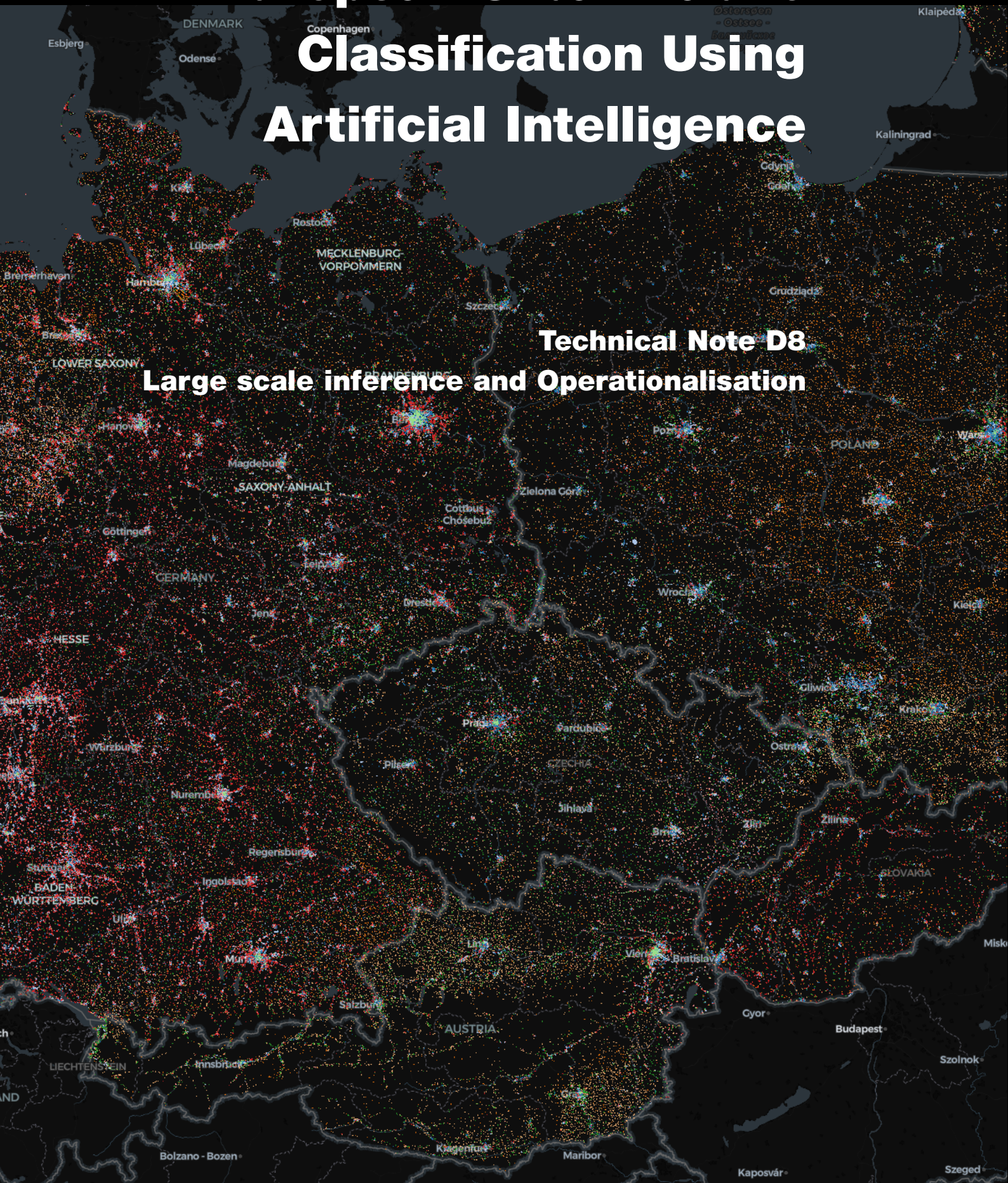
European Urban Fabric Classification Using Artificial Intelligence



Technical Note D8

Large scale inference and Operationalisation

Large scale inference and Operationalisation

Technical Note D8

Krasen Samardzhiev[1], Barbara Metzler[2], Martin Fleischmann[1,*] and Daniel Arribas-Bel[2]

1. Faculty of Science, Charles University, Social Geography and Regional Development, Albertov 6, 128 00, Praha, Czechia
2. The Alan Turing Institute, Science of Cities and Regions, British Library, 96 Euston Road, London NW1 2DB, UK

* *Correspondence:* Martin Fleischmann martin.fleischmann@natur.cuni.cz

The activities reported in this document were carried under a programme of, and funded by, the European Space Agency. View expressed in the publication can in no way be taken to reflect the official opinion of the European Space Agency.

eurofab.org

Basemaps © CARTO, © OpenStreetMap contributors.

Table of contents

1 Introduction **1**

2 European Morphometric Classification Strategy **2**
2.1 Infrastructural, technical and methodological aspects involved in scaling . 2
2.2 Risk assessment 3
2.3 Europe-wide classification strategy 3

3 European Space-Time Urban Fabric Strategy **5**
3.1 Infrastructural, technical and methodological aspects involved in scaling . 5
3.2 Risk assessment 6
3.3 Europe-wide deployment strategy 6

4 Scaled up stakeholder engagement **8**

5 Productionalising **10**

1 Introduction

This Technical Note presents a strategic roadmap for advancing the EuroFab project's urban fabric classification system from its current capabilities to a pan-European operational capacity.

Ultimately, EuroFab seeks to pave the road for a world where stakeholders, from local authorities to supranational organisations, are able to track and monitor the patterns of European urban development across time, in detail directly relevant for planning. The pilot study in Central Europe and the UK consisted of three related parts - a machine learning morphometric pipeline, a AI vision model, and a stakeholder engagement component. The current pilot study demonstrated the viability of using alternative data sources to produce multiscale descriptions of the built environment, based on analysis of individual buildings, street segments and their interactions. Furthermore, the stakeholder engagement part highlighted the multiple applications which stakeholders see for such a dataset.

The overarching strategy for scaling and productionalising the EuroFab system involves three principal phases. First, using the morphometric pipeline to generate detailed urban classification at a pan-European scale. Second, training the AI vision model on the outputs from step one, to fill data gaps and to produce a temporal classification based on historical satellite data. Third, productising the refined classification results and expanding stakeholder engagement activities. This will be crucial for driving user adoption and facilitating the derivation of secondary indicators and specialised datasets tailored to specific applications, such as regional development analysis or climate change impact assessments.

Transitioning the EuroFab system from a successful research project with functional processing pipelines to a robust, reliable, and continuously updated data product presents considerable technical and organisational challenges. Serving a diverse range of stakeholders across Europe with consistent and high-quality urban fabric information requires improvement and upscaling of data processing capabilities, the establishment of a sustainable operational framework that covers potential risks, and a commitment to continuous improvement and user engagement.

The roadmap in this technical note addresses these multifaceted challenges, providing a structured approach to achieve full operationalisation.

2 European Morphometric Classification Strategy

The results from the pilot study demonstrated the potential to deliver pan-European classification of urban fabric as a combination of morphological processing and predictive modelling. Specifically, that it is possible to fill the gaps in openly available cadastral data with alternative data sources, namely Overture Maps and Microsoft Building footprints. The resulting classification can be used by stakeholders on its own, but it will also be the training data for the expanded temporal AI models.

2.1 Infrastructural, technical and methodological aspects involved in scaling

The pan-European pipeline will extend the methodologies developed during the pilot study and will comprise four distinct stages. First, calculating the morphological characteristics on alternative building footprints. Second, assigning labels to subsets of the data, where the cadastral classification is available. Third, training a model using the train/test schemes devised in the pilot study. And lastly, using the models to predict urban fabric types based where cadastral data is not available.

There are three types of expected changes to the current methodology.

First, the methodology will require updates to ensure compatibility with the latest versions of relevant open-source geographical software packages, in order to keep as close as possible to the cadastral classification. The ground truth cadastral classification is a dynamic product, with ongoing updates influencing upstream open-source repositories and benefiting from contributions from a diverse community. Therefore, it changes over time, both adapting and contributing to the state-of-the-art morphological open source software. The effort required for these updates is anticipated to be low to medium, as the work will predominantly involve technical software development rather than fundamental methodological revisions.

Methodologically, it is expected that the first and second stages will need only minor changes. However, a significant data pre-processing effort will be necessary. This includes comprehensively mapping available data resources across Europe and rigorously assessing their quality and consistency. For example, this involves evaluating the completeness,

geometric accuracy, and attribute information of building footprint datasets from various national and regional providers. The third and fourth stage may require more adaptions in order to account for the new scale and data. Potential pitfalls and solutions are discussed in the subsequent sections.

There are no new infrastructural requirements expected. The pilot methodology is expected to scale on the currently available computing resources.

2.2 Risk assessment

The principal risks identified in scaling the morphometric classification include:

- **Sourcing and processing European building footprints.** An extensive mapping of the available data sources needs to be carried out, before applying the methodology to the European continent. A key finding from the pilot work is the difficulty of working with cadastral data, primarily because every country, or even different regions within the same country, have different definitions of what a building is. The available data needs to be downloaded and standardised for the first two steps of the analysis. This could potentially result in data from different sources, in different parts of Europe.
- **Data Gaps in Specific Regions:** It is possible that there will be regions with no available building footprint data at all. These regions will need to be handled differently in order to provide consistent coverage.
- **Unknown urban fabric classes:** It is possible that there are some unique urban fabric classes in Europe, not present in the underlying ground truth cadastral classification. This will lead to exclusion from the model training data and degraded model performance.

2.3 Europe-wide classification strategy

- **Alignment with Cadastral Classification:** Maintaining continuous alignment with the evolving cadastral classification is crucial to keeping the morphometric pipeline up-to-data. There are no major challenges expected for this work, however it will require continuous updates of the entire pipeline, including models and data processing. Nevertheless, it is expected that each update of the ground truth data will improve the performance of the models.
- **Heterogenous data sources:** The scope of the pilot study was already extended so that the models and morphological processing pipeline can handle heterogenous data. The results suggest that better quality data from heterogenous sources, specifically Overture Maps data, leads to more accurate predictions. However,

the underlying building data from new European regions can have previously unseen issues in the pilot study area. These could include issues such as topological inconsistencies (e.g., overlapping building polygons), incomplete coverage, or discrepancies in the street network data used for contextual analysis. While these challenges are considered solvable, they will likely increase the complexity and volume of the required data pre-processing work. Extending and improving the Overture Maps data processing pipeline is a good starting point for all of these and other data integration tasks.

- **Data Gaps:** In regions where there are no building footprints available at all, the vision model can be used directly to fill in the gaps, since Sentinel 2 has full coverage of Europe. However, such a region could potentially represent an unobserved urban class, and therefore needs to be analysed in more detail.
- **Unknown urban fabric types:** The likelihood of unseen urban classes lessens as the geographical scope of the ground truth cadastral classification grows. The current classification covers Germany, Poland, Lithuania, Slovakia, Czechia and Austria. There is cadastral data available for other European Countries such as Spain, Belgium, France and the Netherlands. It is expected that these will be included in subsequent cadastral classifications and therefore increase the types of urban fabric covered. This prospect further increases the importance of continuously aligning the existing morphometric pipeline with the cadastral ground truth, as well as with the latest advancements in open-source geospatial software.
- **Model Tuning and Optimisation:** The current morphometric models are HistGradientBoostingClassifier with unlimited depth and number of nodes, trained on hundreds of thousands of examples per class. However, this scheme will need to change as the size of the datasets increases, or alternatively multiple models will need to be trained and tested. The detailed spatial testing schemes implemented in the pilot study can guide the development of the new models, where the main anticipated challenges will be hyper-parameter tuning, addressing class imbalances and accurate subletting of the training data. The data availability and data pre-processing work will also have a large impact on the model tuning and training. It is possible that only specific types regions lack cadastral data, or are not covered by the ground truth classification and therefore will require most of the modelling focus.

3 European Space-Time Urban Fabric Strategy

3.1 Infrastructural, technical and methodological aspects involved in scaling

The core of EuroFab's current Earth Observation analytical capability is the Spatial Signatures Pipeline, detailed in the eurofab-project/eo GitHub repository. This pipeline is engineered to generate spatial signature predictions using satellite imagery as its primary input. First, it integrates computer vision models to create feature embeddings from Sentinel 2 imagery. These embeddings are subsequently processed by an XGBoost model, which performs the classification of individual image chips -groups of pixels - into predefined urban fabric classes. The following aspects of this baseline processing chain needs to be adapted and scaled:

- **Sourcing and processing historical European Sentinel 2 data:** The current pipeline is processing historical Sentinel 2 data only for the UK. The data processing and storage have to be expanded to cover the whole of Europe, which will exponentially increase the size of the data that needs to be stored and processed.
- **Classification Labels:** The current XGBoost model assigns each grid cell to one of twelve predefined spatial classes and their aggregations. These classes include: 'Accessible suburbia', 'Connected residential neighbourhoods', 'Countryside agriculture', 'Dense residential neighbourhoods', 'Dense urban neighbourhoods', 'Disconnected suburbia', 'Gridded residential quarters', 'Open sprawl', 'Urban buffer', 'Urbanity', 'Warehouse/Park land', and 'Wild countryside'. This typology provides a foundational characterisation of diverse urban and peri-urban environments. However, spatial signatures are restricted to the geographical extent of Great Britain. The pan-European product will use as ground truth, a combination of the existing morphological classification and the predicted classes from the morphological pipeline, where the cadastral data is not available. This change will generate orders of magnitude more training data for the vision and XGBoost models, from all across Europe. This change will not just increase the scale of the results, but also their accuracy as shown by the conclusions of the pilot study.
- **Infrastructural Capacity:** The current inference pipeline and its training is not heavily computationally demanding as it was trained on a single high-end

GPU. However, the training times were not optimal and for a scaled-up pipeline, the computational demands will grow significantly. Therefore, it is expected that training and inference will require a small cluster of high-end GPUs, if the inference will be ran multiple times within a year.

3.2 Risk assessment

The scaling of the AI models for pan-European space-time urban fabric analysis presents three primary challenges:

- **Large scale evaluation of urban fabric change across time**: The pilot analysis demonstrated that urban fabric classifications exhibit distinct temporal and spatial dynamics, reflecting varying levels of stability and diversity over time. Urban and sub-urban classes displayed higher probabilities of transition, indicating active urban transformation. Conversely, rural classes showed significant stability. As highlighted in the pilot's results these observed changes may be more related to classifier uncertainty than actual environmental changes.
- **Generalisability**: Evaluating the generalisability of the AI methodological framework between different regions is crucial for its wider applicability. Due to the limited geographical scale of the pilot study this was not possible.
- **Handling Misclassifications**: Misclassifications typically occur between visually similar urban fabric classes, indicating inherent uncertainty in predictions. Incorporating prediction probabilities into a secondary model could help address this issue. By explicitly using probability scores from the initial classification as input for a refinement model, we could better distinguish between ambiguous cases. This approach may "smooth" predictions, reducing noise and improving overall classification accuracy. The pan-European work should explore how prediction confidence scores can be systematically utilised, either by employing spatial smoothing algorithms or by applying secondary machine learning models trained specifically to correct uncertain predictions.

3.3 Europe-wide deployment strategy

- **Leveraging Copernicus Services:** The Copernicus programme, with its commitment to free, full, and open data access, will be the cornerstone for sourcing satellite imagery. Primarily, data from the Sentinel-2 (multispectral optical) mission will be used to derive consistent historic coverage of the urban fabric in the entire European continent. These data will be the core input for the raster-based analyses performed by the EuroFab pipeline and will enable the temporal classification and change detection of urban fabric.

- **Leveraging the new morphometric classification results as ground truth:** It is expected that replacing the ground truth classes from spatial signatures to the new morphometric results will not lead to significant changes in the pipeline. Although the spatial signatures capture functional information, almost all morphometric characters used in deriving the new classification are already used in the spatial signatures derivation. Therefore, the current AI vision model pipeline is already processing urban form information. There is one significant change that will need to be accounted for. The new ground truth classification focuses exclusively on areas that have buildings, therefore large expanses of non-built up area such as fields and forests are excluded by definition. This will actually reduce processing times and help with class imbalances, however some changes in the processing pipeline will be neccassary.
- **Generalisability Testing**: The performance of the vision model will be tested using the same robust generalisability testing framework used in the pilot study's morphological pipeline. This was not possible for the pilot AI pipeline due to the limited extent of the available data. However, with the expansion of the study area data more complex testing scenarios become viable. Furthermore, the generalisability task becomes easier as the scale of the morphometric pipeline used to generate the training data increases to cover the whole of Europe. Rather than having to predict urban form in multiple entirely unseen contexts, the vision model will have to do in-fill predictions - —classifying urban fabric for regions where contextual information from surrounding, already classified areas is available, thus benefiting from a richer, more representative training dataset.
- **Evaluation of urban predictions across time** Similarly, the larger quantity of data, as well as the new testing frameworks will enable better model performance and more certainty in the temporal classification results. Furthermore, the expanded stakeholder engagement and derivation of secondary datasets will provide valuable real-world validation, and increased confidence in the results.
- **Methodological processing** The existing AI pipeline architecture relies on localised information - processing image pixels and chips with limited local context - to generate embeddings and subsequent predictions. Therefore, it is anticipated that memory and computing requirements will scale in a broadly linear fashion with the increase in data volume, without necessitating fundamental changes to the core methodology. However, it is expected that significantly more resources will be required for preprocessing the data and for training generating the embeddings from the satellite images for the whole of Europe. In order to do these steps in a reasonable timeframe it is expected that training and inference will require a small cluster of high-end GPUs, with the inference ran on a yearly basis.

4 Scaled up stakeholder engagement

The stakeholder work in the pilot study confirmed the need for a detailed urban classification at a granular level. The engagement also resulted in multiple feature requests for secondary indicators, specific functionality and reports. The new pan-European detailed urban classification has the potential to serve as a foundational dataset for a multitude of derived data products, each adapted to particular use cases and stakeholder requirements. Examples of requested applications included: comparative analyses between countries at various spatial scales (neighbourhoods, cities, regions); direct access to building-level morphological data for micro-scale studies; and predictive classification capabilities to assess how newly designed or planned urban areas might be categorised before construction. Furthermore, stakeholders expressed strong interest in combining the EuroFab classification with other complementary datasets, such as socio-economic data, environmental hazard maps, or transport network information.

The expanded geographical scale and enhanced detail of the new pan-European results will enable us to continue and enhance the stakeholder work in several ways.

- **Workshops and conference work:** Experience from the pilot phase demonstrated that workshops and conference presentations are highly effective mechanisms for fostering collaborations and gathering valuable insights. The pan-European scale of the new product will increase the number of relevant and interested stakeholders. Therefore, sustained engagement through targeted workshops and active participation in key European conferences is expected to continue delivering substantial value to the project as a whole.
- **Engagement with Specific Cities and Regions:** Stakeholders consistently expressed interested in their local areas or countries, as well as larger pan-European data. Even the pilot study extent of five Central European countries and the UK proved limiting. For example, the inclusion of SSVA (Construction Sector Development Agency of the Ministry of Environment, Lithuania) as one of the key stakeholder engagements, neccesicitated the inclusion of Lithuania in the some parts of the pilot's research. The comprehensive pan-European scale of the operational product will significantly streamline the process of onboarding new interested stakeholders, as their specific geographical areas of interest will inherently be covered.
- **Collaborations & co-development:** The focus of the current stakeholder work was on establishing relationships and gathering feedback. Extended collaborations

on specific applications and data products was not possible due to the short duration of the pilot. The expanded scope and duration of the pan-European classification project will lead to stakeholder engagement beyond simple feedback collection. This can involve establishing formal partnerships where stakeholders actively contribute resources – such as local datasets for training/validation, domain expertise, or personnel for joint validation campaigns – and, in return, share in the benefits and responsibilities of the new data products. Such collaborative ventures are designed to deeply embed the EuroFab service and its outputs into the stakeholders' existing operational workflows and critical decision-making processes, thereby maximising its real-world impact.

- **Developing derived products:** For specific, strategically important use cases (e.g., monitoring progress towards Sustainable Development Goal 11 on sustainable cities), the project can extend its stakeholder engagement to actively guide the co-development of tailored derived products. This work will also result in detailed examples and documentation of how to use the foundational space-time urban classification product for the development of derived indicators in different policy areas. Some desirable indicators that came up during the development work were combining the classification with land use data and geodemographic factors, generating summaries of the morphometric character of cities, access to the full taxonomy, as well as raw morphometric characters data at the building level.
- **Supporting Third-Party Development of Derived Products:** The pilot study also revealed interest from commercial entities and research groups in developing their own derived products for specific niche applications. These were local consultancies and private companies operating in sectors such as urban planning, real estate analytics, and environmental management. A framework for ongoing engagement and technical support can be established to facilitate and encourage these third-party developments, potentially through API access or developer toolkits.
- **Open source community:** EuroFab's foundational commitment to openness represents a key strategic asset for ensuring its long-term impact, widespread adoption, and overall sustainability. This includes maintaining the public GitHub repositories (e.g., `eurofab-project/morphometrics`, `eurofab-project/eo`) for all relevant code, tools, and documentation. The scope of the open-source work can also be extended to establish and maintain online forums, mailing lists, and developer channels (e.g., Slack, Discord) to provide support to users and foster a collaborative environment for contributors.

5 Productionalising

Moving EuroFab from a research project to an operational service requires a well-defined framework covering deployment, service delivery, quality assurance, and iterative development based on user needs.

- **Defining Output Products:** Based on the stakeholder work there are two types of products stakeholders are interested in. The first type are baseline Products such as the pan-European urban fabric maps and the underlying data based on the core classification scheme. These products should be openly accessible and updated regularly (e.g., annually or biennially). The second type are derived indicators from the base urban fabric data. Examples include metrics related to urban sprawl patterns, structure of cities, countries and regions, as well as access to the underlying morphological data (if available).
- **Access and Integration:** The resulting pan-European datasets will require specialised hosting dissemination solutions, due to their volume. The successful development and deployment of the morphological web application during the pilot phase demonstrated the viability of leveraging national research infrastructures, such as the Czech national research infrastructure network, for hosting and serving such data products. These infrastructures often provide scalable cloud storage solutions and are well-suited for handling large geospatial datasets. To maximise interoperability and efficiency, the pan-European datasets should be stored and disseminated in cloud-native, analysis-ready formats, such as Cloud-Optimised GeoTIFFs (COGs) for raster data from the AI vision model and GeoParquet for vector-based outputs from the morphometric models. Furthermore, to maximise impact and avoid creating isolated systems, EuroFab services can be integrated with or exposed through relevant existing European platforms and cloud data standards.
- **User Adoption and Trust:** Continuous stakeholder and end-user engagement has been a cornerstone of the pilot study and will remain critical for the operational phase. The pan-European classification will actively seek feedback regarding data quality, utility, and potential errors, gathered from stakeholders, will be systematically addressed and, where appropriate, implemented through transparent and publicly accessible Calibration/Validation (Cal/Val) procedures. The further investment in stakeholder engagement, co-design processes training, and clear documentation will also improve user confidence in the data product and drive adoption.

- **Timeline:** The indicative timeline for achieving full operationalisation of the pan-European EuroFab service is projected to be two years. This work carried out during this time can cover increasing the scale the methodology, building up the additional services based on stakeholder requirements, operationalise the results and incorporating the user feedback mechanisms.
- **Team Composition:** The proposed core team composition mirrors that of the successful pilot study, ensuring continuity of expertise and established collaborative workflows. This comprises a dedicated six-person team, with two expert members drawn from each of the three key institutions already involved in the EuroFab consortium. The only difference with the current structure is the change of the affiliation of Turing Team, which will be under the University of Liverpool in the UK. The division of responsibilities will remain : the OECD team will continue to lead stakeholder engagement, outreach, and policy relevance activities; the Charles University team will work on the scaling and refinement of the morphometric classification pipeline, as well as the overall product components; and the University of Liverpool team will focus on advancing the AI vision model and the space-time analytical components.
- **Funding:** The project team will actively pursue a diversified funding strategy. This will involve preparing and submitting competitive grant applications to relevant international and national funding bodies, such as UK Research and Innovation (UKRI), Horizon Europe programmes (e.g., those focused on Earth observation, digital transformation, or sustainable development), and the European Space Agency (ESA).