



**POLITECNICO
MILANO 1863**

**SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE**

EXECUTIVE SUMMARY OF THE THESIS

From Multimedia Data Analysis to the creation of a Context-aware Framework for the New European Bauhaus

LAUREA MAGISTRALE IN COMPUTER ENGINEERING - INGEGNERIA INFORMATICA

Author: TOMMASO ANDALORO

Advisor: PROF. LETIZIA TANCA

Co-advisor: EMILIA LENZI, CARLO ANDREA BIRAGHI

Academic year: 2025-2026

1. Introduction

In recent years, the increasing availability of heterogeneous urban data has profoundly transformed the way cities can be analysed and understood. Information coming from structured, semi-structured and unstructured sources such as geographic databases, multimedia content and public repositories provides complementary perspectives on the urban environment; however, extracting meaningful knowledge from these datasets requires the integration of different disciplines and methodologies. In this context, Computer Vision enables the automatic extraction of information from visual data, while data analysis techniques support the integration and interpretation of heterogeneous sources. Their combination provides new opportunities for Urban Analysis and Design, enabling the development of data-driven frameworks capable of supporting more objective, scalable and context-aware analyses. In particular, this thesis aims at implementing some indicators taken from the New European Bauhaus initiative [1], which are computed exploiting heterogeneous datasets. The latter are made of both structured, semi-structured and unstructured data which, if properly processed, can deliver insightful information

which can support urban analysis and design. Leveraging specific tools, such as semantic interpreters, libraries supporting machine learning, and other Python-based systems, those indicators will be implemented and they will finally contribute to the creation of a context-aware framework, which supports specific urban analysis issues. Every step of the project will be supported by two pillar documents: the New European Bauhaus self-assessment [2] tool and the CrAFt Impact Model [3]. The overall pipeline of the proposed methodology is summarised in Figure 1. Starting from the identified research gaps and the underlying theoretical background, the framework follows two complementary development paths: the automatic computation of existing New European Bauhaus indicators through Computer Vision techniques, and the implementation of a new accessibility indicator by integrating heterogeneous datasets. The resulting indicators are then enriched with additional contextual information and organised within a Context Dimension Tree (CDT), ultimately enabling context-aware urban analysis and supporting data-driven decision-making in urban design.

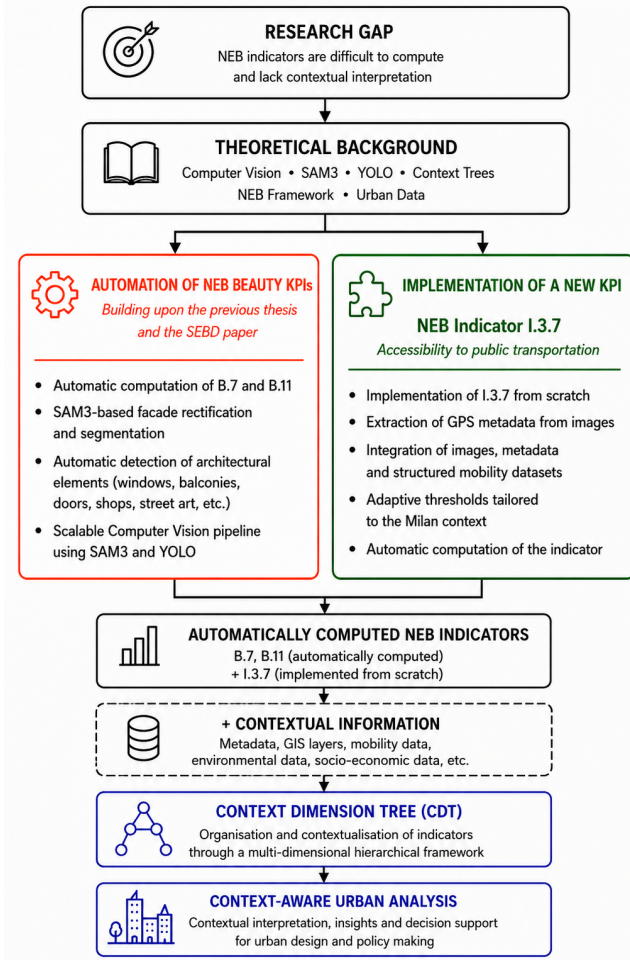


Figure 1: Overview of the thesis' work.

2. Background and State of the Art

The increasing availability of heterogeneous urban data has profoundly changed the way cities can be analysed and interpreted. Geographic information systems, multimedia datasets and public repositories continuously generate large amounts of information describing different aspects of the built environment. As a consequence, urban analysis is progressively shifting from traditional approaches towards data-driven methodologies, where heterogeneous sources are integrated to provide a more comprehensive representation of urban contexts.

Within this scenario, the New European Bauhaus (NEB) represents one of the most relevant European initiatives for promoting urban environments that are sustainable, inclusive and beautiful. To translate these principles into measurable objectives, the European Commission introduced a collection of Key Performance

Indicators (KPIs), whose computation often requires the integration of multiple data sources. Among the initiatives supporting the practical implementation of the NEB, the CrAFt Impact Model provides a framework for evaluating urban interventions through context-sensitive indicators.

Recent advances in Computer Vision have further expanded the possibilities for urban analysis. Deep learning models such as YOLO and SAM3 make it possible to automatically detect and segment urban elements from street-level imagery, transforming unstructured visual information into structured data that can support the computation of urban indicators.

However, extracting information is only the first step of the analysis. Urban elements need to be interpreted within their surrounding context in order to describe the complexity of the built environment. For this reason, this thesis adopts the Context Dimension Tree (CDT), a hierarchical graph-based model that organises heterogeneous information and preserves the relationships among different urban entities. This representation provides the conceptual basis for the context-aware framework proposed in this work.

3. Computation of the NEB Indicators

The implementation of the New European Bauhaus indicators constitutes the operational core of the proposed framework: indeed, the Context Dimension Tree presented in the following chapter relies on the information extracted through the computation of the KPIs, which constitute the quantitative layer upon which the contextual representation is built. For this reason, this thesis focuses on two complementary contributions: on the one hand, the automatic computation of the existing Beauty indicators, whose original implementation required several manual operations which limited the scalability of the overall process; on the other hand, the implementation from scratch of the Accessibility to Public Transportation indicator, whose computation demonstrates how heterogeneous data sources can be integrated within the same analytical pipeline. Although the two implementations rely on different methodologies, they share the same objective, transforming heterogeneous information into quantitative indicators which

can subsequently be contextualised inside the proposed framework.

3.1. Computation of the Beauty Indicators

The implementation of the Beauty indicators starts from the pipeline developed in a previous thesis, called "*Multimedia data analysis: A pipeline for the New European Bauhaus KPIs*" [4], which has been subsequently consolidated in the discussion paper presented at SEBD 2026 [5]; however, the overall workflow underwent substantial modifications, whose main objective is the automation of almost every manual operation previously required during the computation of the KPIs. In particular, the original implementation relied on manually prepared datasets and manually annotated architectural elements, making the overall process accurate but particularly time-consuming and difficult to scale when analysing large collections of images. The proposed pipeline starts from street-level photographs of building facades and automatically prepares the dataset before the actual computation of the indicators takes place. First of all, the main facade is extracted through SAM3 semantic segmentation; the resulting mask is then refined and used to estimate the façade boundaries, allowing the original perspective to be corrected through a rectangularisation process. This preliminary step provides a standardised representation of the facade, reducing the influence of the acquisition point of view and producing a suitable input for the following stages of the pipeline. Once the facade has been normalised, the architectural elements required for the computation of the indicators are automatically detected. The semantic capabilities of SAM3 are initially exploited to identify the main facade components; afterwards, a YOLO model was trained to recognise the architectural elements required by the implementation, such as windows, balconies, entrances, shops and graffiti. This second step considerably reduces the computational cost of the detection process, while preserving a satisfactory level of accuracy; moreover, the resulting bounding boxes can be directly exploited to compute the different components of the Beauty indicators, replacing the manual annotation process adopted in the previous implementation. The

obtained results demonstrate that the proposed methodology is capable of reliably detecting the architectural elements composing the facades, thus making the computation of the Beauty indicators considerably more scalable and user-friendly. Figure 2 reports an example of the generated bounding boxes together with the corresponding confusion matrix in Figure 3 obtained during the validation process, showing the overall performances achieved by the trained model.



Figure 2: Generated bounding-boxes.



Figure 3: Final confusion matrix.

3.2. Computation of the Accessibility Indicator

Unlike the previous indicators, whose computation mainly relies on the visual analysis of a

multimedia dataset, the implementation of the Accessibility to Public Transportation indicator aims at demonstrating how heterogeneous sources of information can be integrated to compute a New European Bauhaus KPI. In this case, the image itself does not provide sufficient information to assess the accessibility of the analysed building; on the contrary, it represents the starting point of a broader pipeline where unstructured data, metadata and structured datasets progressively contribute to the computation of the final score. The process starts from the extraction of the geographical coordinates stored inside the metadata associated with each image; this information is subsequently integrated with structured datasets describing the public transportation network of Milan, containing both underground stations and surface transportation stops. Once the different data sources have been integrated, the system evaluates the accessibility of each building by analysing the surrounding transportation infrastructure and by estimating the contribution of the closest public transport services to the final indicator. Although the New European Bauhaus provides a set of predefined thresholds to evaluate the accessibility conditions, their direct adoption proved unsuitable for the urban context analysed throughout this work. For this reason, an adaptive parametrisation was introduced: instead of relying on fixed thresholds, the proposed methodology estimates the density of the transportation network surrounding the analysed building and automatically adjusts the parameters used during the computation of the indicator. This allows the KPI to better adapt to different urban contexts, making the overall methodology more flexible while preserving the rationale proposed by the New European Bauhaus framework. Finally, the resulting scores are integrated in a dataset, containing the information of all the buildings in Milan, producing a spatial representation of the accessibility conditions across the city. Figure 4 reports the resulting accessibility map of Milan, highlighting how the integration of metadata and structured datasets makes it possible to automatically compute a context-dependent New European Bauhaus indicator.

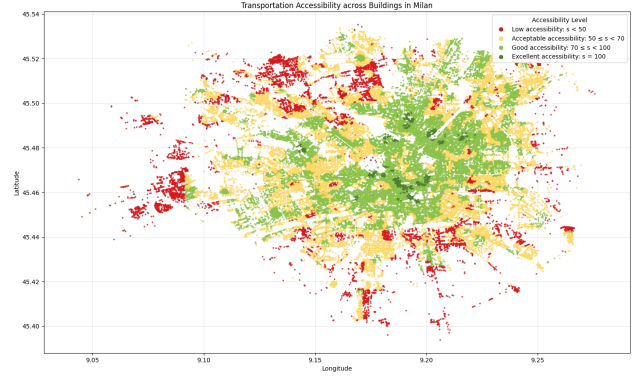


Figure 4: Map of Milan’s buildings, according to I.3.7.

4. The Context Dimension Tree for Urban Analysis

The automatic computation of the New European Bauhaus indicators represents only the first step of the proposed methodology: although the resulting KPIs provide quantitative information describing different aspects of the built environment, they are still isolated values which cannot fully represent the complexity of an urban context. In particular, a numerical indicator does not explicitly describe which information contributed to its computation, which datasets have been exploited nor how it relates to the other dimensions characterising the analysed scenario. For this reason, the final contribution of this thesis consists in the implementation of a Context Dimension Tree (CDT) [6]: a context-aware representation whose objective is to organise heterogeneous information according to the relationships existing among data, indicators and urban knowledge. The proposed framework combines the information provided by the CrAFt Impact Model and by the New European Bauhaus self-assessment tool, exploiting the former to identify the urban issue under analysis and the latter to retrieve the corresponding KPIs and the information required for their computation. Starting from a textual description of the problem provided by the user, semantic similarity techniques identify the most relevant impact categories; afterwards, the corresponding NEB indicators are automatically selected together with the datasets required for their implementation. This process makes it possible to generate a contextual representation which adapts to the analysed urban issue, in-

stead of relying on a predefined and static knowledge structure. The resulting Context Dimension Tree is organised around three dimensions: the *data type* dimension, describing the available structured, semi-structured and unstructured datasets; the *analysis* dimension, containing the New European Bauhaus indicators selected according to the analysed urban issue; and the *information* dimension, which represents the information required to compute those indicators. Once the indicators presented in the previous section have been computed, their numerical values are automatically associated with the corresponding nodes of the tree; similarly, the available datasets and the extracted information are connected to the appropriate branches, producing a contextual representation where every component contributes to the overall analysis. Figure 5 shows the complete implemented pipeline, which describes the overall process from the query input to the filtered contextual view.

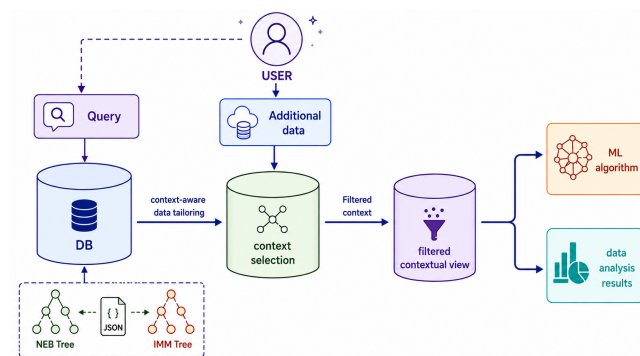


Figure 5: Pipeline for context-aware data selection

Figure 6 reports the final Context Dimension Tree generated throughout this work. The highlighted nodes identify the information effectively available inside the proposed framework: the multimedia dataset is represented through the unstructured image branch, the geographical coordinates extracted from the metadata support the computation of the Accessibility indicator by connecting buildings with the surrounding transportation infrastructure and finally, the automatically computed NEB indicators populate the analysis dimension, becoming part of a unified context-aware representation. In this way, the proposed methodology moves beyond the isolated computation of individual KPIs and

provides a flexible framework capable of integrating heterogeneous information inside a single analytical model, supporting more comprehensive and informed urban analyses through contextual representation.

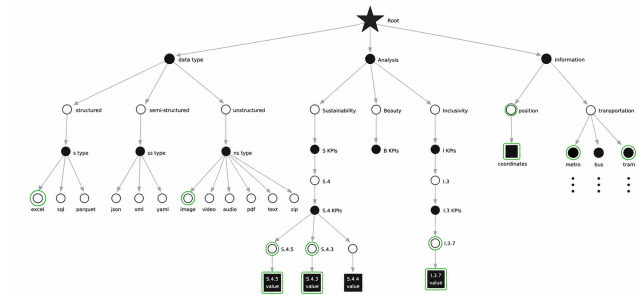


Figure 6: Final Context Dimension Tree generated through the proposed framework.

5. Conclusion and Future Developments

This thesis proposed a context-aware framework for Urban Analysis based on the integration of heterogeneous data sources, combining Computer Vision techniques, structured datasets and semantic models to support the implementation of some New European Bauhaus indicators. The automatic computation of the Beauty indicators considerably simplifies the original workflow, reducing the amount of manual intervention required during the analysis; similarly, the implementation of the Accessibility to Public Transportation indicator demonstrates how metadata and structured datasets can be integrated to compute urban KPIs which cannot be directly inferred from visual data alone. Finally, the computed indicators are organised within a Context Dimension Tree, providing a unified representation where data, information and analysis are explicitly connected, thus supporting more comprehensive and context-aware urban analyses.

Although the proposed framework already demonstrates the effectiveness of the adopted methodology, several opportunities for further development remain. A first possible extension concerns the evolution of the CrAFt Impact Model from a hierarchical knowledge structure into a simulation environment, where the relationships among different impact categories could be exploited to estimate the effects gen-

erated by urban interventions and to support decision-making processes. A second development regards the automatic integration of user-provided datasets: instead of relying on manually defined associations, the framework could automatically analyse heterogeneous data sources, identify the available information and populate the Context Dimension Tree accordingly. Such extensions would further increase the flexibility of the proposed methodology, allowing the framework to dynamically adapt both to different urban issues and to the datasets available for their analysis.

References

- [1] European Commission. New european bauhaus, 2026.
- [2] Konstantinos Gkatzogiannis, Romano Romano, and Paola Negro, editors. *A Practical Guide to the New European Bauhaus Self-Assessment Method and Tool*. European Commission, Joint Research Centre, 2024.
- [3] Han Vandevyvere and Annemie Wyckmans. Neb impact model: Summary for cities. Project deliverable d1.1, CrAFt – Creating Actionable Futures, June 2023. Version 30.06.2023.
- [4] Davide Ciuffreda and Marco Ciammaichella. Multimedia data analysis: A pipeline for the new european bauhaus kpis. Master’s thesis, Politecnico di Milano, 2026. Master’s Degree in Computer Science and Engineering.
- [5] Tommaso Andalaro, Ilaria Bartolini, Carlo Andrea Biraghi, Marco Ciammaichella, Davide Ciuffreda, Emilia Lenzi, Marco Patella, and Letizia Tanca. A pipeline for multimedia data preparation in the context of the new european bauhaus. In *Proceedings of the 34th Symposium on Advanced Database Systems (SEBD 2026)*, Senigallia, Italy, June 2026. Discussion Paper.
- [6] Cristiana Bolchini, Carlo A. Curino, Giorgio Orsi, Elisa Quintarelli, Rosalba Rossato, Fabio A. Schreiber, and Letizia Tanca. And what can context do for data? *Communications of the ACM*, 52(11):136–140, 2009.