

Opening of a Ph.D. Position on *Urban Knowledge Hub for Evolving Cities*

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

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 Pagoda¹ and UD-SV² 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), Dr. Emmanuel Coquery (emmanuel.coquery@univ-lyon1.fr) and Dr. John Samuel (john.samuel@cpe.fr) with the following:

1. CV
2. Motivation letter

¹ <https://projet.liris.cnrs.fr/pagoda/latest/>

² <https://github.com/VCityTeam/UD-SV/>

3. Academic transcripts
4. List of publications (if any)
5. 2 reference letters (or persons to contact)

References

- [1] *Towards Limiting Semantic Data Loss in 4D Urban Data Semantic Graph Generation*, Diego Vinasco-Alvarez, **John Samuel**, Sylvie Servigne, **Gilles Gesquière**, 3D GeolInfo 2021, le 11-14 octobre 2021, En-ligne, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, VIII-4/W2-2021, 37-44
- [2] *Integrating multimedia documents and time-evolving 3D city models for web visualization and navigation*, Vincent Jaillot, Valentin Rigolle, Sylvie Servigne, **John Samuel**, and **Gilles Gesquière**, Transactions in GIS, mars 2021
- [3] *Representation of concurrent points of view of urban changes for city models*, **John Samuel**, Sylvie Servigne, **Gilles Gesquière**, Journal of Geographical Systems, février 2020
- [4] *UrbanCo2Fab: Comprehension Of Concurrent Viewpoints Of Urban Fabric Based On GIT*, **John Samuel**, Sylvie Servigne, **Gilles Gesquière**, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume IV-4/W6, 2018, 13th 3D GeolInfo Conference, le 1-2 octobre 2018, Delft, Hollande
- [5] *Integration of Multiple Heterogeneous and Autonomous Web Services using Mediation Approach: Open Challenges*, **John Samuel**, Christophe Rey, Special Issue on Digital Information Management, Journal on Advances in Theoretical and Applied Informatics - JADI 2016, Volume 2, Numéro. 2, Pages 38-46
- [6] *Heterogeneous Sensor Data Exploration and Sustainable Declarative Monitoring Architecture: Application to Smart Building*, Sylvie Servigne, Yann Gripay, Ozgun Pinarer, **John Samuel**, Atay Ozgovde, and Jacques Jay, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume III-4/W1, 2016, 1st International Conference on Smart Data and Smart Cities, 30th UDMS, le 7–9 septembre 2016, Split, Croatie
- [7] *Visualization of Documented 3D Cities*, Clément Chagnaud, **John Samuel**, Sylvie Servigne, **Gilles Gesquière**, Eurographics Workshop on Urban Data Modelling and Visualisation, UDMV 2016, Liège, Belgique, le 8 décembre, 2016
- [8] *Representation and Visualization of Urban Fabric through Historical Documents*, **John Samuel**, Clémentine Périnaud, Sylvie Servigne, Georges Gay, **Gilles Gesquière**, 14th EUROGRAPHICS Workshop on Graphics and Cultural Heritage, Gênes, Italie, le 5-7 octobre, 2016
- [9] *Combining Graph Exploration and Fragmentation for Scalable RDF Query Processing*, Abdallah Khelil, Amin Mesmoudi, Jorge Galicia, Ladjel Bellatreche, Mohand-Saïd Hacid, **Emmanuel Coquery**, Information Systems Frontiers, vol. 23, no. 1, février 2021, p. 165 —83.
- [10] *Beyond the Hype: Big Data Concepts, Methods, and Analytics*, Amir Gandomi, Murtaza Haider, International Journal of Information Management, vol. 35, no 2, avril 2015, p. 137–144.
- [11] *Access Control to Materialized Views: An Inference-Based Approach*, Sarah Nait Bahloul, **Emmanuel Coquery**, Mohand-Saïd Hacid, Proceedings of the 2011 Joint EDBT/ICDT Ph.D. Workshop on - PhD '11, ACM Press, 2011, p. 19–24.
- [12] *RQL: A Query Language for Rule Discovery in Databases*, Brice Chardin, **Emmanuel Coquery**, Marie Pailloux, Jean-Marc Petit, Theoretical Computer Science, vol. 658, janvier 2017, p. 357–374.
- [13] *Inference Leakage Detection for Authorization Policies over RDF Data*, Tarek Sayah, **Emmanuel Coquery**, Romuald Thion, Mohand-Saïd Hacid, Data and Applications Security and Privacy XXIX, Springer International Publishing, 2015, p. 346–361.
- [14] *Interactive Mapping Specification with Exemplar Tuples*, Angela Bonifati, Ugo Comignani, **Emmanuel Coquery**, Romuald Thion, ACM Transactions on Database Systems, vol. 44, no 3, juin 2019, p. 10:1-10:44.