Opening of a Ph.D. Position on

**Urban Knowledge Hub for Evolving Cities**

<table>
<thead>
<tr>
<th>Title</th>
<th>Urban Knowledge Hub for Evolving Cities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supervisors</td>
<td>Dr. Emmanuel Coquery, Dr. John Samuel, Prof. Gilles Gesquière</td>
</tr>
<tr>
<td>University</td>
<td>Université Lyon 1 (UCBL), LIRIS Laboratory (CNRS UMR 5205)</td>
</tr>
<tr>
<td>Date of recruitment</td>
<td>1 September 2022</td>
</tr>
<tr>
<td>Duration</td>
<td>36 months (3 years)</td>
</tr>
<tr>
<td>Keywords</td>
<td>Knowledge representations, multidimensional data, data evolution, code evolution, knowledge evolution, reproducibility</td>
</tr>
</tbody>
</table>

**Context**

This project proposal for a Ph.D. position has been selected in the IADoc@UdL call. The goal is about thinking about how to capitalize on the amount of knowledge developed during the last decade and use it in a multidisciplinary context for understanding city evolution and its capacity to become more sustainable and resilient. This proposal is made possible thanks to a strong collaboration between LIRIS Laboratory and Metropole of Lyon. This project is fully integrated within the LIRIS laboratory where it brings the necessary transversal approach to talk about smart cities between data science specialists (BD team) and graphic computing specialists (Origami team) since it is a question of keeping the strong link between vector data and associated semantics. The LIRIS is also a key partner when it comes to working in the field of Artificial Intelligence (AI).

As an element of context, research is not just about publications and providing code. Working in an urban context implies rethinking and discovering the unexplored part of knowledge (data, code, literature, and process) like exploring the submerged part of an iceberg. In this thesis, the major theme is **how to propose a real aggregation of this knowledge about the city, an aggregation made in a diachronic context?**

In this proposal, we seek a candidate, with good knowledge of computer sciences (level Bac+5 in France). The position is available in Lyon (campus LyonTech La Doua). More information on this subject is given below.

**More information on this subject**

**Evolution of Knowledge**

The knowledge about the city evolves and each new study can lead to revising assumptions or completing the knowledge about some objects that make up the city. For example, a simple photo exhumed from a newly found archive can narrow the range of a building's existence by demonstrating that it was built later than previously known. The evolution of knowledge also implies taking into account the evolution of the meaning of the vocabulary, in the representation of knowledge about the city. For example, a building that met accessibility standards for disabled people in the 1980s meets certain criteria that may not be sufficient in 2022, as the rules have evolved. The use of this type of information may therefore be different depending on when the
Building was declared compliant, but also on when this information appeared in the knowledge base. Moreover, the knowledge linked to the data itself can be used to intelligently publish this data. For example, when publishing data from the Lyon metropolitan area, the question arises as to whether published derived data will be identical to previously published data or whether a new version should be published. This raises the question of the reproducibility of a calculation, reproducibility linked not only to the initial data but also to the codes used to calculate the derived data. **How can we ensure the evolution of knowledge, as well as the evolution of induced knowledge?**

**Sharing knowledge for everyone**

Moreover, urban studies have often been built on mono-disciplinary approaches (in urban planning, geography, environmental science, sociology, economics, law, etc.). The resulting data and knowledge are therefore by construction heterogeneous [1,2,3,6]. In order to comply with regulations, local authorities have produced a very large mass of data that are quite uneven in terms of their documentation, quality, and capacity to represent the evolution of the territory [3,7,8]. In the need for transparency, it is also necessary to go further by delivering, in addition to the data, the algorithms that govern certain mechanisms of public life. An alignment of this knowledge [2,3,9] is necessary in order to aggregate this knowledge, but also to make it available to a large community. It is necessary to propose a diachronic approach allowing us to build knowledge in the long run. **How can we allow a non-data scientist user to mobilize this heterogeneous knowledge?**

**Towards reproducible results for research and decision making**

In addition, it is also important to integrate process and project knowledge into the scientific workflow in accordance with Linked Data principles. To this end, we propose to build an open and reproducible workflow to integrate [5,14], extract, question [6], reason, infer [11,12,13], and automatically generate new knowledge from multiple, autonomous, heterogeneous, and evolving urban data sources. Repeatability (obtaining the same results by repeating the exact same computations) and reproducibility (obtaining the same results by different means) are both key aspects of research. They allow us to confirm results or lead to a better understanding. The replicable approach will allow all stakeholders to not only access and interact with the data and knowledge produced, but also to verify and share the data, knowledge, code and this automated process to explain the reasoning behind their decisions. **How to ensure knowledge related to the processes of production of data and algorithms to ease the reproducibility of results for both research and decision making?**

**Scientific challenges related to the project**

In this thesis, several challenges are identified:

- Management of temporal graph data on knowledge (related to code, data, project, workflow)
- Ability to generate new information about derived knowledge; how to manage it in a context of knowledge versioning needs, in particular in a context of reproducibility of experiments.
- To propose new concepts of organization and interrogation of knowledge allowing it to scale.

**Technical challenges related to the project**

- Ability to propose a prototype based on the LIRIS Pagoda\(^1\) and UD-SV\(^2\) platforms and to demonstrate the feasibility of the proposed approaches.
- Ability to propose the necessary tools for use by non-data scientists.

**Application**

Interested persons should send a mail to Prof. Gilles Gesquière ([gilles.gesquiere@universite-lyon.fr](mailto:gilles.gesquiere@universite-lyon.fr)), Dr. Emmanuel Coquery ([emmanuel.coquery@univ-lyon1.fr](mailto:emmanuel.coquery@univ-lyon1.fr)) and Dr. John Samuel ([john.samuel@cpe.fr](mailto:john.samuel@cpe.fr)) with the following:

1. CV
2. Motivation letter

---

1. [https://projet.liris.cnrs.fr/pagoda/latest/](https://projet.liris.cnrs.fr/pagoda/latest/)
2. [https://github.com/VCityTeam/UD-SV/](https://github.com/VCityTeam/UD-SV/)
3. Academic transcripts
4. List of publications (if any)
5. 2 reference letters (or persons to contact)

References


