

# Colouring Cities: A Global Network and Open Platform for Collaborative Mapping of Building Stocks

Polly Hudson <sup>a</sup>, Oriol Gavalda <sup>b</sup>, Robert Hecht <sup>c</sup>, Falli Palaiologou <sup>d</sup>, Merve Keskin <sup>e,\*</sup>

<sup>a</sup> CCRP PI, Colouring Cities Research Programme, phudson@turing.ac.uk

<sup>b</sup> Next-Generation Cities Institute, University Concordia - oriol.gavalda@concordia.ca

<sup>c</sup> Leibniz Institute of Ecological Urban and Regional Development, Dresden, Germany - r.hecht@ioer.de

<sup>d</sup> CCRP Co-I, School of Architecture Building and Civil Engineering, Loughborough University, UK, g.palaiologou@lboro.ac.uk

<sup>e</sup> Department of Geoinformatics (Z\_GIS), Paris-Lodron University of Salzburg (PLUS), merve.keskin@plus.ac.at

\* Corresponding author

**Keywords:** Participatory mapping, platform, buildings, SDG, open source, open data, CCRP

## Abstract:

Any academic research relating to UN Sustainable Development Goals (SDGs) involving geospatial analysis, will first require access to geospatial data. Where open, high-quality, standardised geospatial data are made available, analysis necessary for SDG-related problem-solving, in the environmental, economic and socio-cultural contexts, is likely to be faster and easier. Furthermore, when opportunities for multidisciplinary, cross-country collaborations and the application of artificial intelligence (AI) and machine learning (ML) to uncover insights into spatial data patterns also exist, an even more substantial acceleration of progress is possible.

The Colouring Cities Research Programme (CCRP) is an informal, non-competitive research initiative that addresses issues with access, quality and reliability of spatial data relating to buildings, and built and green/blue infrastructure (Hudson, 2025). Also, it addresses common blockers to collaboration within academia. The CCRP involves an informal consortium of international researchers interested in knowledge and resources, and co-developing reproducible open tools to help accelerate SDG-related research at the local, national and global scale. The CCRP has been developed within academia over the past decade as a decentralised system, and today engages academic institutions from over thirty countries (Fig 1.), grouped into six CCRP Global Hubs that operate Colouring Cities platforms following CCRP protocols.

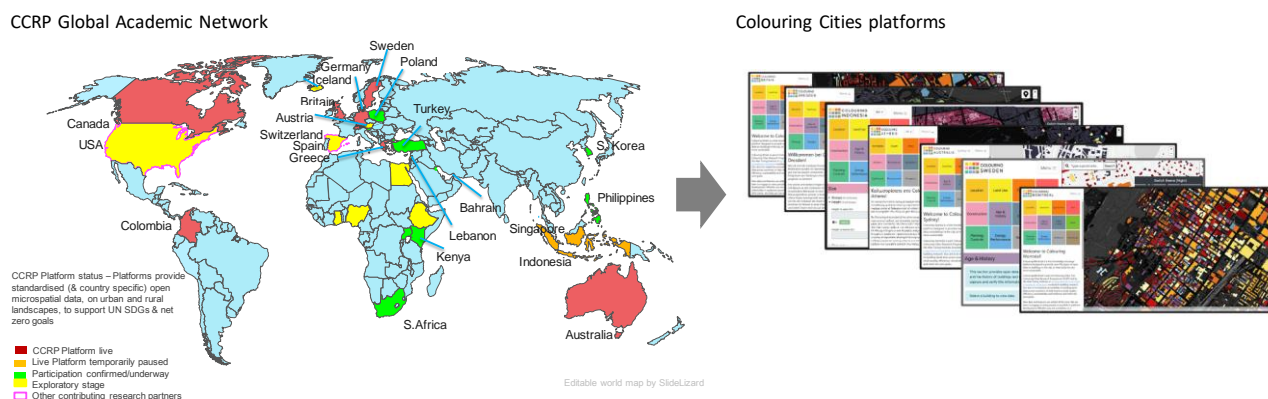


Figure 1. Overview map of the global academic network and examples of the Colouring Cities platforms

(Source: <https://colouringcities.org/>)

At the core of CCRP are a network of open-source, national data platforms that facilitate knowledge collation and sharing of microspatial data on buildings and built and green/blue infrastructure. CCRP academic partners independently manage national platforms and employ standard interfaces to collect, collate, standardise, verify, visualise and release spatial data. All CCRP codes, data and methods are made available under open licences. Initially run by research groups interested in specific SDG-related research questions, CCRP platforms are designed to evolve incrementally into low-cost, permanent open national databases managed by academia and co-created with/used by government, non-profits, industry and citizens. The platforms are supported by CCRP Expert Data Groups, which are set up for each of the CCRP's main data

categories. Experts provide specialized guidance on specific data classes and formats, as well as security and privacy considerations, and collaboratively investigate data applications, while also exploring the strategic integration of Artificial Intelligence (AI) and Machine Learning (ML) in data analysis and modeling to drive informed decision-making and innovation. The CCRP also has an international technical group, developing CCRP features and plug-ins across countries, an impact section for case studies, and an informal operations group which oversees the CCRP GitHub open code repository, platform manuals and other core open resources.

The web-based, interactive CCRP platforms facilitate the collection of over 150 data classes related to the built and natural environment, structured within twelve main data categories. These categories capture information on current building attributes, such as location, land use, typology, size, materials/construction, and age, as well as long and short-term dynamic behavior, planning controls, building performance, building team and energy aspects, blue and green infrastructure context, and the extent of damage to buildings in disaster situations. Most data are standardized, but locally specific bespoke data can be collected, provided the related open code is shared. Academic oversight prioritizes the ethical collection of data and the privacy and security of users and building occupiers in all aspects of CCRP design and operation. For example, the CCRP does not release data on the interior of homes (in most countries) and does not link free text boxes to georeferenced building polygons. Specific trust models are being discussed within the CCRP to enable the secure collection and analysis of more sensitive data within academia.

Platforms are coded to integrate multiple methods of spatial data capture. The focus is on existing bulk uploads (moderated by academia), live/near-time data streamed from official APIs, new large-scale computationally inferred data, and crowdsourcing at the building level. Each method is designed to support specific stakeholder groups and maximize data quality, accuracy, and completeness. APIs, for example, are well-suited to stream government datasets, such as planning data, where frequent updates are necessary. Crowdsourcing of data building by building is effective for local residents, schools, civic societies, and local historians, as buildings are coloured immediately when data is added. Users are responsible for determining whether the data are fit for purpose and are provided with as much information as possible to assess data reliability. This includes information on data capture methods, source links, edit history (similar to OpenStreetMap), and uncertainty quantification features, which are also planned.

CCRP partners are also beginning to work on feedback loops between core data collection methods to drive cross-referencing, with integration of uncertainty quantification features also planned. Researchers behind the German initiative Colouring Dresden (Danke et al. 2024), which has a focus on citizen science approaches, are now investigating the possibilities of using crowdsourced data as training data to derive building characteristics for the entire city using machine learning. This derived data can in turn be uploaded and verified by the crowd. The algorithms and methods used in the study will be published and shared for reproduction and improvement in collaboration with other CCRP research partners.

One of the last future developments of the platform is to allow for the easier deployment of instances while guaranteeing robustness between the versions. A database-based web deployment structure is being evaluated, combined with a plugin architecture, to facilitate the maintenance and updating of the core repository and enable seamless integration of each of the different platforms.

### Acknowledgements

The authors would like to thank all contributors to the Colouring Cities Research Programme.

### References

- CCRP, 2025. Colouring Cities Research Programme. Website. Available from: <https://colouringcities.org/>
- Hudson P., 2025. Colouring Cities Research Programme's Open Manual. Wiki. Available from: <https://github.com/colouring-cities/manual/wiki/>
- Danke, T., R. Hecht, and T. Rieche. 2024. Closing Knowledge Gaps on the Building Stock with Citizen Science: Introducing the Colouring Dresden Platform. *ARPHA Proceedings*, 159–164 6. <https://doi.org/10.3897/ap.e125607>