic population. For each of the agent or decision unit, complete their territorial practices, notably locations of their basic life functions (e.g. housing location, work/education location, purchase location, healthcare location). If possible, it could also be interesting to correlate one’s location (or purpose) with his/her means of transportation while describing the emission characteristics of these means used; for each of the represented agent (we can also call one agent ‘a decision unit’ in economic term).
2) *Pre-evaluation of the quality of data*: you would be verifying if the data provides attributes or variables that allow indicators’ calculation, which is vital for post-micromodelling evaluations; as well as if the data are of good quality and complete.
3) *Data fusion, completion, enrichment and scaling problem:* as the enrichment process is firstly done at the megalopolis level (for the Greater Paris Region), we should verify that the **ESP**’s lifestyles are also realistic and reliable at local level. It means that we should also target local accuracy of mobile and immobile behaviours. This double-levelled enrichment could be useful by both increasing spatial resolution and gaining time in calculation.
4) *Validation of the enrichment test*: it is necessary to remind that enrichment test is successful only if the enrichment algorithm could be duplicated, for different past years. Indeed, we might find different results/performances stemming from different enrichment algorithms. Therefore, we must choose the most appropriate one and repeat it for several years so as to create an **ESP** recording covering a relatively long period of time.
5) *Data-driven prediction*: once a longitudinal series of ESP is recorded, we could train the formulated ESP with AI methods to make future predictions, without the need to recollect all the necessary data, which will allow us to create ESP for the future. Indeed, the time range of this recording is expected to be as long as possible, which might permit cyclic learning process, and allow to compare different AI or learning methods, paving the road for ESP prediction and scenarios making. The ESP prediction refers to a trend scenario, and alternative scenarios refer to hypothesis of sustainable megalopolis scheme and/or healthy lifestyles.

To summarize, the postdoctoral researcher must show a strong capacity in dealing with geographical and environmental data, with focused interests to location data and/or to vehicle data.

Candidate profile

- Holds a PhD degree in statistics, computer engineering, transport engineering, geoinformatics or data science.
- Experience of working with statistical models and skills for data fusion, data completion and data enrichment;
- Knowledge of the transport and mobility, and/or issues related to urbanism particularly to suburbanization would be a plus;
- Spirit of exchange and respect for deadlines.

Working environment

- MATRiS, CY Cergy Paris Université. 33 Boulevard du port, 95000, Cergy

- ETIS, CY Cergy Paris Université. 2 Avenue Adolphe Chauvin, 95302, Pontoise

Supervising Team

Liu LIU, Associate Professor, CY Cergy Paris University, lab MATRiS, liu.liu@cyu.fr

Dimitris KOTZINOS, Professor, CY Cergy Paris University, lab ETIS, Dimitrios.Kotzinos@cyu.fr

Candidatures

Applications (CV, transcripts and a motivation letter) must be sent by email until 3 March 2025 to:

Liu Liu liu.liu@cyu.fr and Dimitris KOTZINOS Dimitrios.Kotzinos@cyu.fr