NUS National University of Singapore

urban analytics lab

(Geo)AI for Buildings and Cities

Filip Biljecki
Assistant Professor
National University of Singapore
Urban Analytics Lab

This mosaic was formed using 5,720 street view images contributed by Mapillary and KartaView users.

Content

Challenges and opportunities

- Understanding emerging data — we need further data sources
- Data gaps and data quality —we need better data
- Multi-scale integration — we need to integrate data across scales
- Overlooked human aspect — we need to bring humans in the loop

Goals

- Raise awareness of emerging forms of data and their quality
- Present several use cases in the context of climate change mitigation and adaptation
- Examples of human-centric AI
- Learn about research directions

urban analytics lab

buds lab building and urban data science

BEAM

NUS National University of Singapore

Street view imagery is one of the emerging datasets that supports research on climate change (e.g. urban form)

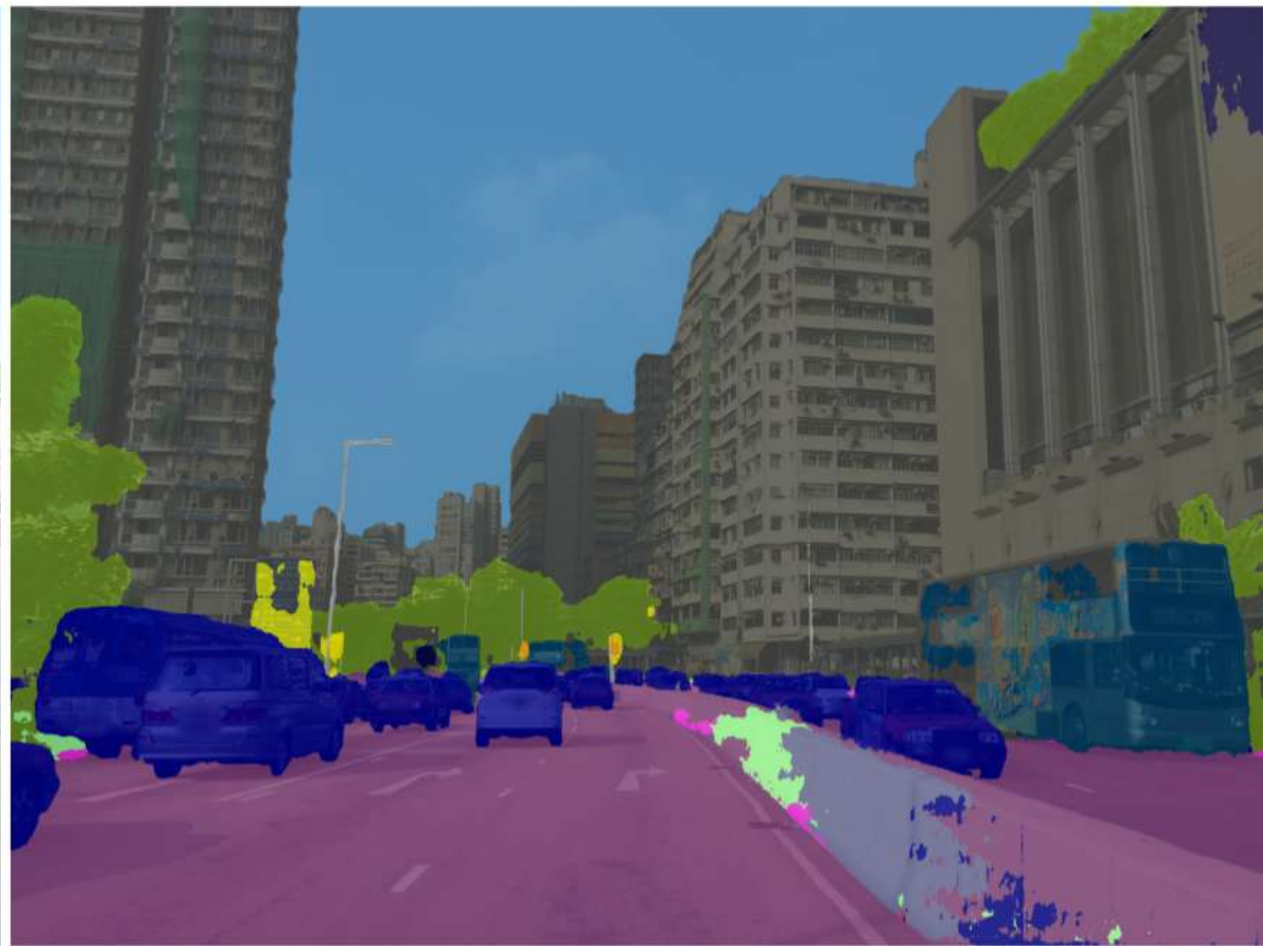
Obtained from Google Street View

Human perception
Appearance, audits, walkability, and socio-economic studies

Infrastructure
Spatial data collection, real estate, transport, and health studies

Activities
Human traffic, place semantics, vibrancy, and economic activity

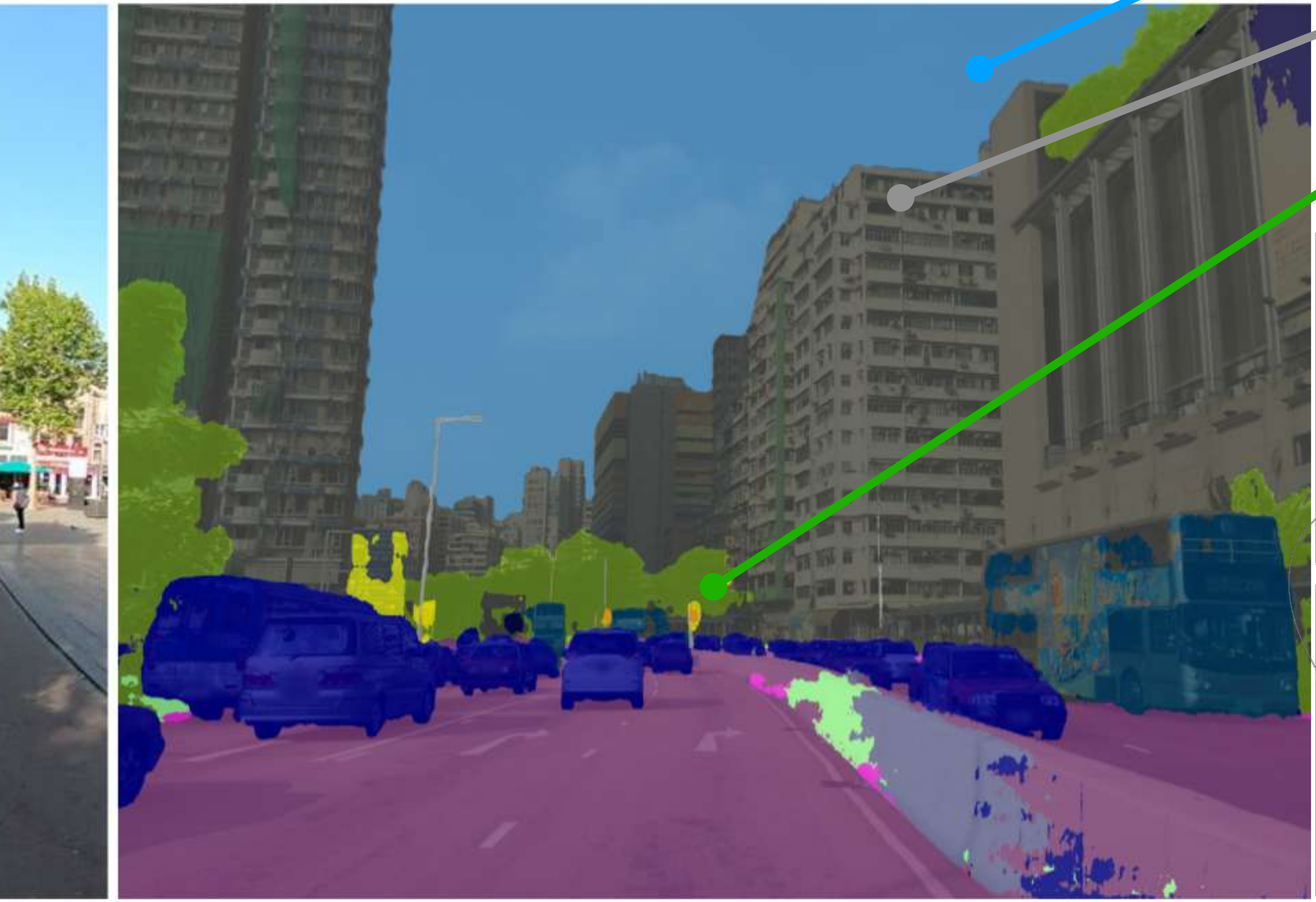
Urban form
Sky view factor and share of vegetation in urban canyons

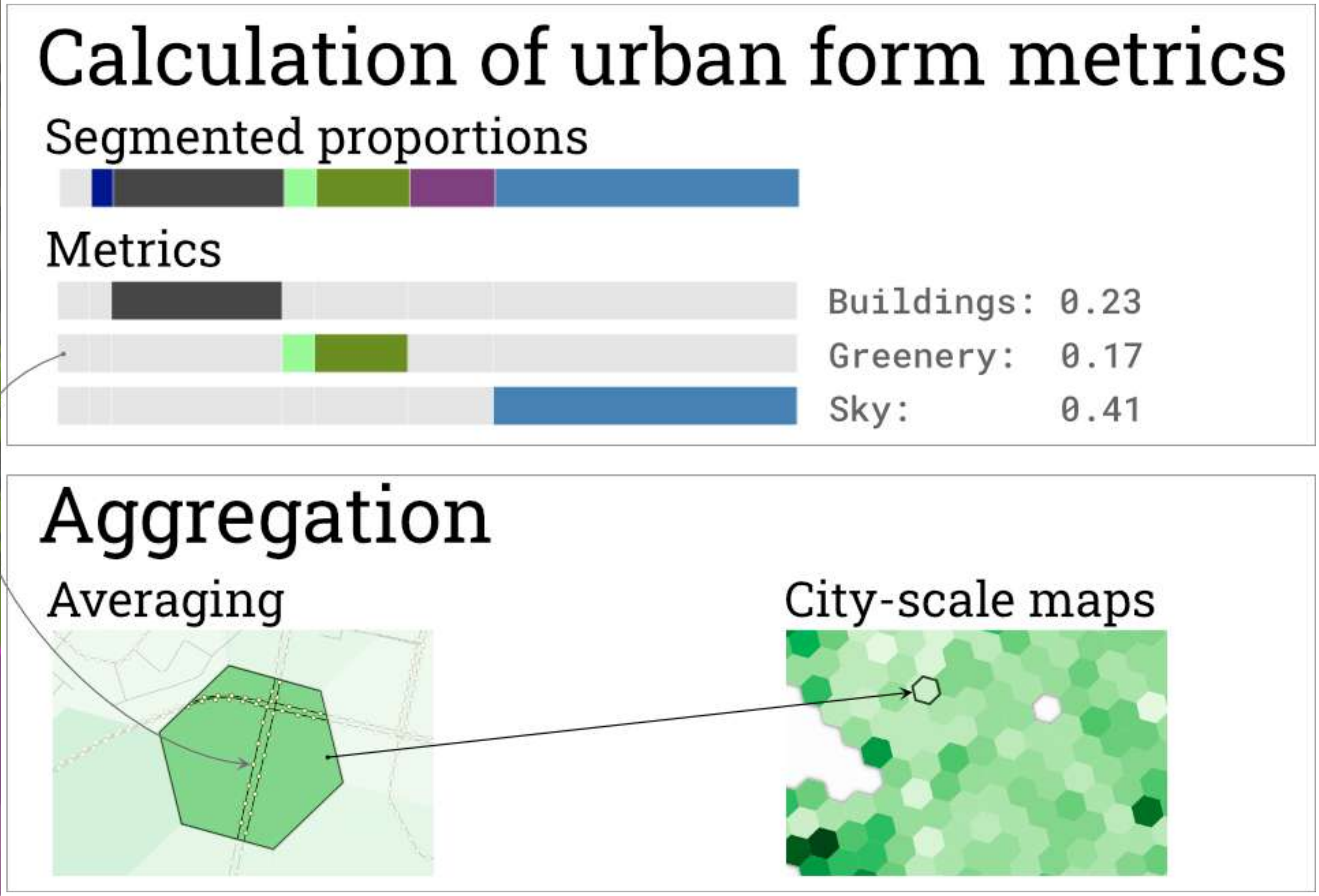


Images are obtained from Mapillary

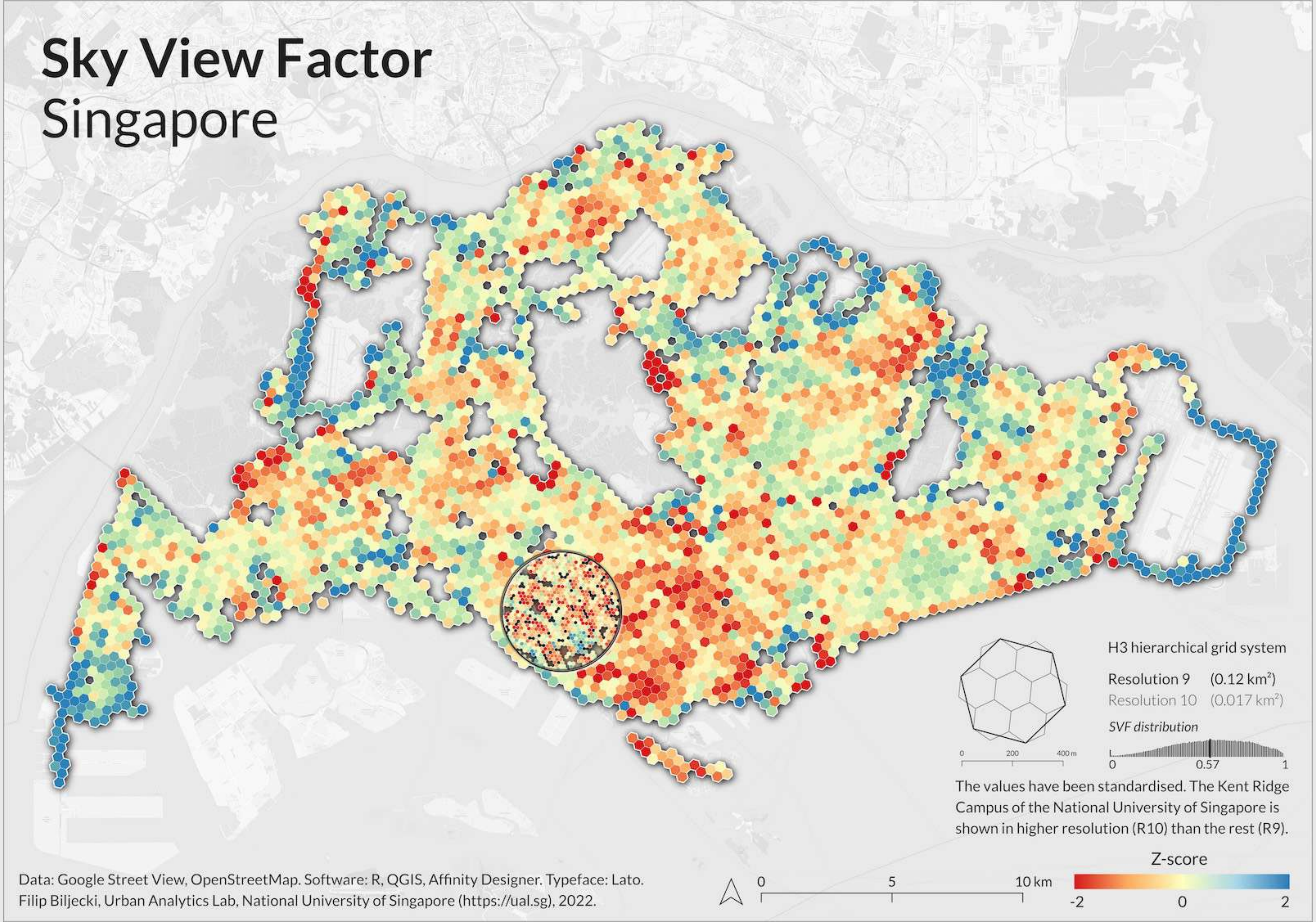
Urban form

Sky view factor and share of vegetation in urban canyons





Images are obtained from Mapillary following a manual exploration


Sky View Factor
Singapore
H3 hierarchical grid system
Resolution 9 (0.12 km²)
Resolution 10 (0.017 km²)
SVF distribution
0 0.57 1
0 200 400 m
The values have been standardised. The Kent Ridge Campus of the National University of Singapore is shown in higher resolution (R10) than the rest (R9).
Z-score
-2 0 2
0 5 10 km
Data: Google Street View, OpenStreetMap. Software: R, QGIS, Affinity Designer. Typeface: Lato.
Filip Biljecki, Urban Analytics Lab, National University of Singapore (https://ual.sg), 2022.

User-generated geographic information such as OpenStreetMap has also surged in the past years…

Search Where is this?

More than 600 million buildings mapped in OSM

… supporting research on urban form and other relevant applications at the city and building scales

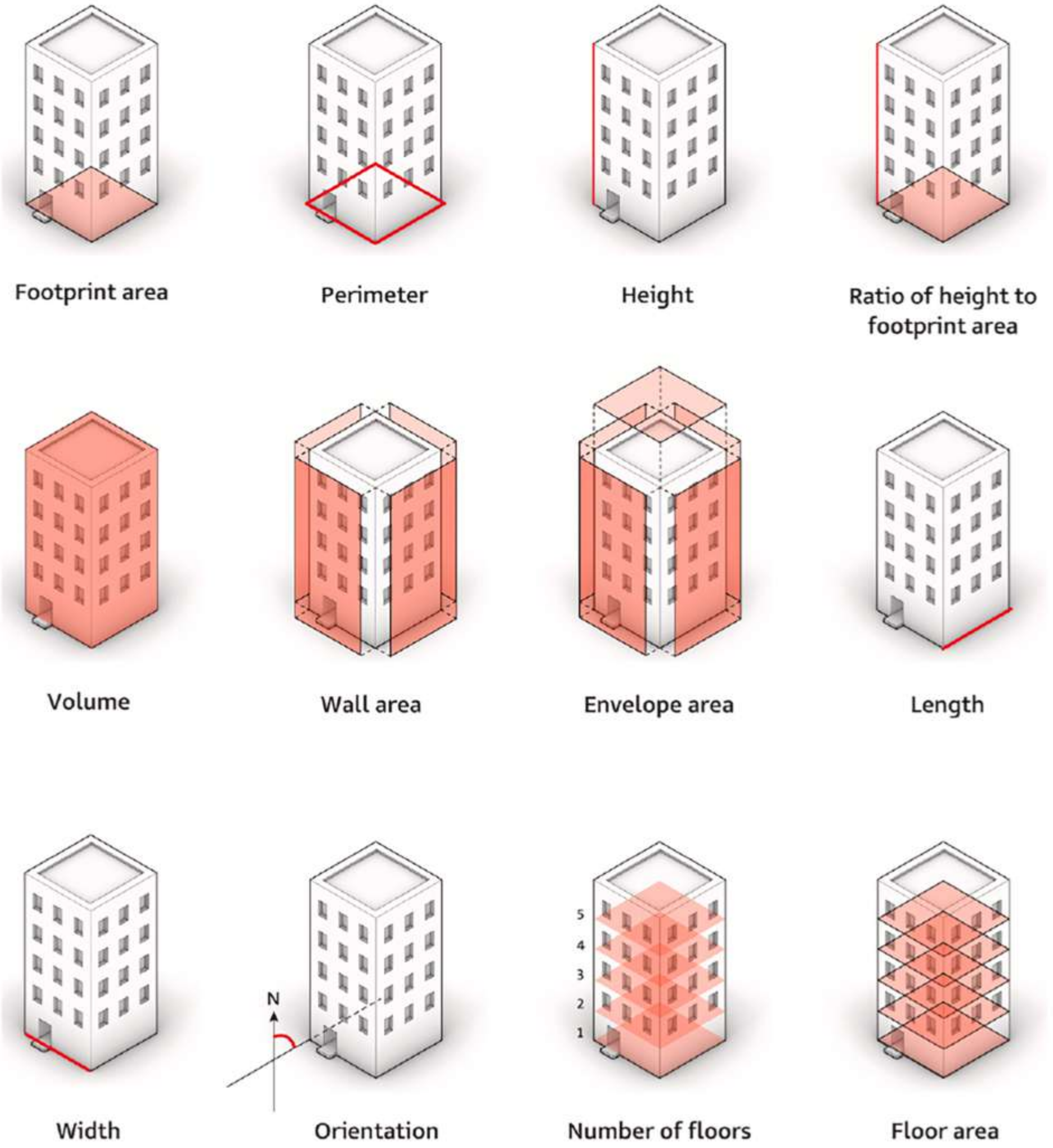
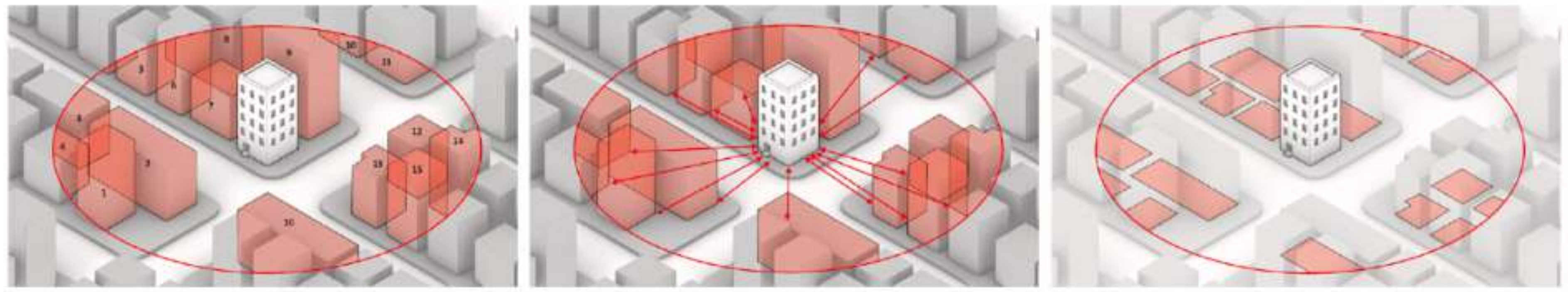
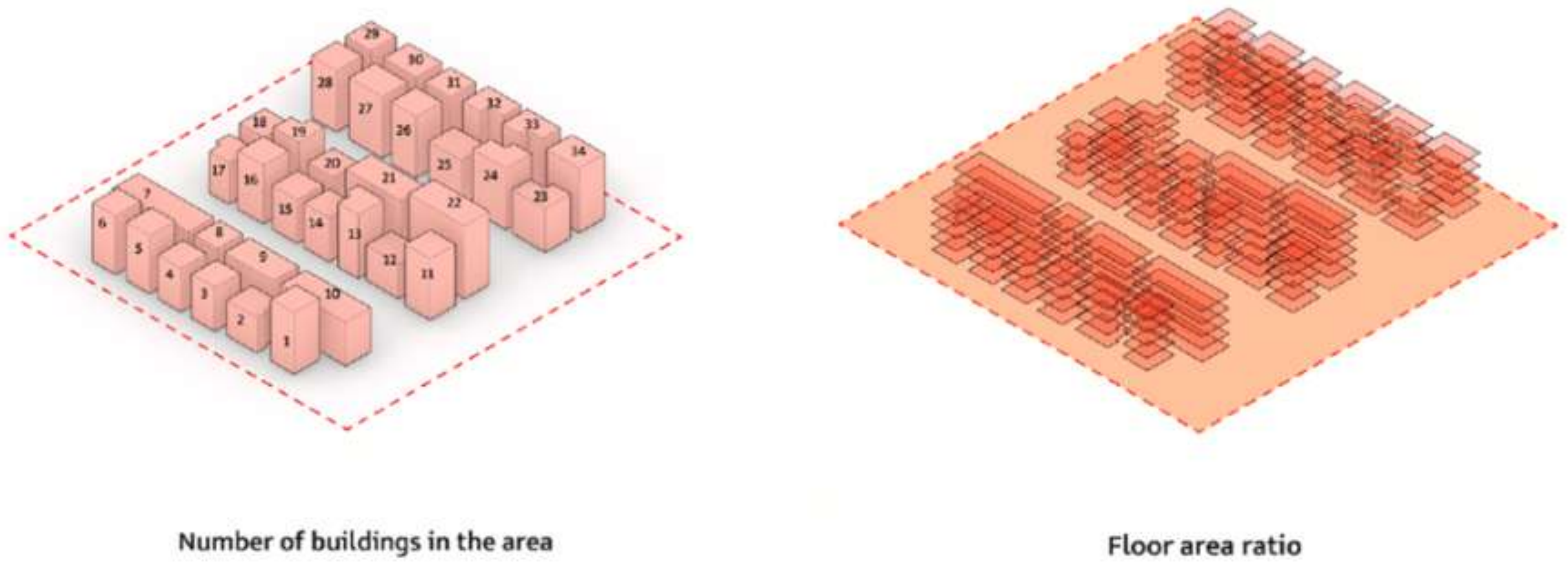


Fig. 1. Illustration of independent indicators at the building level.

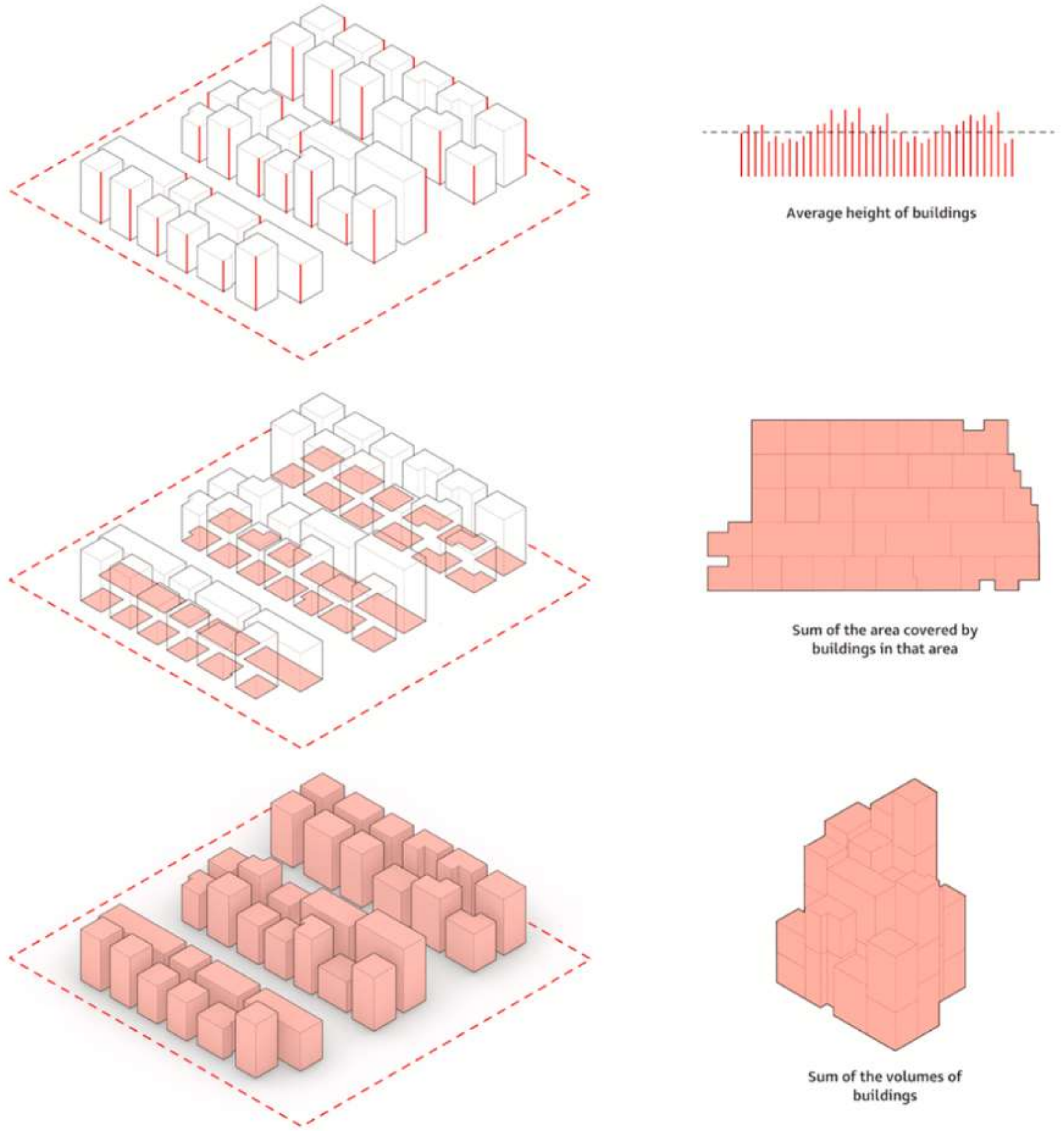


(a) Count. (b) Distances. (c) Area.

Fig. 2. Illustration of contextual aspects computed at the building level based on its surroundings.



(a) Aggregated indicators that are calculated directly based on one or more indicators at the building-level.

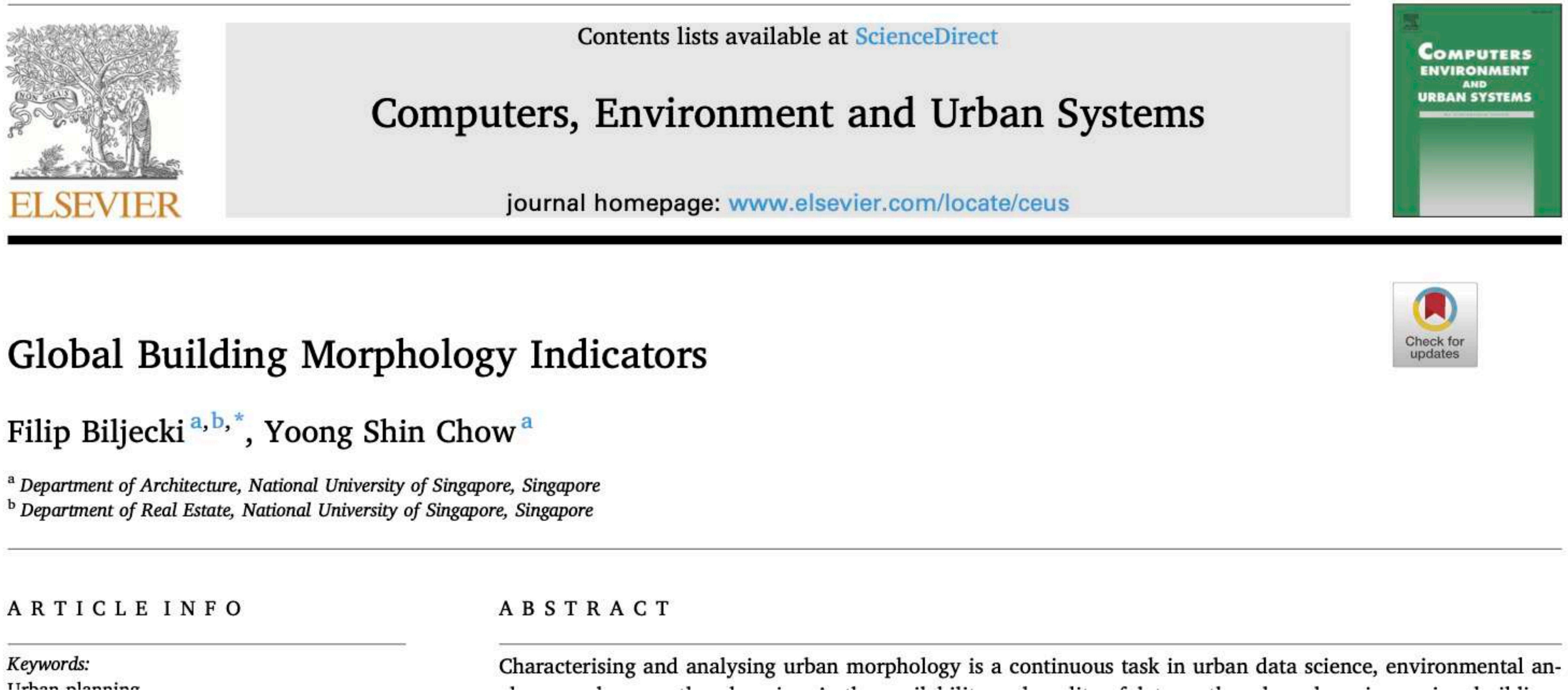


(b) Examples of aggregated indicators that are based on summary statistics from an array of values such as building heights. Each of these indicators has several counterparts pertaining to the same array of values, such as minimum value and standard deviation.

GBMI

Global Building Morphology Indicators

- 1. Catalogue of metrics (~600)
  - e.g. site coverage,
    st. dev. of footprint complexity
- 2. Open-source tool to compute them
  - At any scale and anywhere
  - From OpenStreetMap
- 3. Open dataset
  - 20+ cities and a few countries


Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

ELSEVIER

Global Building Morphology Indicators

Filip Biljecki [a,b,*], Yoong Shin Chow [a]

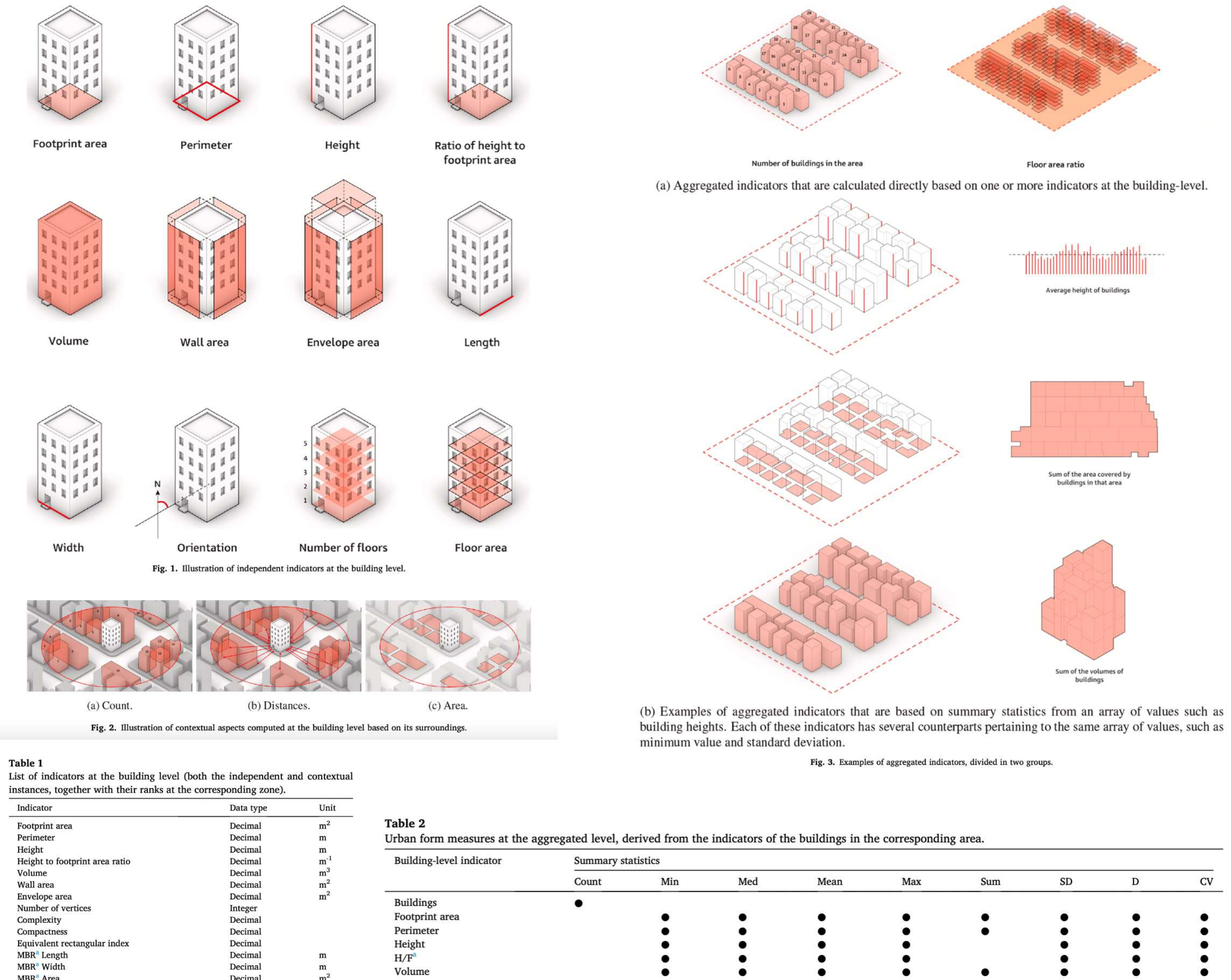
[a] Department of Architecture, National University of Singapore, Singapore
[b] Department of Real Estate, National University of Singapore, Singapore

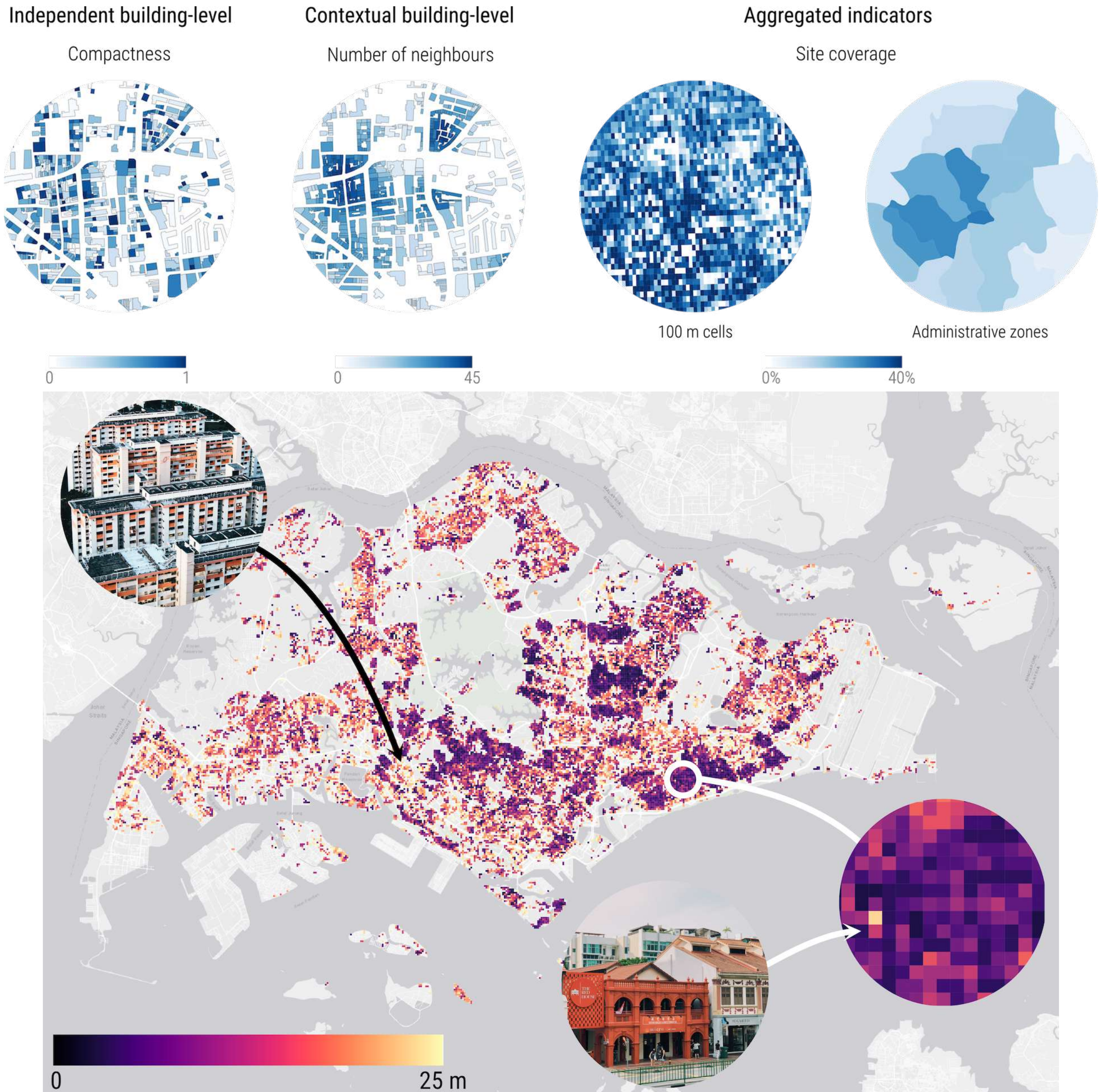
ARTICLE INFO

Keywords:

ABSTRACT

Characterising and analysing urban morphology is a continuous task in urban data science, environmental an-




Independent building-level
Compactness
0
1
Contextual building-level
Number of neighbours
0
45
Aggregated indicators
Site coverage
100 m cells
Administrative zones
0%
40%
0
25 m

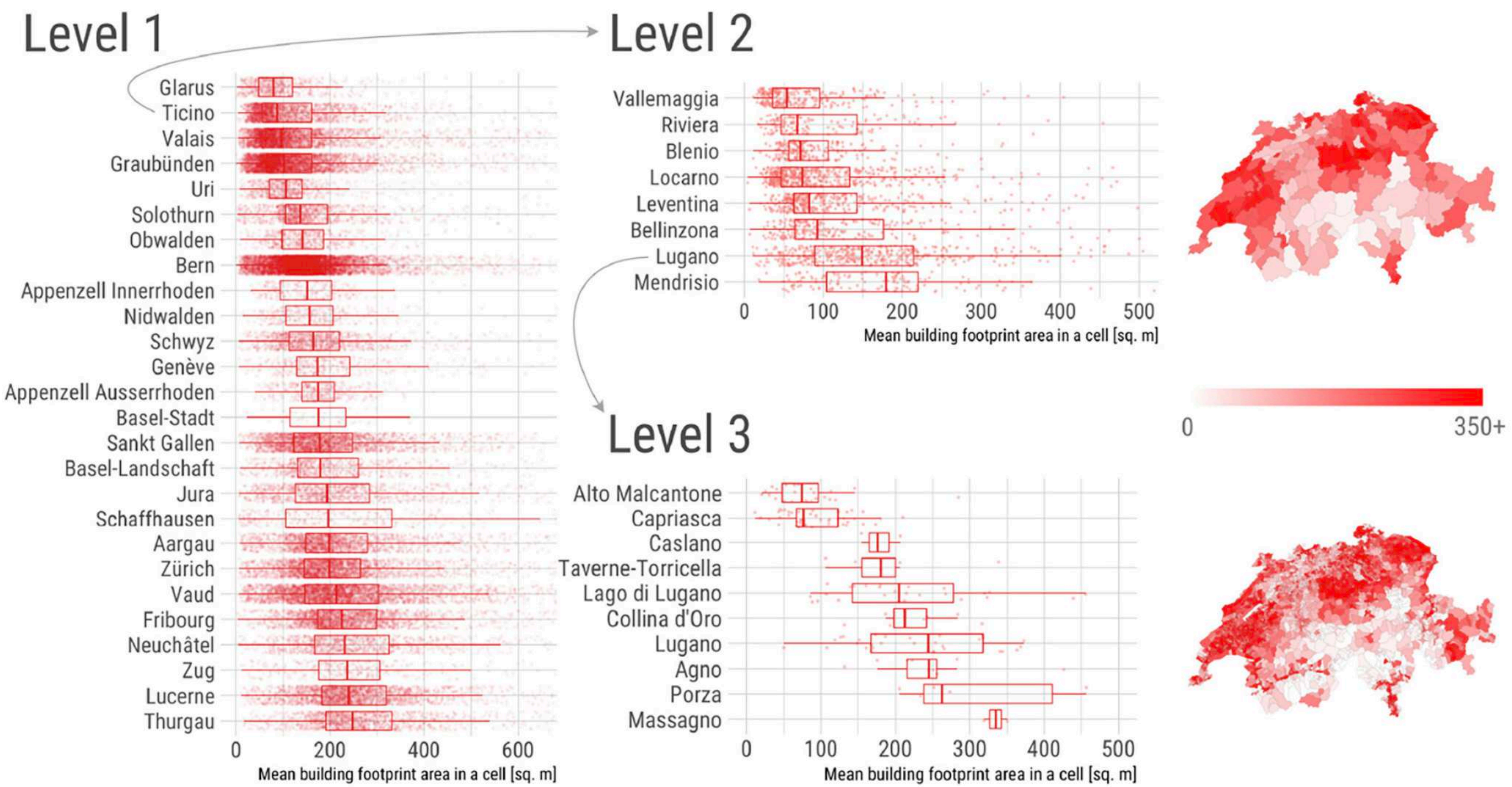


Fig. 7. Hierarchical and structured integration of data. These plots and maps were derived from footprint areas of all buildings in Switzerland, and aggregated at multiple levels. Note that for space constraints the plot does not indicate all levels, and in the lowest level we omitted some zones.

OpenStreetMap
Edit | History | Export
GPS Traces | User Diaries | Communities | Copyright | Help | About | Log In | Sign Up

© OpenStreetMap contributors ♥ Make a Donation. Website and API terms

What about the quality of this data?

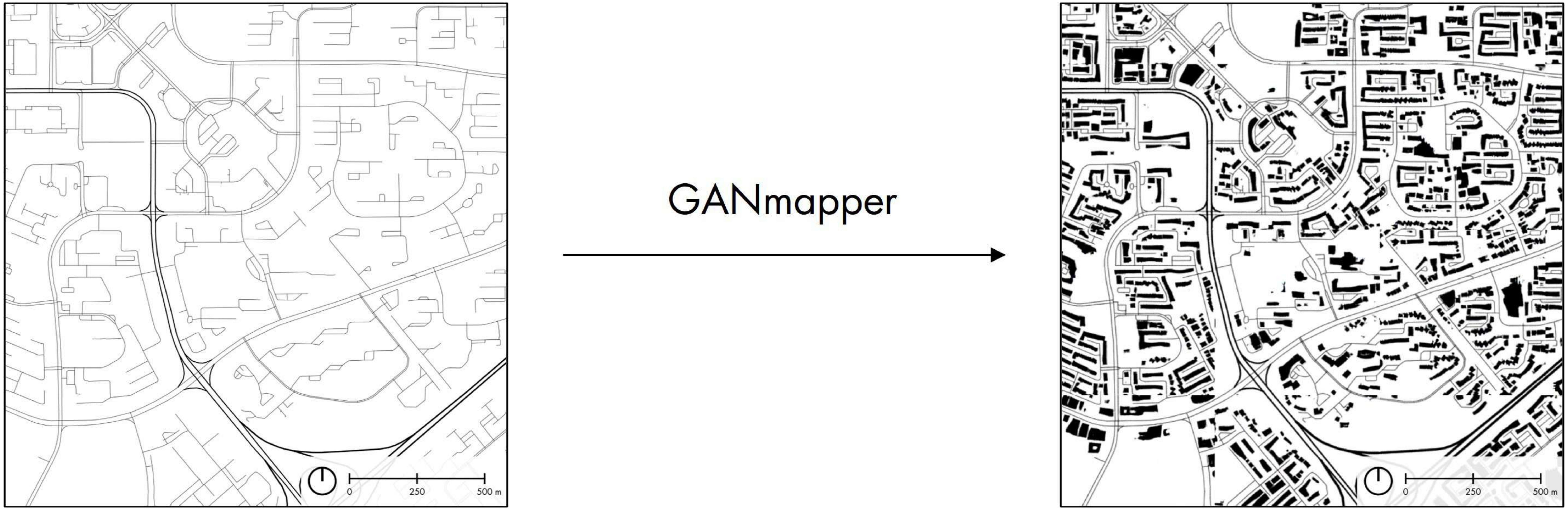

Jalan Sambon - Kartasura
15
Jalan Pandawa
IAIN Surakarta
Jalan Slamet Riyadi
Jalan Brigadir Jenderal Katamso
Jalan Jenderal Ahmad Yani
Kali Gembongan
Pabelan
Gembongan
Jalan Raya Gembongan
Kartasura
Jalan Sanggrahan
Jalan Gumuksari I
Mess Grup
2 Kopassus
Jalan Jegongan
Jalan Tanggul
Ngadirejo
Jalan Jatiurip - Ngemplak
KartasuraKartas,
NgadirejoNgadireGempol,
Gumpang
Jalan Ngemplak - Papungan
Jalan Mayang Gatak
Jalan Raya Gawok
(c) OpenStreetMap contributors

Can we use AI to improve data on buildings?

GANmapper — geographical data translation

by Abraham Noah Wu

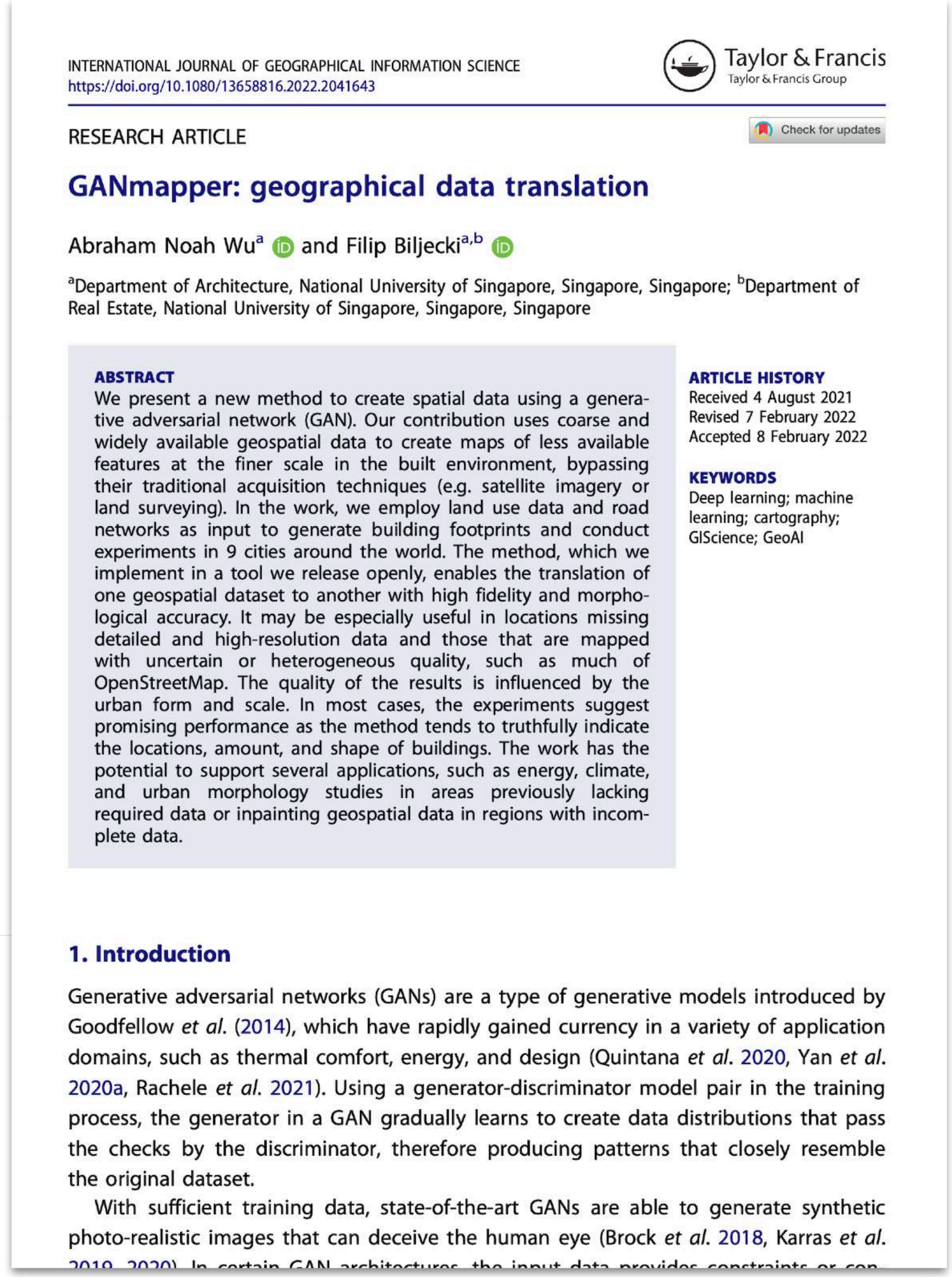
New means of building data generation using Generative Adversarial Network



Published in the International Journal of Geographical Information Science. Credit: Abraham Noah Wu

GANmapper — geographical data translation

New means of building data generation using Generative Adversarial Network



INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
https://doi.org/10.1080/13658816.2022.2041643

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Check for updates

GANmapper: geographical data translation

Abraham Noah Wu[a] and Filip Biljecki[a,b]

[a]Department of Architecture, National University of Singapore, Singapore, Singapore; [b]Department of Real Estate, National University of Singapore, Singapore, Singapore

ABSTRACT

We present a new method to create spatial data using a generative adversarial network (GAN). Our contribution uses coarse and widely available geospatial data to create maps of less available features at the finer scale in the built environment, bypassing their traditional acquisition techniques (e.g. satellite imagery or land surveying). In the work, we employ land use data and road networks as input to generate building footprints and conduct experiments in 9 cities around the world. The method, which we implement in a tool we release openly, enables the translation of one geospatial dataset to another with high fidelity and morphological accuracy. It may be especially useful in locations missing detailed and high-resolution data and those that are mapped with uncertain or heterogeneous quality, such as much of OpenStreetMap. The quality of the results is influenced by the urban form and scale. In most cases, the experiments suggest promising performance as the method tends to truthfully indicate the locations, amount, and shape of buildings. The work has the potential to support several applications, such as energy, climate, and urban morphology studies in areas previously lacking required data or inpainting geospatial data in regions with incomplete data.

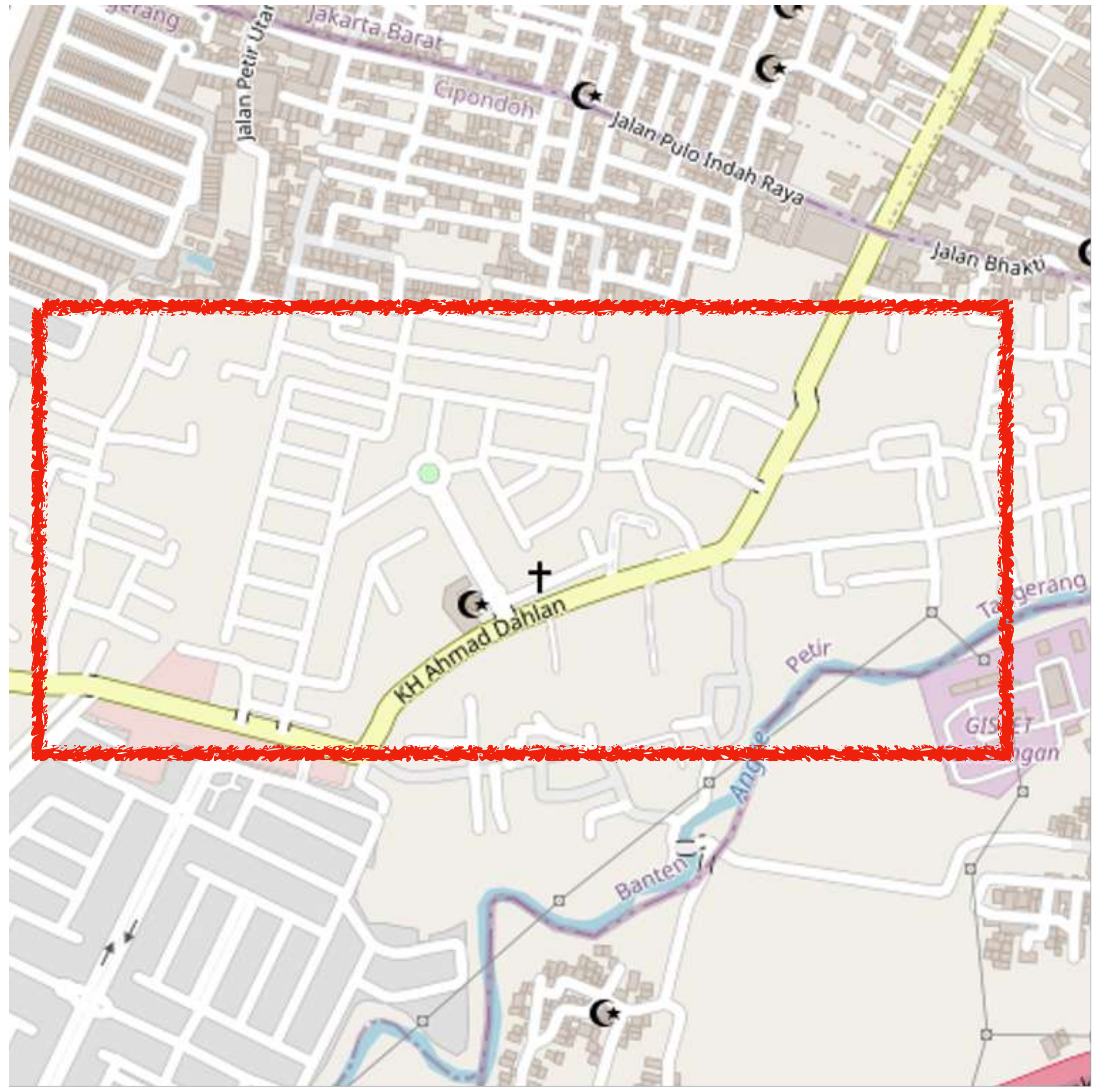
ARTICLE HISTORY
Received 4 August 2021
Revised 7 February 2022
Accepted 8 February 2022

KEYWORDS
Deep learning; machine learning; cartography; GIScience; GeoAI

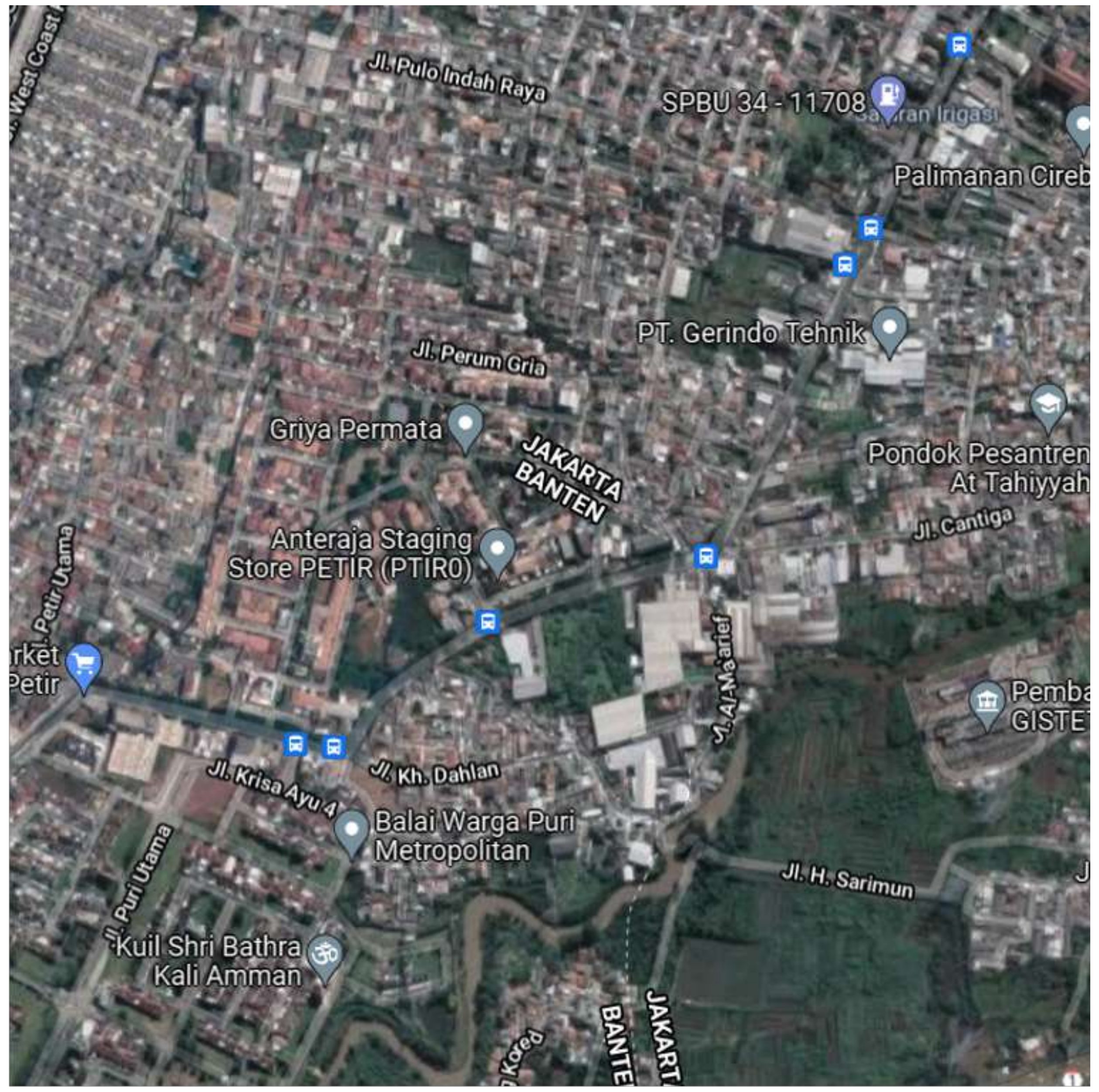
1. Introduction

Generative adversarial networks (GANs) are a type of generative models introduced by Goodfellow *et al.* (2014), which have rapidly gained currency in a variety of application domains, such as thermal comfort, energy, and design (Quintana *et al.* 2020, Yan *et al.* 2020a, Rachele *et al.* 2021). Using a generator-discriminator model pair in the training process, the generator in a GAN gradually learns to create data distributions that pass the checks by the discriminator, therefore producing patterns that closely resemble the original dataset.

With sufficient training data, state-of-the-art GANs are able to generate synthetic photo-realistic images that can deceive the human eye (Brock *et al.* 2018, Karras *et al.* 2019, 2020). In certain GAN architectures, the input data provides constraints or con-



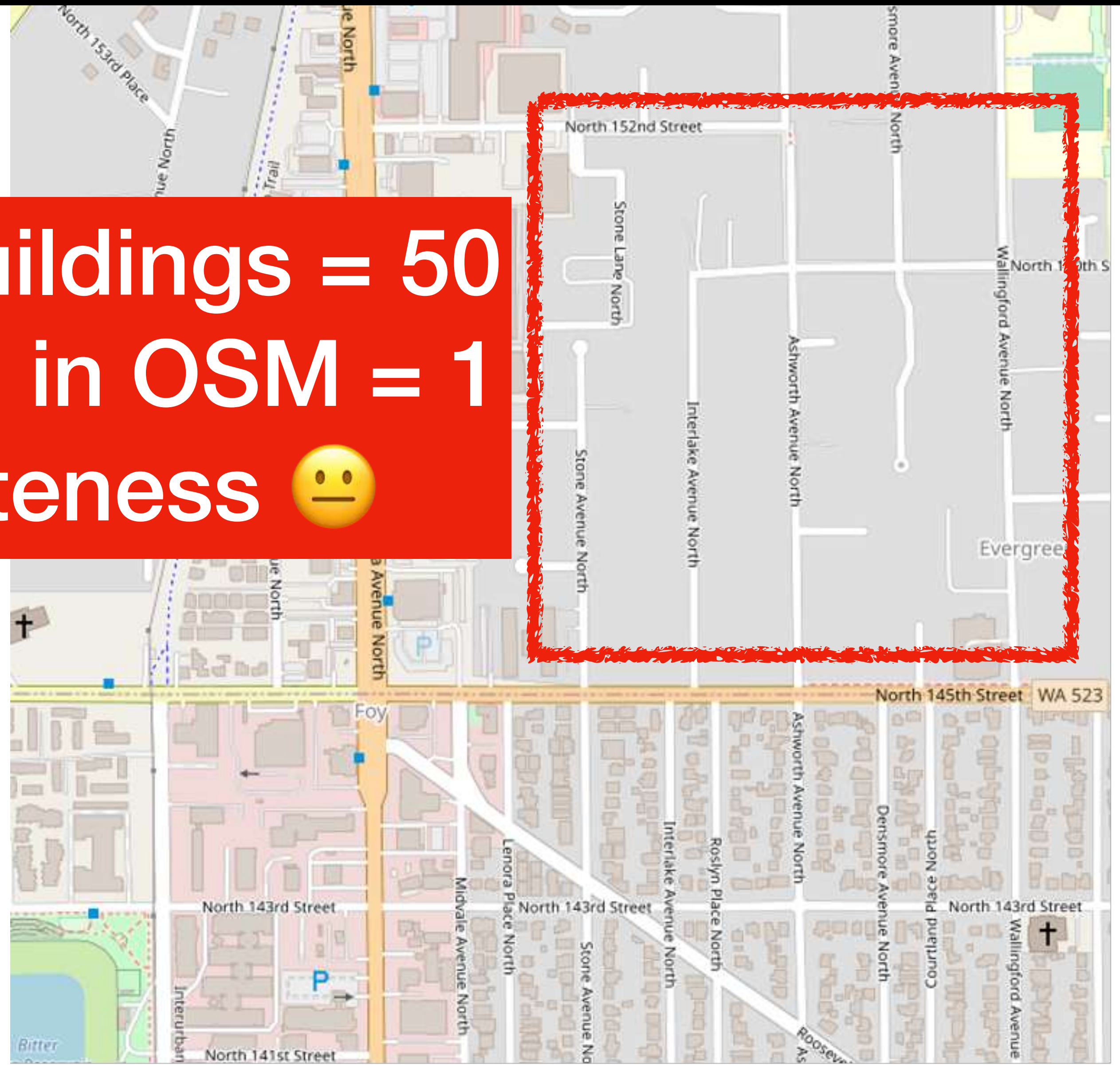
Jakarta (OSM)



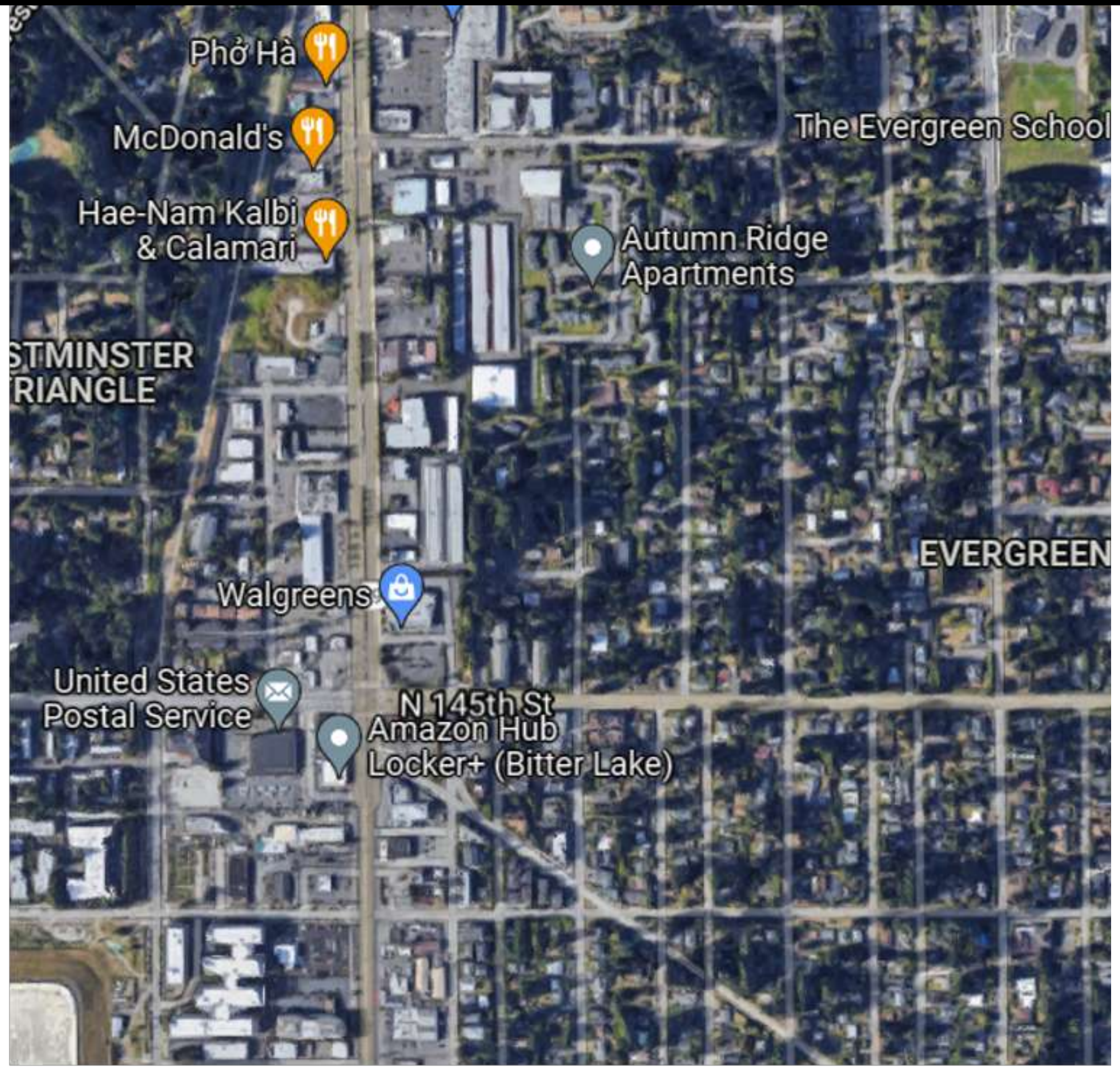
Jakarta (Satellite)

GANmapper for spatial data quality control (completeness)?

Predicted no. of buildings = 50
No. of buildings mapped in OSM = 1
🔍 Poor OSM completeness 😐



Seattle (OSM)



Seattle (Satellite)

ISPRS Journal of Photogrammetry and Remote Sensing 195 (2023) 90–104

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

InstantCITY: Synthesising morphologically accurate geospatial data for urban form analysis, transfer, and quality control

Abraham Noah Wu [a], Filip Biljecki [a,b,*]

[a] *Department of Architecture, National University of Singapore, Singapore*
[b] *Department of Real Estate, National University of Singapore, Singapore*

ARTICLE INFO

Keywords:
Deep learning
Machine learning
Volunteered Geographic Information
GIScience
GeoAI

ABSTRACT

Generative Adversarial Network (GAN) is widely used in many generative problems, including in spatial information sciences and urban systems. The data generated by GANs can achieve high quality to augment downstream training or to complete missing entries in a dataset. GANs can also be used to learn the relationship between two datasets and translate one into another, e.g. road network data into building footprint data. However, such approach has not been developed in the geospatial and urban data science context, its usability remains unknown, and the methods are not fully developed. We develop a new Geographical Data Translation algorithm based on GAN to generate high-resolution vector building data solely from street networks, which may be used to predict the urban morphology in absence of building data, also enabling studies in unmapped or undermapped urban geographies, among other advantages. Experiments on 16 cities around the world demonstrate that the generated datasets are largely successful in resembling ground truth morphologies. Thus, the approach may be used in lieu of traditional data for tasks that are often hampered by lack of data, e.g. urban form studies, simulation of urban morphologies in new contexts, and spatial data quality assessment. Our work proposes a novel rapid approach to generate building footprints in replacement of procedural methods and it introduces a new intrinsic method for large-scale spatial data quality control, which we test on OpenStreetMap by predicting missing buildings and suggesting the completeness of data without the usually required authoritative counterparts. The code, sample model, and dataset are available openly.

1. Introduction

The rapid growth of geospatial data in the past decades have accelerated our understanding of the world. Insights and products derived from geospatial data are solving a myriad of complex problems down the gamut from increasing fuel efficiency, monitoring logistic networks, to sustainability-driven urban designs (Lee and Kang, 2015; Lipson et al., 2022; Wagner et al., 2022; Arribas-Bel and Fleischmann, 2022; Wang et al., 2022; Li et al., 2022b). The datasets that enabled these advancements are mostly collected using a variety of methods such as surveying, remote sensing, mobile mapping, and crowdsourcing (Li et al., 2016; Huang and Wang, 2020; Jin et al., 2022; Heikinheimo et al., 2020; Luo et al., 2022; Yan and Huang, 2022).

Today, such datasets are more abundant than ever, but their quality and coverage can still vary dramatically. This phenomenon is described as Geospatial Data Asymmetry in which some geospatial data are mapped more extensively than another *correlated* data (Wu and Biljecki, 2022). This problem is more acute in open access and volunteered platforms such as OpenStreetMap (OSM). For example, more than 80% of roads has been mapped in OSM (Barrington-Leigh and Millard-Ball, 2017). However, building data, which is usually strongly associated with street networks, is estimated to be disproportionately less complete (Biljecki, 2020; Yeboah et al., 2021; Leonard et al., 2022), hampering their applications in many areas around the world.

Such asymmetry exists due to the fact that it is more complex and time-consuming to map certain features such as building footprints than tracing street networks. Automated approached used in research and industry also face similar limitations despite recent advances (Sun et al., 2020).

One novel take on the issue of data asymmetry is the introduction of Geographic Data Translation (GDT). Instead of simply collecting data from the real world, GDTs use the problem of data asymmetry to their advantage, generating a less abundant geospatial dataset by learning associations from another more available dataset. For example, Wu and Biljecki (2022) have developed GANmapper, an approach that uses Generative Adversarial Networks (GAN) to transform one spatial dataset to another without the need of any other external data sources

* Correspondence to: Department of Architecture, College of Design and Engineering, National University of Singapore, Singapore.
E-mail addresses: abrahamwu@u.nus.edu (A.N. Wu), filip@nus.edu.sg (F. Biljecki).

https://doi.org/10.1016/j.isprsjprs.2022.11.005
Received 5 February 2022; Received in revised form 10 November 2022; Accepted 11 November 2022
Available online 29 November 2022

OpenStreetMap | Edit | History | Export

Search | Where is this? | Go

Way: Preston Residential College (264858358)

Version #4

Columbia changes

Edited over 6 years ago by Royoriti
Changeset #50817763

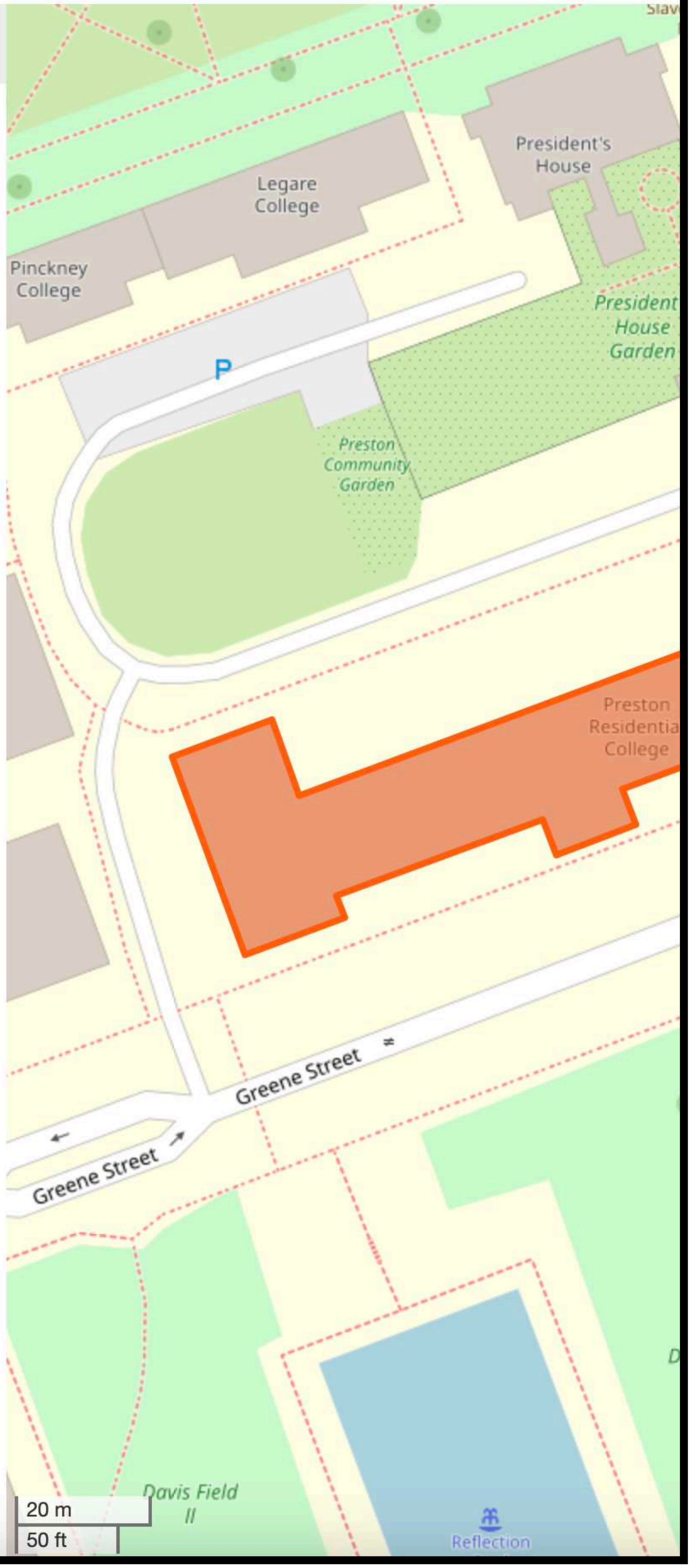
Tags

addr:city	Columbia
addr:housenumber	1323
addr:postcode	29225
addr:street	Greene Street
building	dormitory
building:levels	3
name	Preston Residential College

Nodes

▸ 17 nodes

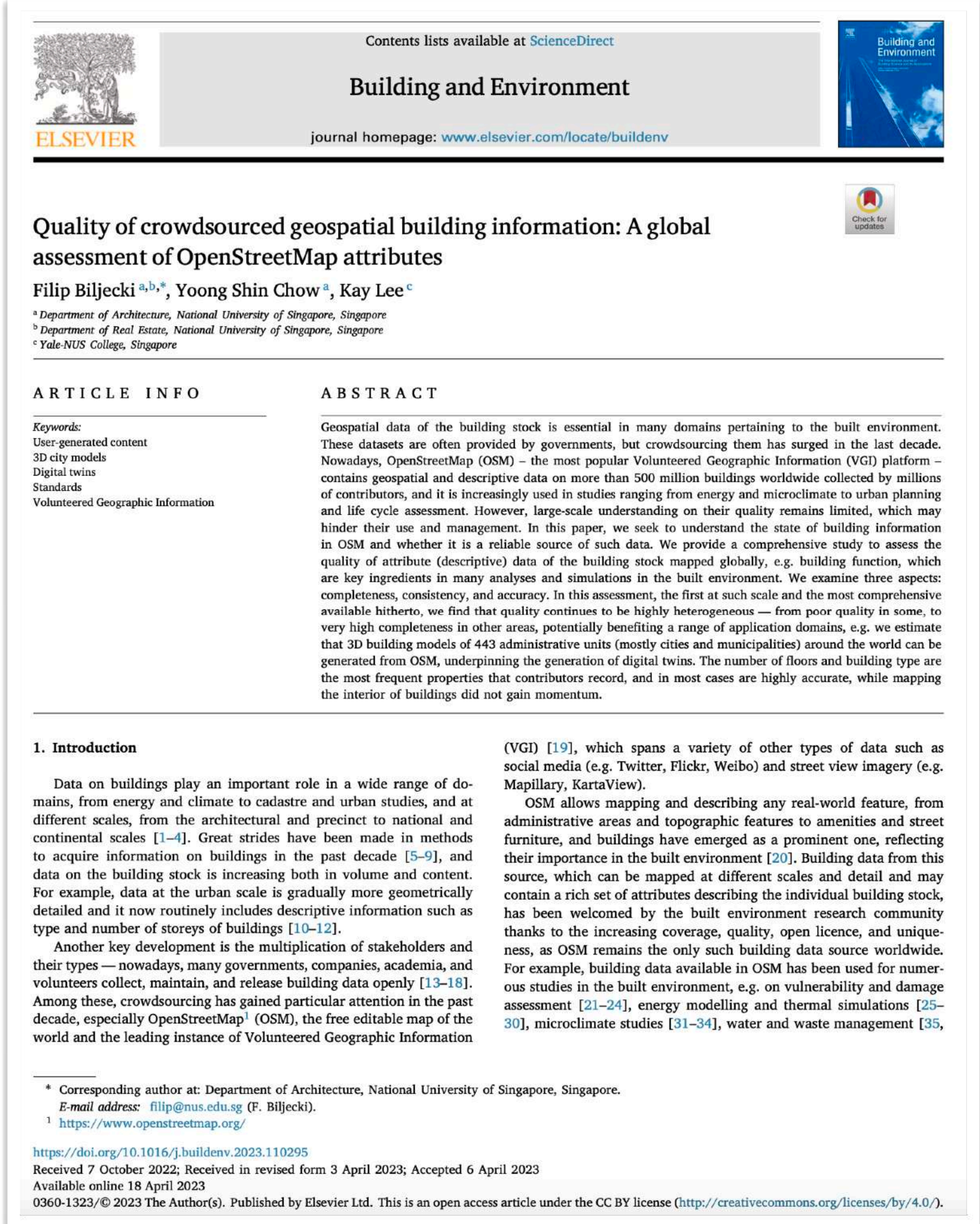
Download XML · View History



Quality of attributes of buildings in OSM

Global quality analysis: completeness and accuracy

- More than 600M buildings in OpenStreetMap
  - 22,000+ tags
  - Attributes have not been much in focus
- Findings
  - A few predominant tags around the world
    - 19.5% type of a building
    - 4.6% number of storeys
  - Highly heterogeneous completeness



Contents lists available at ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv

Quality of crowdsourced geospatial building information: A global assessment of OpenStreetMap attributes

Filip Biljecki [a,b,*], Yoong Shin Chow [a], Kay Lee [c]

[a] Department of Architecture, National University of Singapore, Singapore
[b] Department of Real Estate, National University of Singapore, Singapore
[c] Yale-NUS College, Singapore

ARTICLE INFO

Keywords:
User-generated content
3D city models
Digital twins
Standards
Volunteered Geographic Information

ABSTRACT

Geospatial data of the building stock is essential in many domains pertaining to the built environment. These datasets are often provided by governments, but crowdsourcing them has surged in the last decade. Nowadays, OpenStreetMap (OSM) – the most popular Volunteered Geographic Information (VGI) platform – contains geospatial and descriptive data on more than 500 million buildings worldwide collected by millions of contributors, and it is increasingly used in studies ranging from energy and microclimate to urban planning and life cycle assessment. However, large-scale understanding on their quality remains limited, which may hinder their use and management. In this paper, we seek to understand the state of building information in OSM and whether it is a reliable source of such data. We provide a comprehensive study to assess the quality of attribute (descriptive) data of the building stock mapped globally, e.g. building function, which are key ingredients in many analyses and simulations in the built environment. We examine three aspects: completeness, consistency, and accuracy. In this assessment, the first at such scale and the most comprehensive available hitherto, we find that quality continues to be highly heterogeneous — from poor quality in some, to very high completeness in other areas, potentially benefiting a range of application domains, e.g. we estimate that 3D building models of 443 administrative units (mostly cities and municipalities) around the world can be generated from OSM, underpinning the generation of digital twins. The number of floors and building type are the most frequent properties that contributors record, and in most cases are highly accurate, while mapping the interior of buildings did not gain momentum.

1. Introduction

Data on buildings play an important role in a wide range of domains, from energy and climate to cadastre and urban studies, and at different scales, from the architectural and precinct to national and continental scales [1–4]. Great strides have been made in methods to acquire information on buildings in the past decade [5–9], and data on the building stock is increasing both in volume and content. For example, data at the urban scale is gradually more geometrically detailed and it now routinely includes descriptive information such as type and number of storeys of buildings [10–12].

Another key development is the multiplication of stakeholders and their types — nowadays, many governments, companies, academia, and volunteers collect, maintain, and release building data openly [13–18]. Among these, crowdsourcing has gained particular attention in the past decade, especially OpenStreetMap[1] (OSM), the free editable map of the world and the leading instance of Volunteered Geographic Information (VGI) [19], which spans a variety of other types of data such as social media (e.g. Twitter, Flickr, Weibo) and street view imagery (e.g. Mapillary, KartaView).

OSM allows mapping and describing any real-world feature, from administrative areas and topographic features to amenities and street furniture, and buildings have emerged as a prominent one, reflecting their importance in the built environment [20]. Building data from this source, which can be mapped at different scales and detail and may contain a rich set of attributes describing the individual building stock, has been welcomed by the built environment research community thanks to the increasing coverage, quality, open licence, and uniqueness, as OSM remains the only such building data source worldwide. For example, building data available in OSM has been used for numerous studies in the built environment, e.g. on vulnerability and damage assessment [21–24], energy modelling and thermal simulations [25–30], microclimate studies [31–34], water and waste management [35,

* Corresponding author at: Department of Architecture, National University of Singapore, Singapore.
E-mail address: filip@nus.edu.sg (F. Biljecki).

1 https://www.openstreetmap.org/

https://doi.org/10.1016/j.buildenv.2023.110295
Received 7 October 2022; Received in revised form 3 April 2023; Accepted 6 April 2023
Available online 18 April 2023

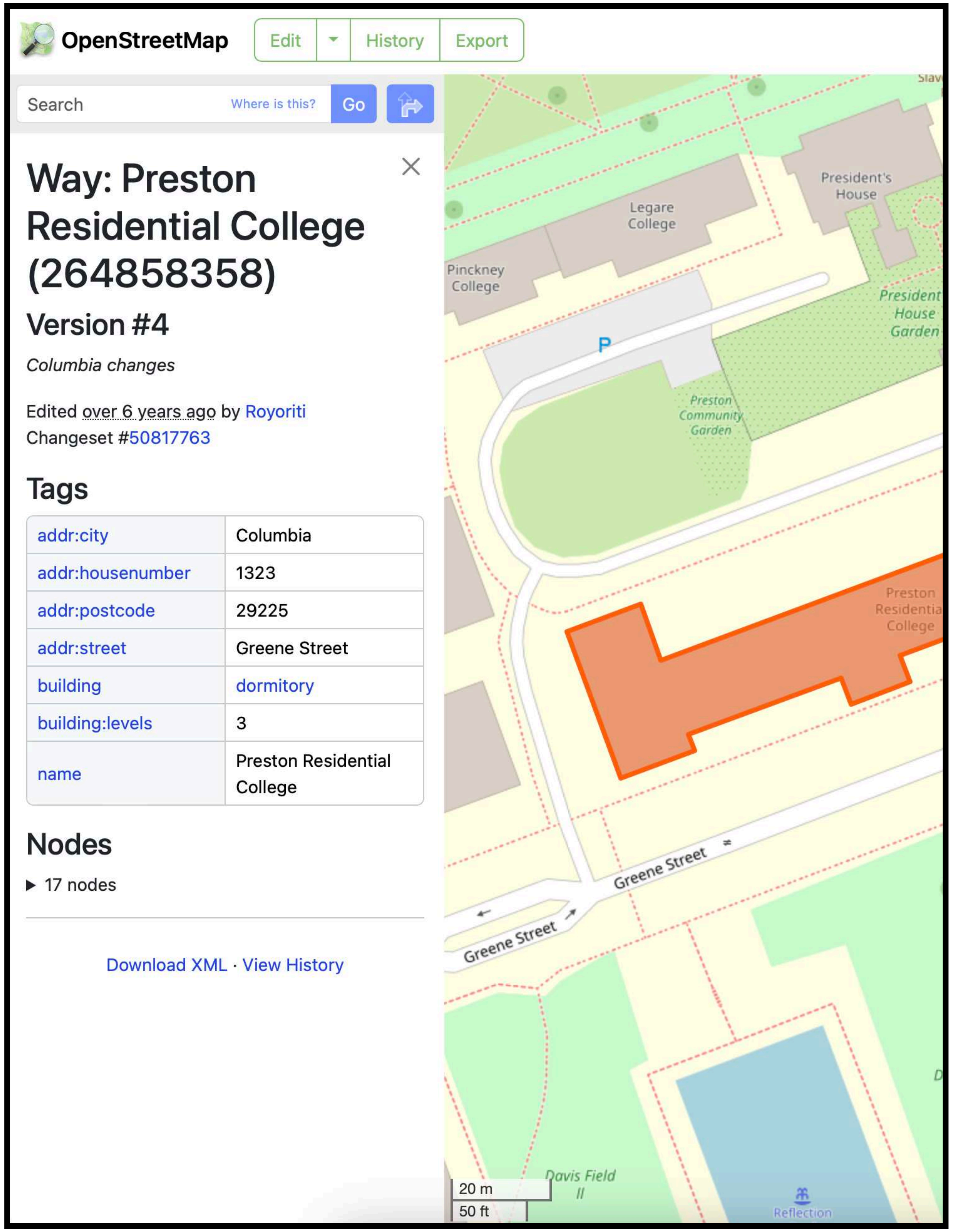
Biljecki F, Chow YS, Lee K (2023): Quality of crowdsourced geospatial building information: A global assessment of OpenStreetMap attributes. Building and Environment 237: 110295.

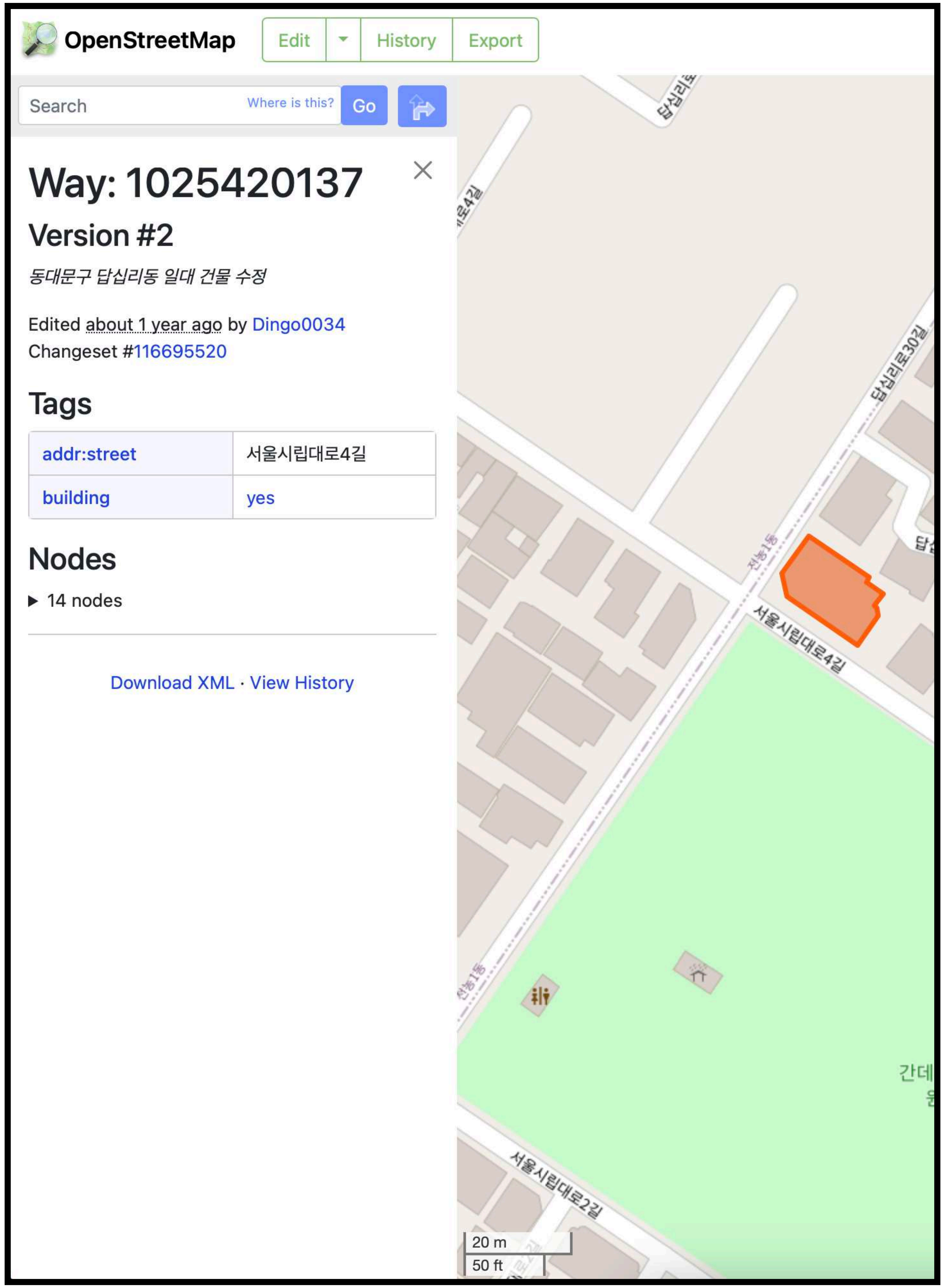

Alger, Algeria
Berlin, Germany
Brisbane, Australia
Brno, Czechia
Calgary, Canada
Chicago, United States
Dallas, United States
East Lothian, Scotland
Dar Es Salaam, Tanzania
Dublin, Ireland
Greater Accra, Ghana
Greater London, England
Madrid, Spain
Kiev, Ukraine
Lima, Peru
Minsk, Belarus
Moscow, Russia
Oslo, Norway
Pulau Penang, Malaysia
Zurich, Switzerland
0
25
50
75
100
building:levels
Completeness [%]
Argentina
Australia
Azerbaijan
Belarus
Bolivia
Czechia
Eritrea
Germany
Iran
Kazakhstan
Pakistan
Poland
South Korea
Ukraine
Venezuela
0
25
50
75
100
building:levels
Completeness [%]

Accuracy of tags

A few key attributes

- Thousands of buildings
  - Type 84%
  - Storeys 72%
  - Shape of the roof 82%
- High completeness + High accuracy = OSM can be quite good in thousands of districts
  - But some attributes important for certain domains (e.g. material) are very scarce


OpenStreetMap
Edit
History
Export
Search
Where is this?
Go
Way: Preston Residential College (264858358)
Version #4
Columbia changes
Edited over 6 years ago by Royoriti
Changeset #50817763
Tags
addr:city Columbia
addr:housenumber 1323
addr:postcode 29225
addr:street Greene Street
building dormitory
building:levels 3
name Preston Residential College
Nodes
17 nodes
Download XML · View History
Legare College
President's House
Pinckney College
President House Garden
Preston Community Garden
Preston Residential College
Greene Street
Greene Street
Davis Field
Reflection
20 m
50 ft


OpenStreetMap
Edit
History
Export
Search
Where is this?
Go
Way: 1025420137
Version #2
동대문구 답십리동 일대 건물 수정
Edited about 1 year ago by Dingo0034
Changeset #116695520
Tags
addr:street 서울시립대로4길
building yes
Nodes
14 nodes
Download XML · View History
서울시립대로4길
서울시립대로2길
20 m
50 ft

Can we use AI to improve data on buildings?
Part 2

Hedges Park Condominium

Condominium

Overview | Home Finance | Price Insights | Location | Shortlist | Share

S$ 2,088,888 Negotiable

Shortlist | Hide | Share | PDF

Report Listing

4 4 1539 sqft S$ 1,357.30 psf

Est. Repayment S$ 5,791 /mo Get the best rates

Hedges Park Condominium

81 Flora Drive 506886 Changi Airport / Changi Village (D17)

MAP

Details

Property Type Condominium For Sale	Floor Size 1539 sqft
Developer Tripartite Developers Pte Limited	PSF S$ 1,357.30 psf
Furnishing Unfurnished	Floor Level Ground Floor
Tenure 99-year Leasehold	TOP June, 2015
Listing ID 24359099	Currently Tenanted No
Listed On 32 seconds ago	Maintenance Fee S$ 450.00 /mo

Are real estate ads a type of user-generated geographic information that has been ignored in GIScience?

Source: PropertyGuru

Introducing new means of acquisition of building data

By Xinyu Chen

Urban Informatics Paper of the Year Award (2023)

Chen and Biljecki *Urban Informatics* (2022) 1:12
https://doi.org/10.1007/s44212-022-00012-2

Urban Informatics

ORIGINAL ARTICLE Open Access

Mining real estate ads and property transactions for building and amenity data acquisition

Xinyu Chen[1] and Filip Biljecki[2,3*]

Abstract

Acquiring spatial data of fine and dynamic urban features such as buildings remains challenging. This paper brings attention to real estate advertisements and property sales data as valuable and dynamic sources of geoinformation in the built environment, but unutilised in spatial data infrastructures. Given the wealth of information they hold and their user-generated nature, we put forward the idea of real estate data as an instance of implicit volunteered geographic information and bring attention to their spatial aspect, potentially alleviating the challenge of acquiring spatial data of fine and dynamic urban features. We develop a mechanism of facilitating continuous acquisition, maintenance, and quality assurance of building data and associated amenities from real estate data. The results of the experiments conducted in Singapore reveal that one month of property listings provides information on 7% of the national building stock and about half of the residential subset, e.g. age, type, and storeys, which are often not available in sources such as OpenStreetMap, potentially supporting applications such as 3D city modelling and energy simulations. The method may serve as a novel means to spatial data quality control as it detects missing amenities and maps future buildings, which are advertised and transacted before they are built, but it exhibits mixed results in

Situation
Real estate data
(1) Real estate ads
(2) Property transactions
Extracting relevant data
(A) Detecting features (computer vision)
Swimming Pool
(B) Collecting the location of a building
N1.295423, E103.8357
(C) Mining descriptions/characteristics
e.g. building type and year of construction
Type
TOP
Condominium For Sale
1993
e.g. height from address and floor area
#24-63 42
#23-65 97
#23-64 84
#23-63 42
Spatial database and downstream applications
(i) Quality control: revealing unmapped features
River Valley Road
Jalan Kuala
Saint Thomas Wk
Opposite Saint Thomas Wk
detected missing swimming pool
The Morningside
(ii) Sensing future buildings
(e.g. planned ones that are not yet mapped)
(iii) Quality control of existing attributes
type condominium
start_date 1997
(iv) Data enrichment: new attributes
levels 24
avg_floor_area 80
Example use cases: supporting energy simulations and 3D model generation
Floor area provides value for energy consumption estimations
Number of floors, which can be inferred from corresponding property transactions, can be used to extrude existing footprints, resulting in 3D building models

How can AI help us develop new use cases at the building and city scales?

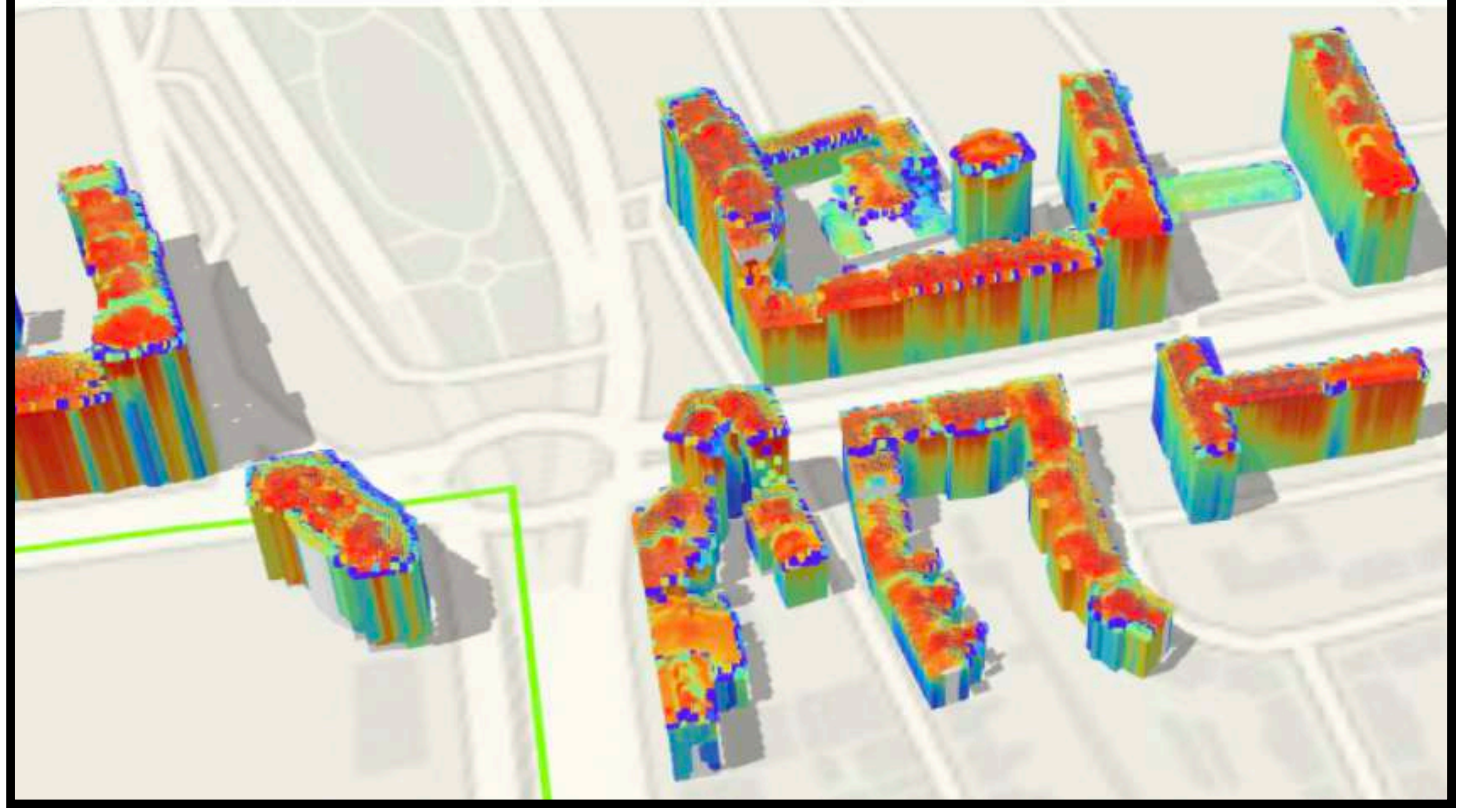
Unlocking rooftops

Potential for green roofs, farms, solar panels…

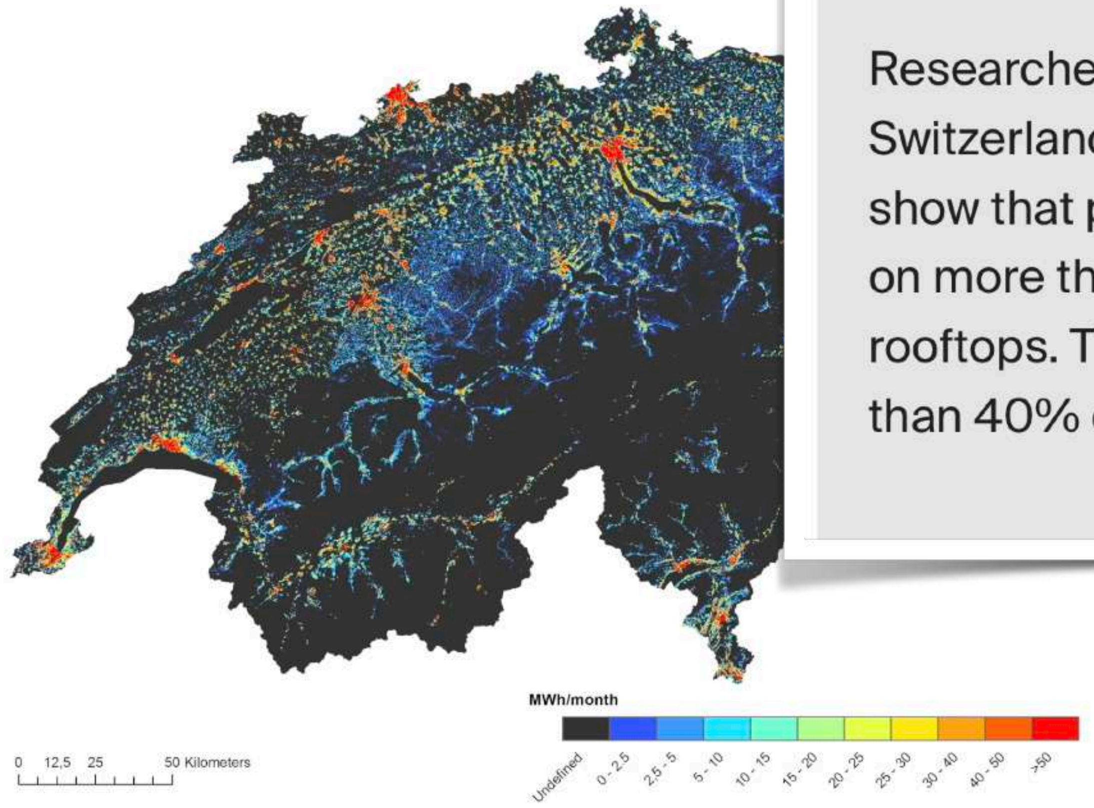
Unlocking rooftops

Research on potential for solar panels, green roofs, farms, …

(1) Identifying the best locations



(2) Quantifying the potential at large-scale



EPFL NEWS EPFL ENAC SB STI IC SV CDM CDH All EN

› News

What if half of Switzerland's rooftops produced electricity?

Researchers at EPFL are assessing Switzerland's solar power potential. Their results show that photovoltaic panels could be installed on more than half of the country's 9.6 million rooftops. The resulting power would meet more than 40% of Swiss electricity demand.

Brito, M. C., Redweik, P., Catita, C., Freitas, S. & Santos, M. 3D Solar Potential in the Urban Environment: A Case Study in Lisbon. Energies 12, 3457–13 (2019).

https://actu.epfl.ch/news/what-if-half-of-switzerland-s-rooftops-produced-el/

We know the potential in future. But what about the current situation and actual status *today*? What is the rooftop utilisation rate?

Roofpedia

Global open registry of roofs for urban sustainability

Roofpedia

Global open registry of roofs for urban sustainability

Published in Landscape and Urban Planning. Credit: Abraham Noah Wu by Abraham Noah Wu

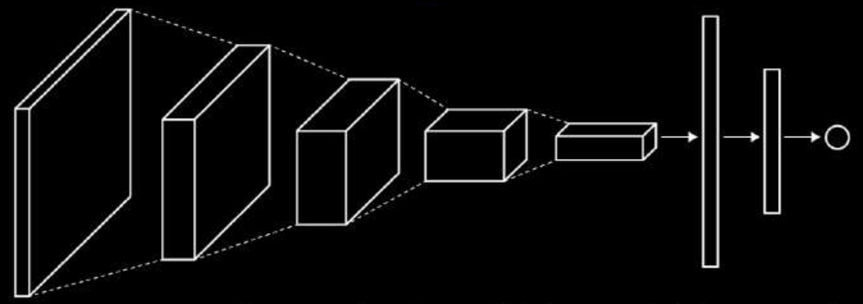
ROOFPEDIA

Automated Roof Mapping + Geospatial Roof Registry + Sustainable Roof Index

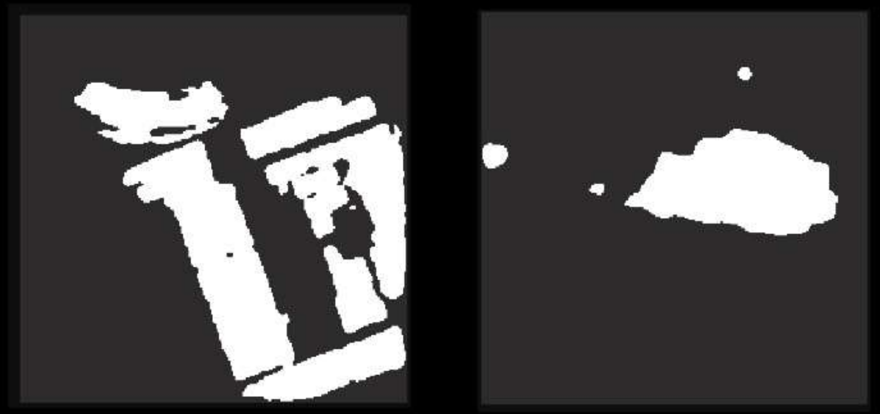
NUS National University of Singapore | urban analytics lab

Automated Classification

Satellite Images



Convolutional Neural Network

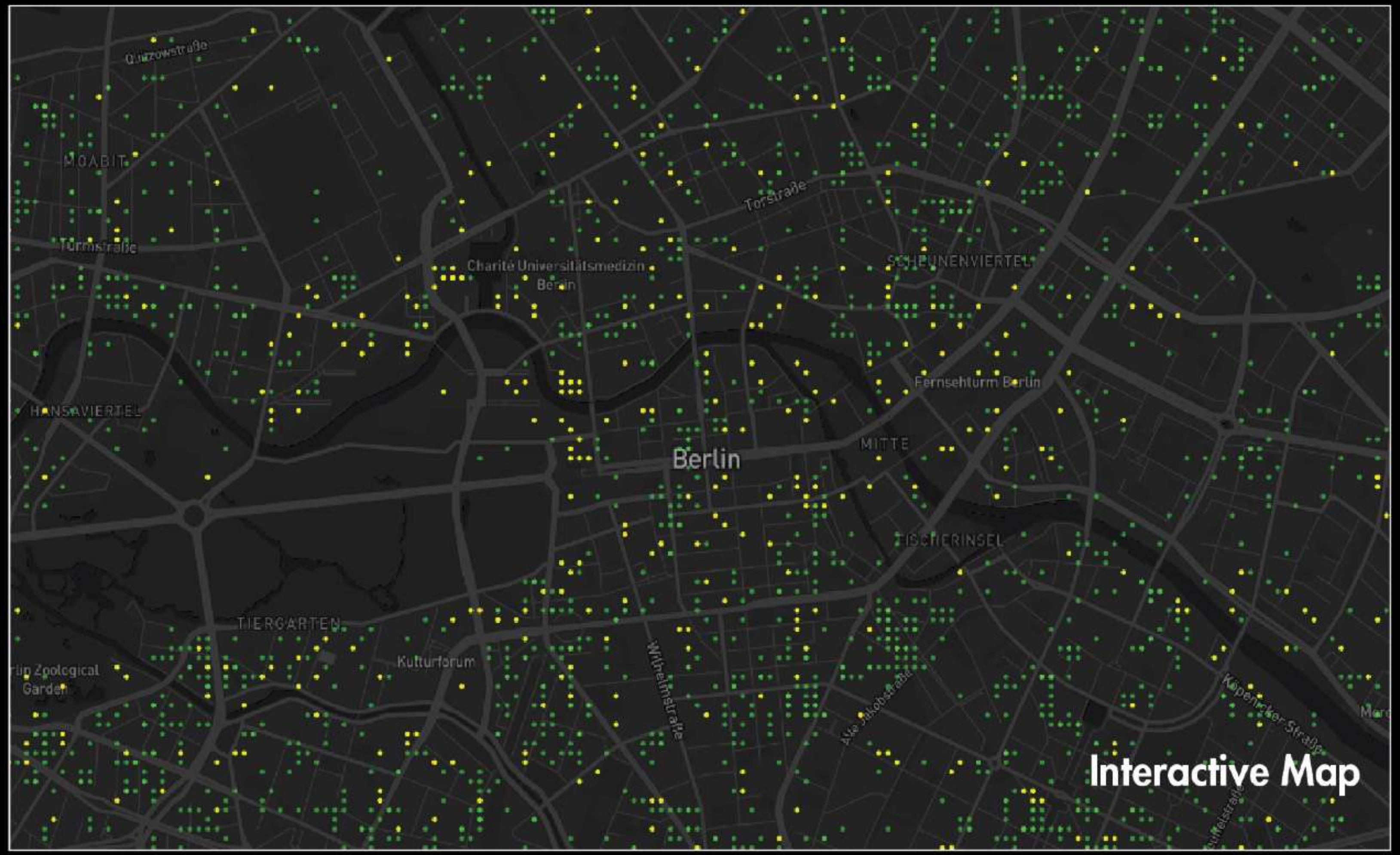


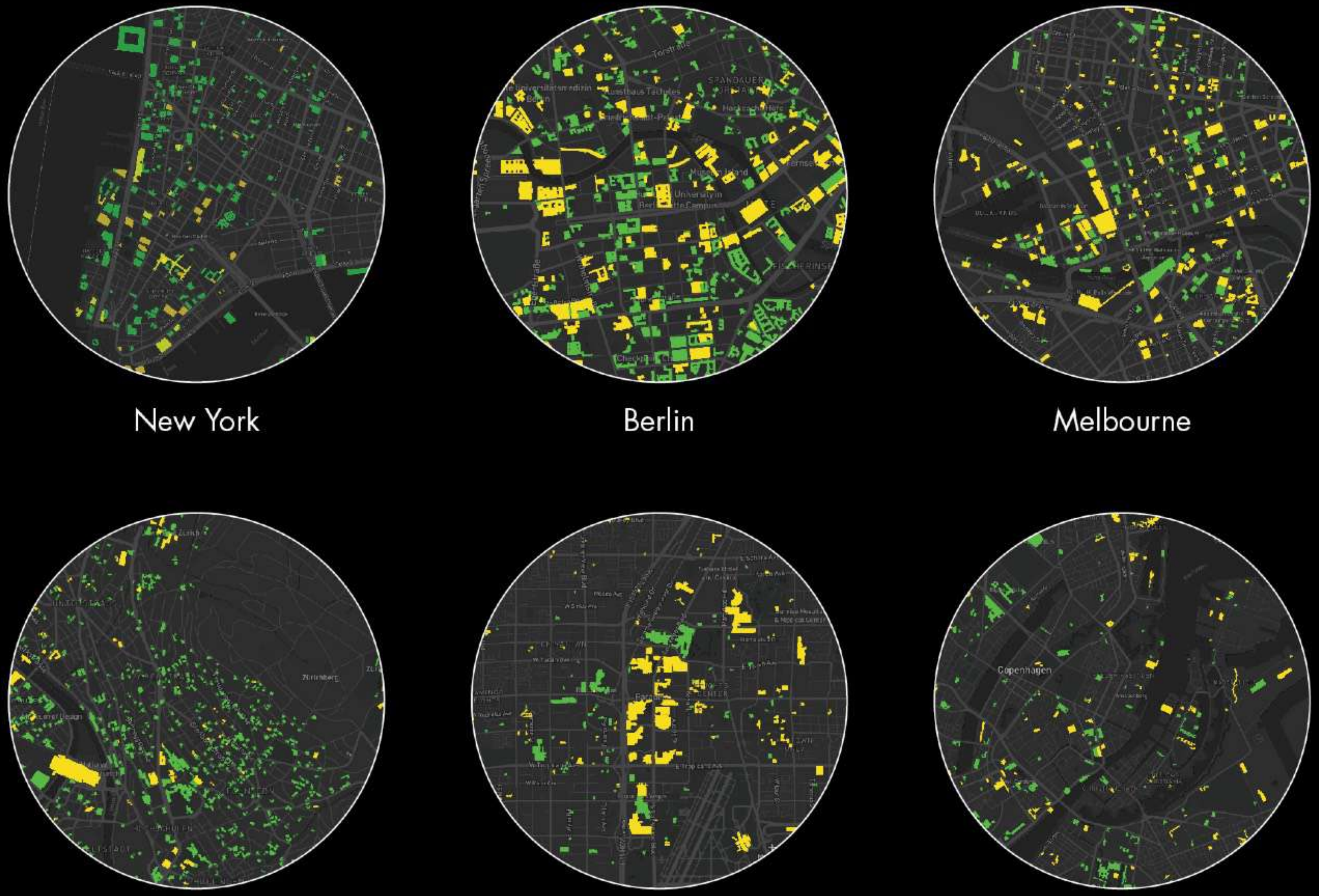
Rooftop Solar Panels — Rooftop Vegetation

GIS Processing

Solar Roofs — Green Roofs

Roofpedia Registry





New York — Berlin — Melbourne

Zurich — Las Vegas — Copenhagen

Roofpedia Indices

Solar Roof Index

Las Vegas	86
Zurich	81
Singapore	75
Phoenix	75
Melbourne	74
Berlin	57
Copenhagen	45
New York	42
Paris	42
San Diego	24
Los Angeles	20
Seattle	13
San Jose	12
Portland	10
San Francisco	9
Luxembourg City	7
Vancouver	0

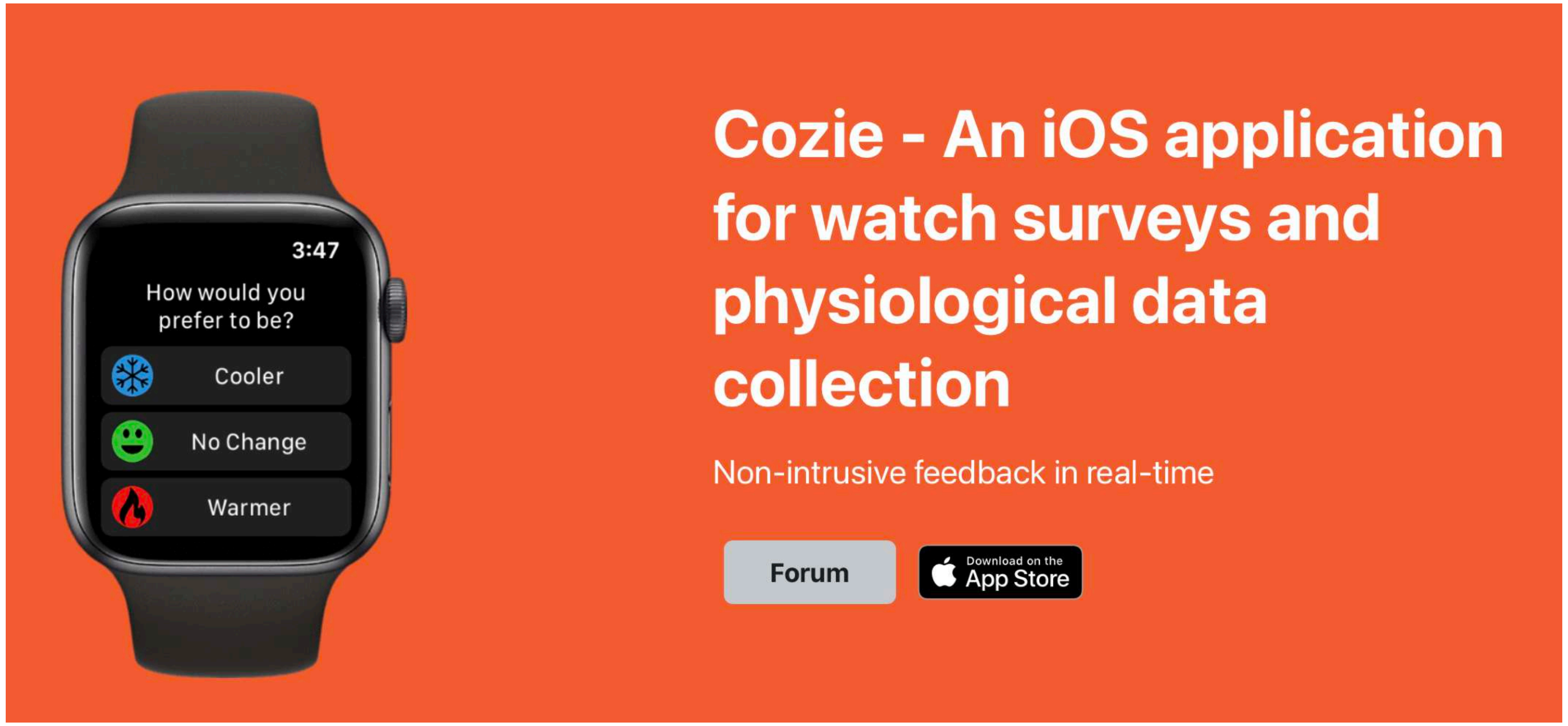
Green Roof Index

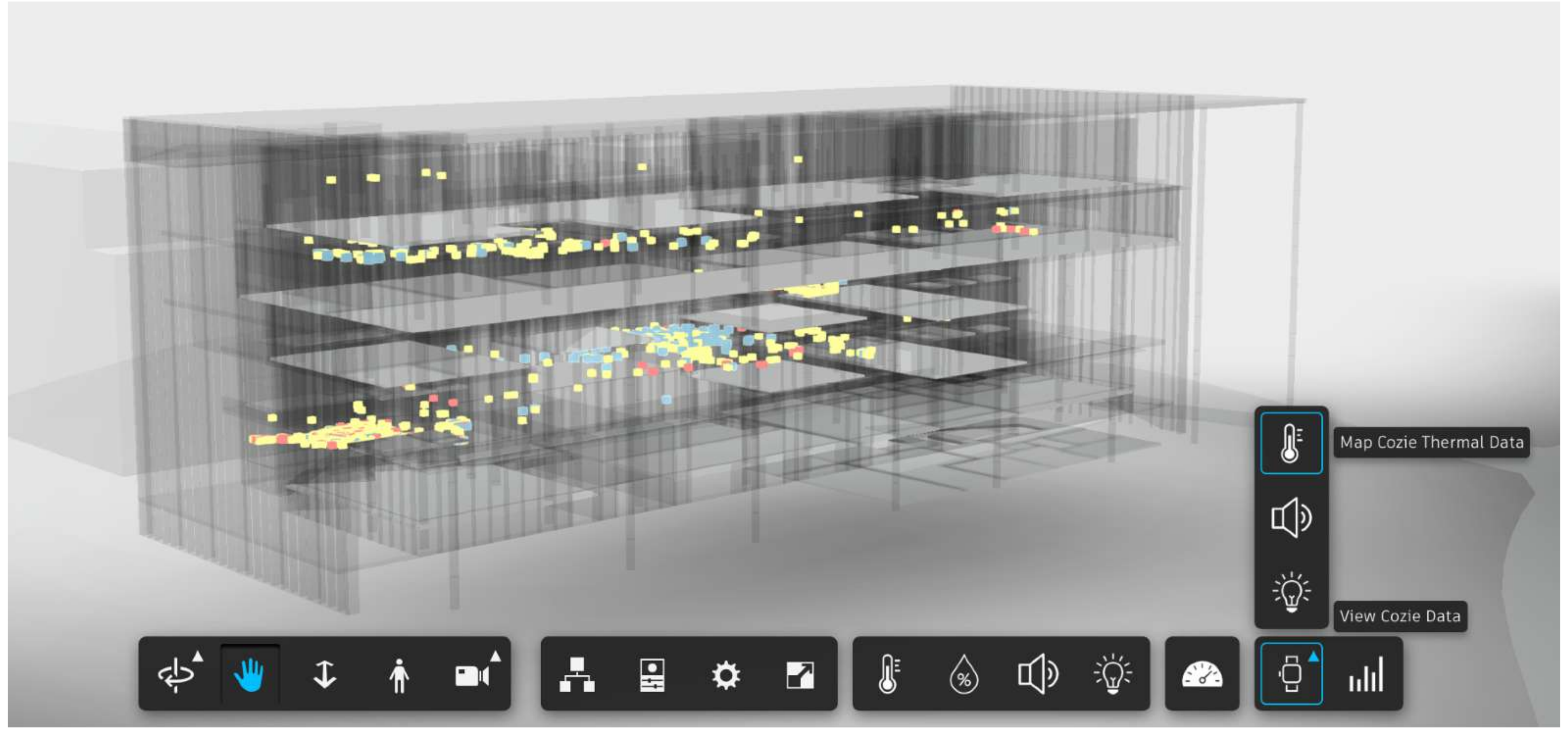
Zurich	100
Berlin	51
New York	28
Copenhagen	22
Paris	18
San Diego	14
San Jose	13
Phoenix	13
Melbourne	11
Las Vegas	9
Seattle	6
Los Angeles	6
Luxembourg City	4
Portland	3
San Francisco	2
Vancouver	0

Data sources: Mapbox and OpenStreetMap contributors

Where are humans?

Cozie Platform: Collecting Occupant Data at Scale in the Built Environment





buds lab
building and urban data science

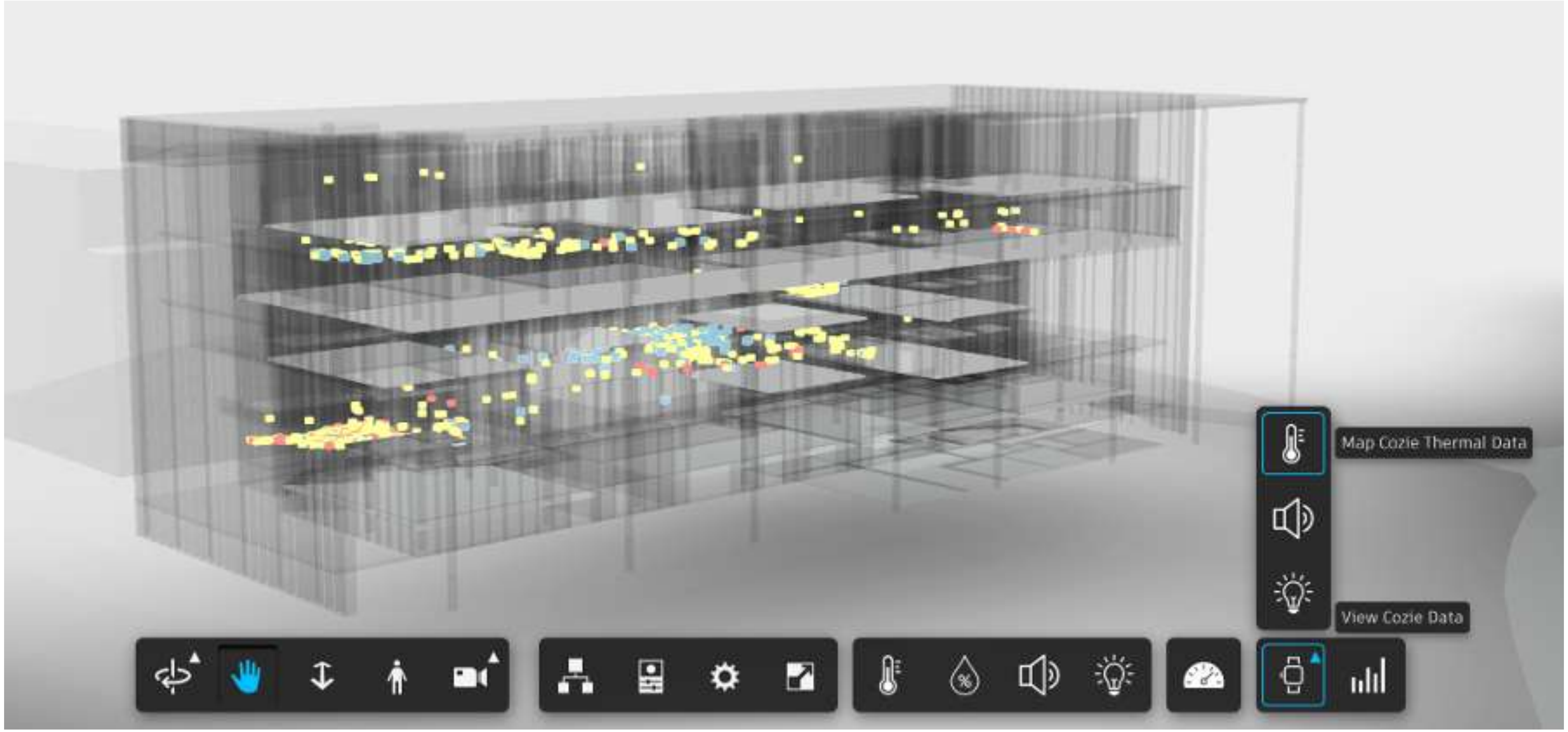
https://cozie-apple.app/
Jayathissa P, Quintana M, Abdelrahman M, Miller C. Humans-as-a-Sensor for Buildings—Intensive Longitudinal Indoor Comfort Models. Buildings. 10: 174, 2020. https://doi.org/10.3390/buildings10100174

Micro-survey (EMA) Watch-based Question Flows

buds lab
building and urban data science

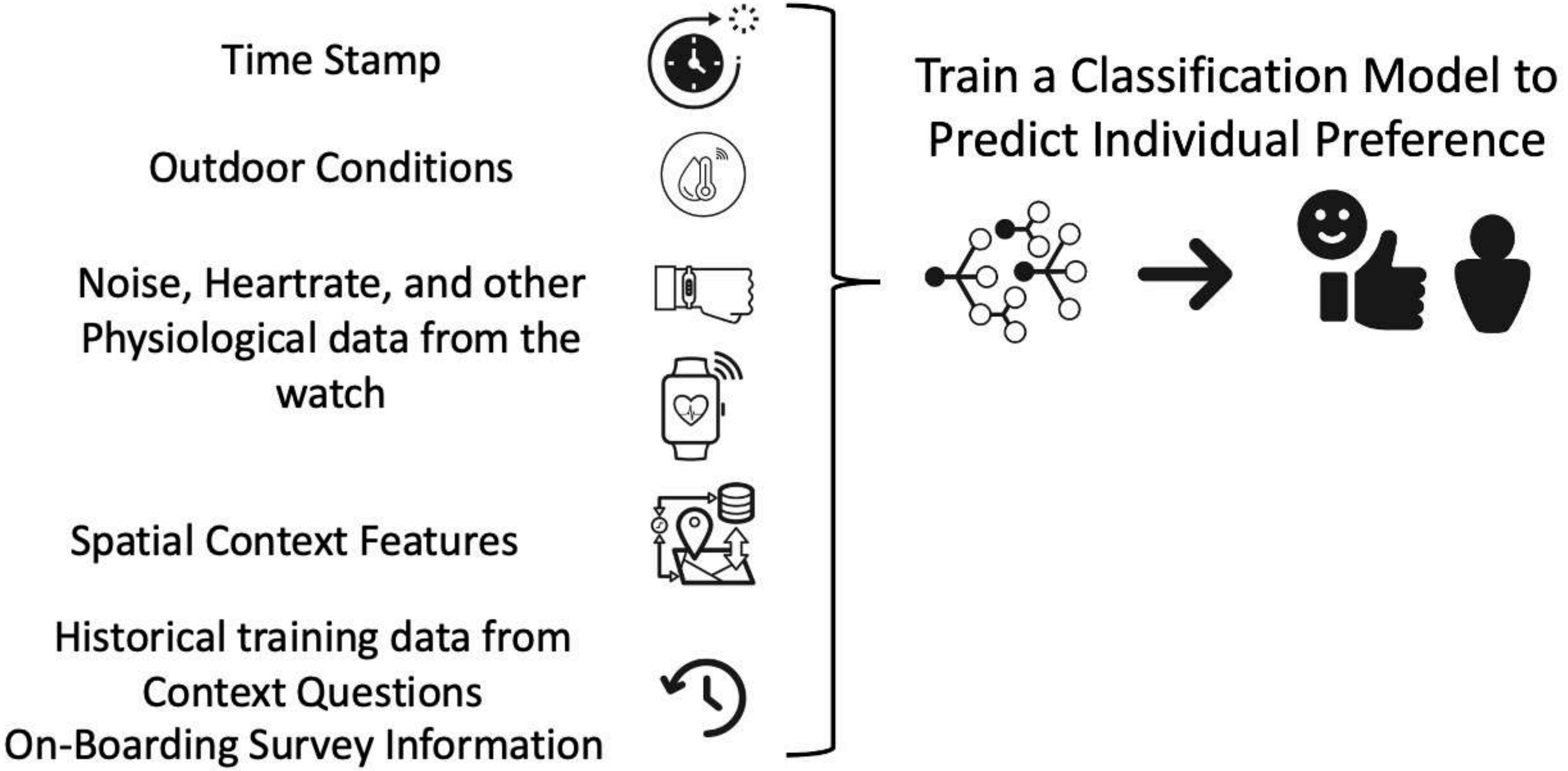
Tartarini, F., Miller, C., & Schiavon, S. (2023). Cozie Apple: An iOS mobile and smartwatch application for environmental quality satisfaction and physiological data collection. Journal of Physics. Conference Series. https://doi.org/10.1088/1742-6596/2600/14/142003

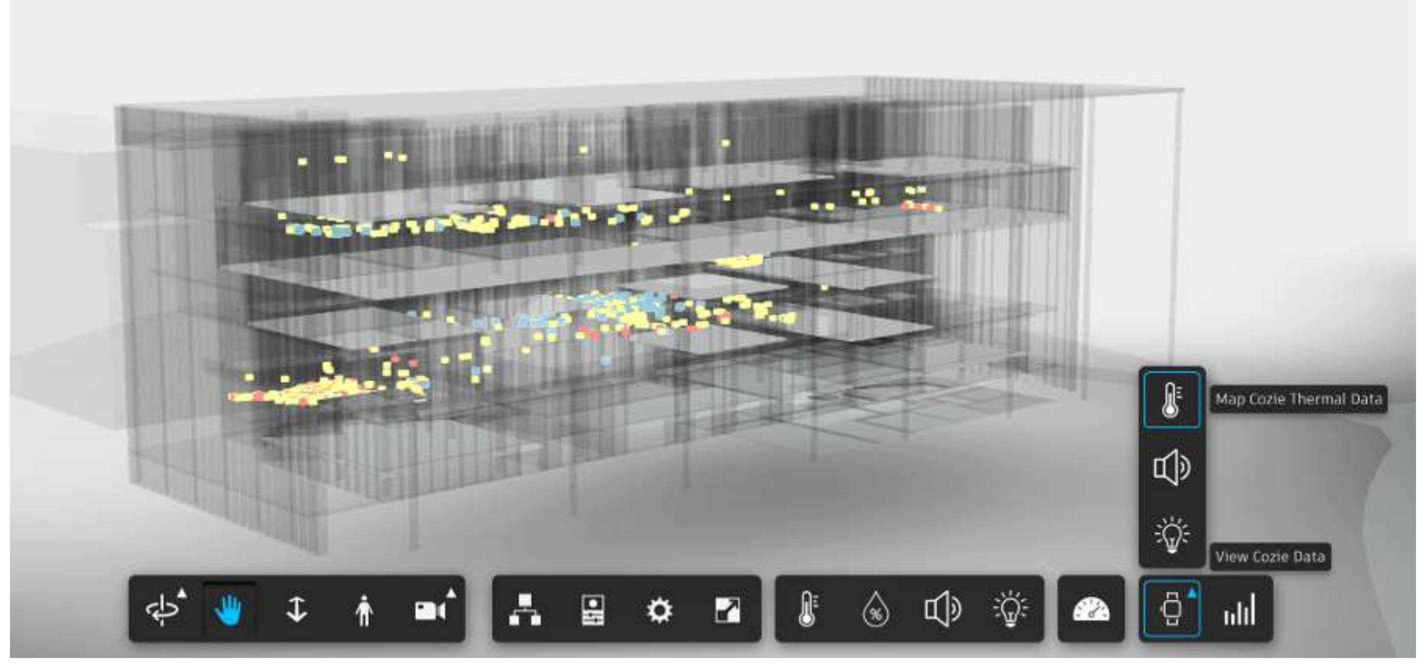
Scalable Field-based Data Collection



https://cozie-apple.app/
Jayathissa P, Quintana M, Abdelrahman M, Miller C. Humans-as-a-Sensor for Buildings—Intensive Longitudinal Indoor Comfort Models. Buildings. 10: 174, 2020.
https://doi.org/10.3390/buildings10100174

Developing AI Comfort Models for different ‘Personality Types’



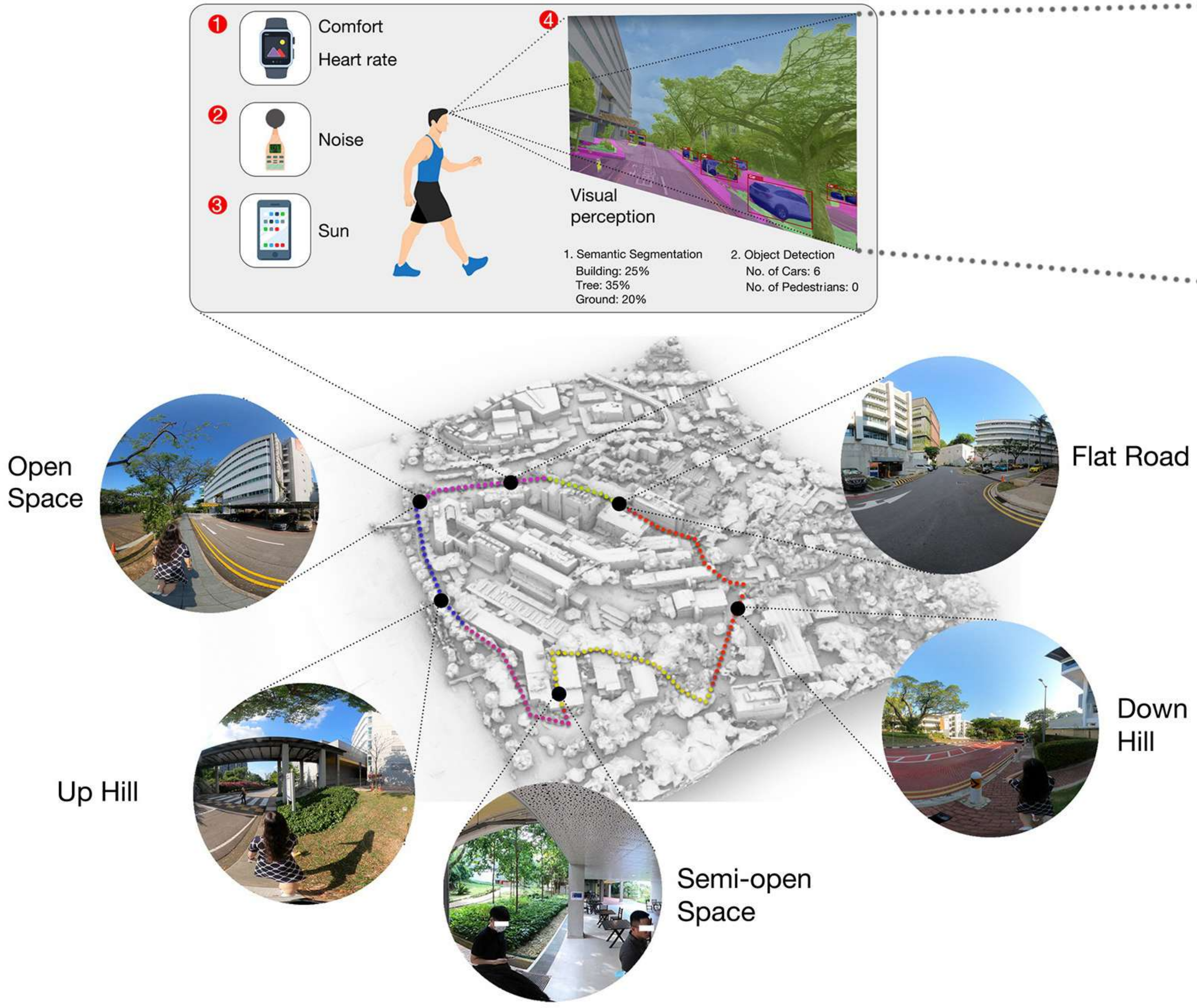


Jayathissa P, Quintana M, Abdelrahman M, Miller C. Humans-as-a-Sensor for Buildings—Intensive Longitudinal Indoor Comfort Models. Buildings. 2020;10: 174. https://doi.org/10.3390/buildings10100174
Quintana, M., Schiavon, S., Tartarini, F., Kim, J., & Miller, C. (2023). Cohort comfort models — Using occupant’s similarity to predict personal thermal preference with less data. Building and Environment, 227, 109685. https://doi.org/10.1016/j.buildenv.2022.109685
Miller, C., Quintana, M., Frei, M., Chua, Y. X., Fu, C., Picchetti, B., Yap, W., Chong, A., & Biljecki, F. (2023). Introducing the Cool, Quiet City Competition: Predicting Smartwatch-Reported Heat and Noise with Digital Twin Metrics. Proceedings of the 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, 298–299. https://doi.org/10.1145/3600100.3626269

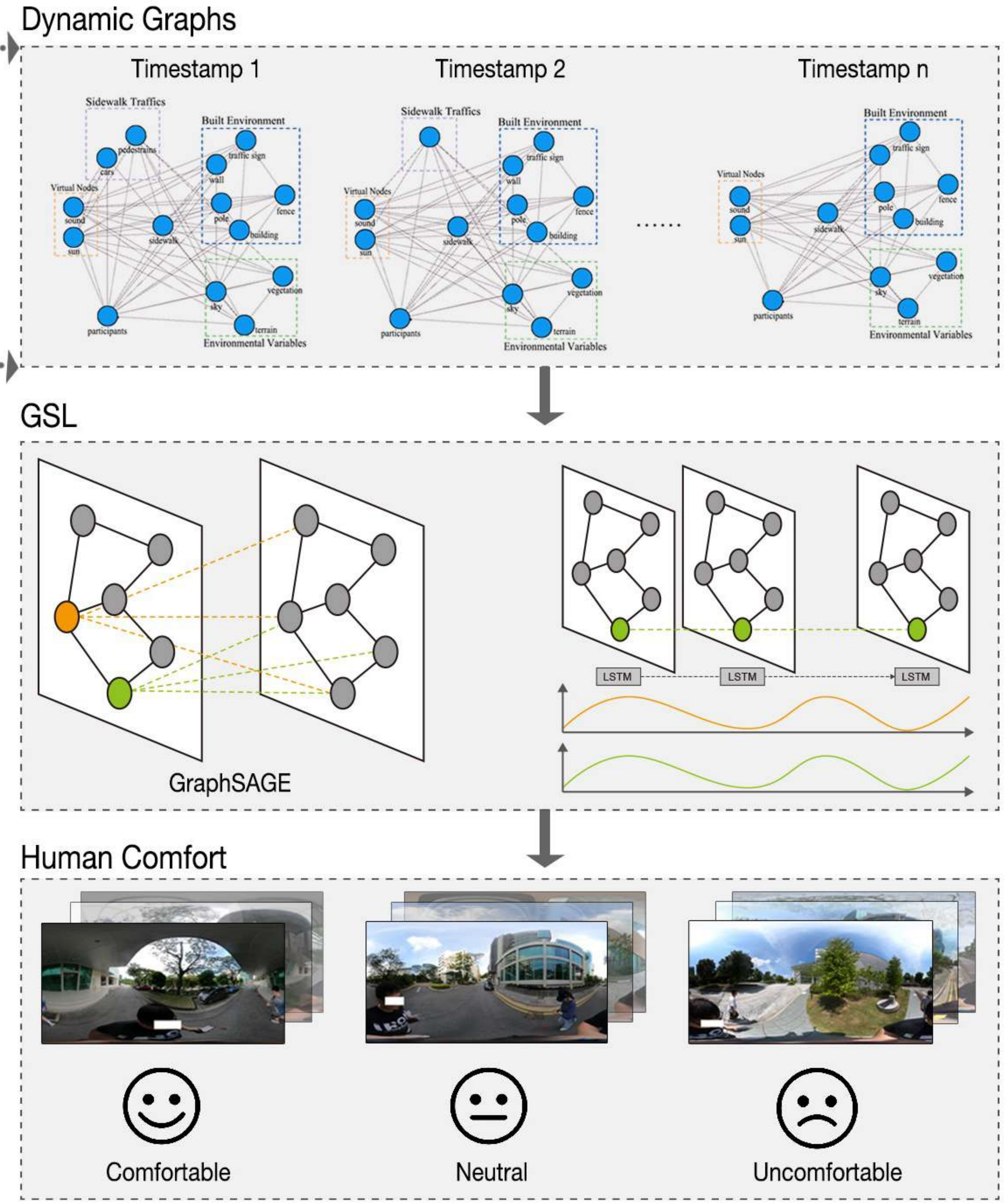
Thinking beyond thermal comfort and buildings…

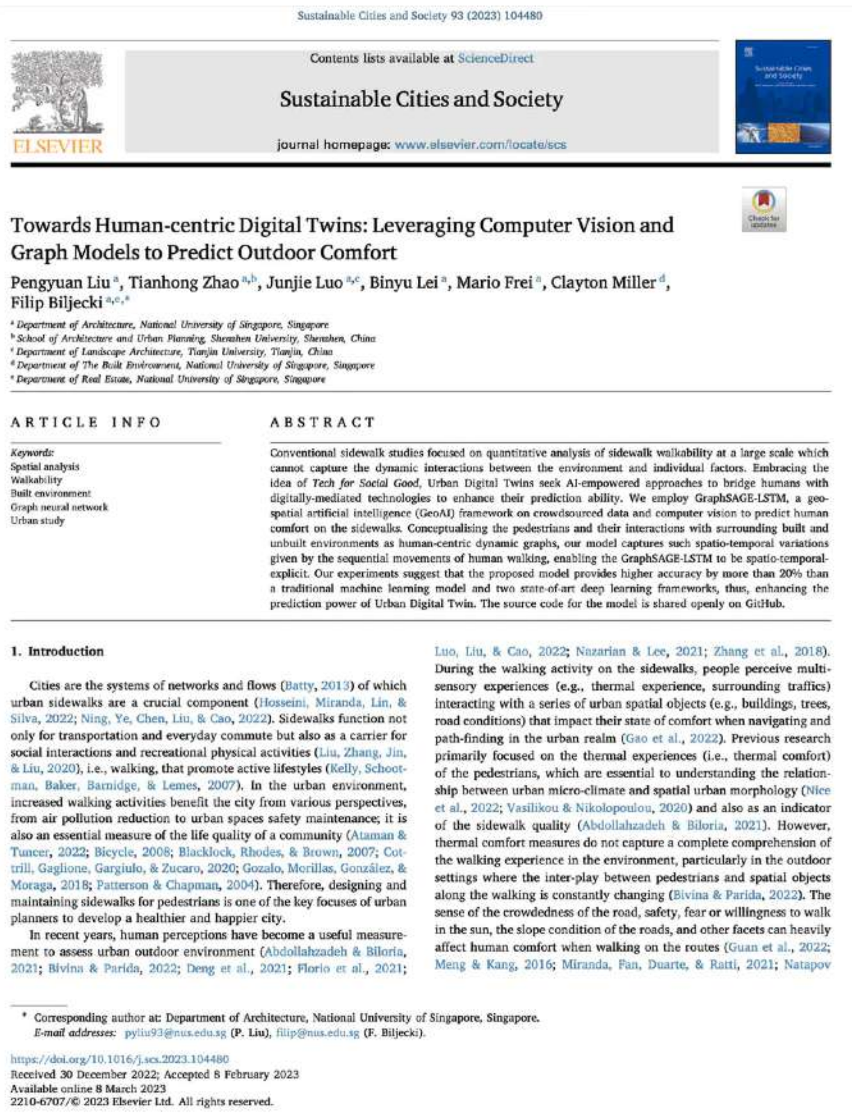
- Comfort is more than thermal comfort
- Comfort is influenced by myriads of factors
- AI can help us understand that

Predicting Human Comfort on the Sidewalk



Methodology





Sustainable Cities and Society 93 (2023) 104480

Sustainable Cities and Society

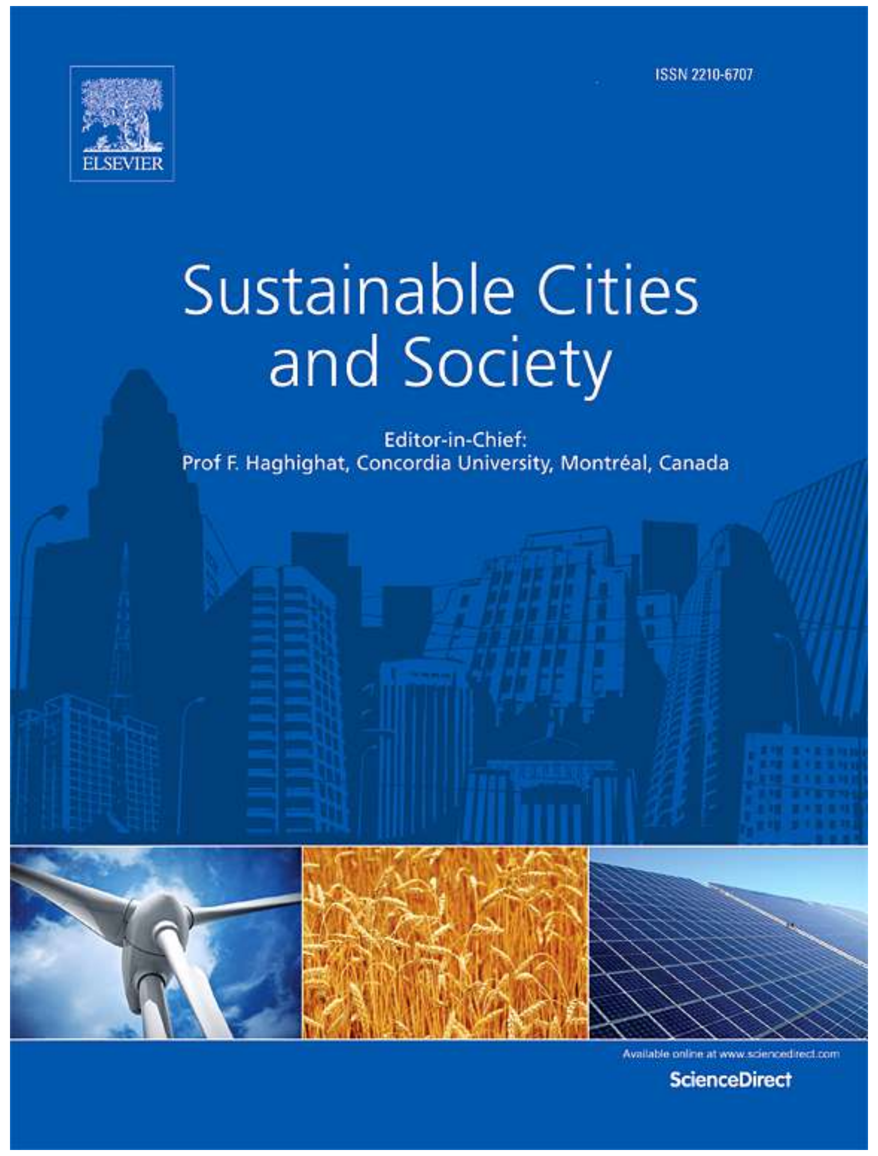
Towards Human-centric Digital Twins: Leveraging Computer Vision and Graph Models to Predict Outdoor Comfort

Pengyuan Liu, Tianhong Zhao, Junjie Luo, Binyu Lei, Mario Frei, Clayton Miller, Filip Biljecki

ARTICLE INFO

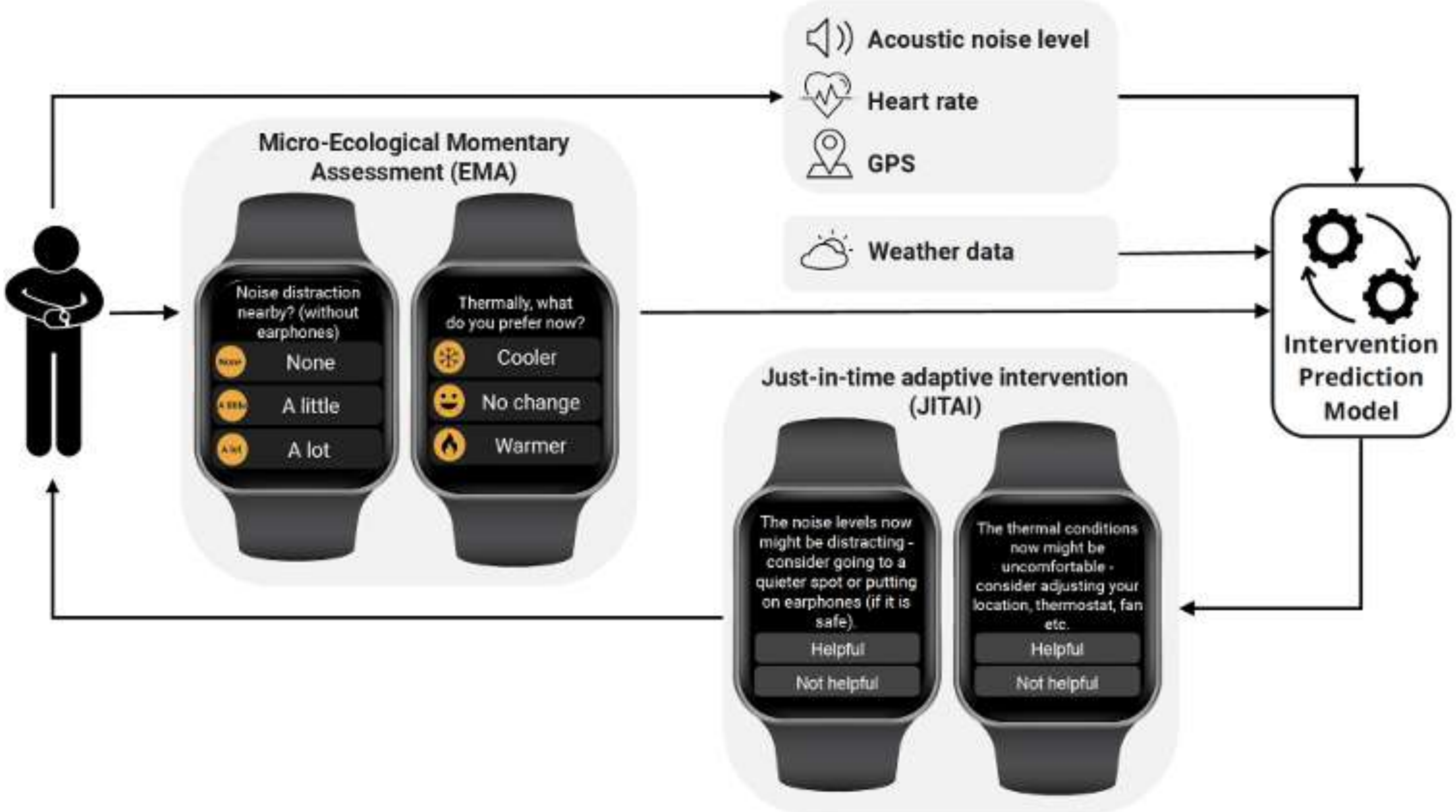
ABSTRACT

1. Introduction



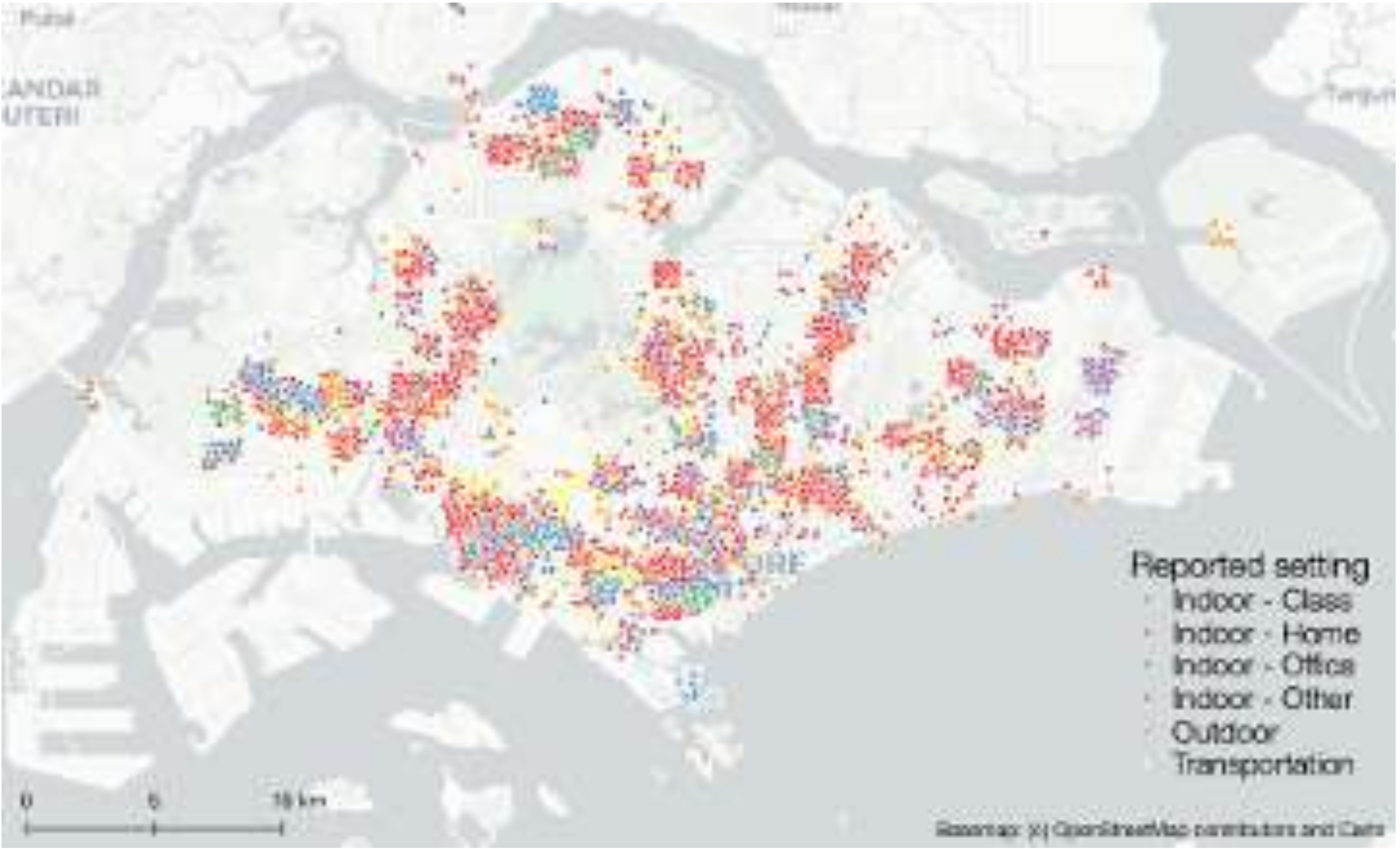
Liu P, Zhao T, Luo J, Lei B, Frei M, Miller C, and Biljecki F (2023): Towards Human-centric Digital Twins: Leveraging Computer Vision and Graph Models to Predict Outdoor Comfort. Sustainable Cities and Society 93: 104480

Our current experiment: Just-in-time Adaptive Intervention (JITAI) Messages



Miller, C., Chua, Y. X., Frei, M., & Quintana, M. (November 9-10 2022). Towards smartwatch-driven just-in-time adaptive interventions (JITAI) for building occupants. The 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation. https://doi.org/10.1145/3563357.3566135

Scalability of Data Collection across a city

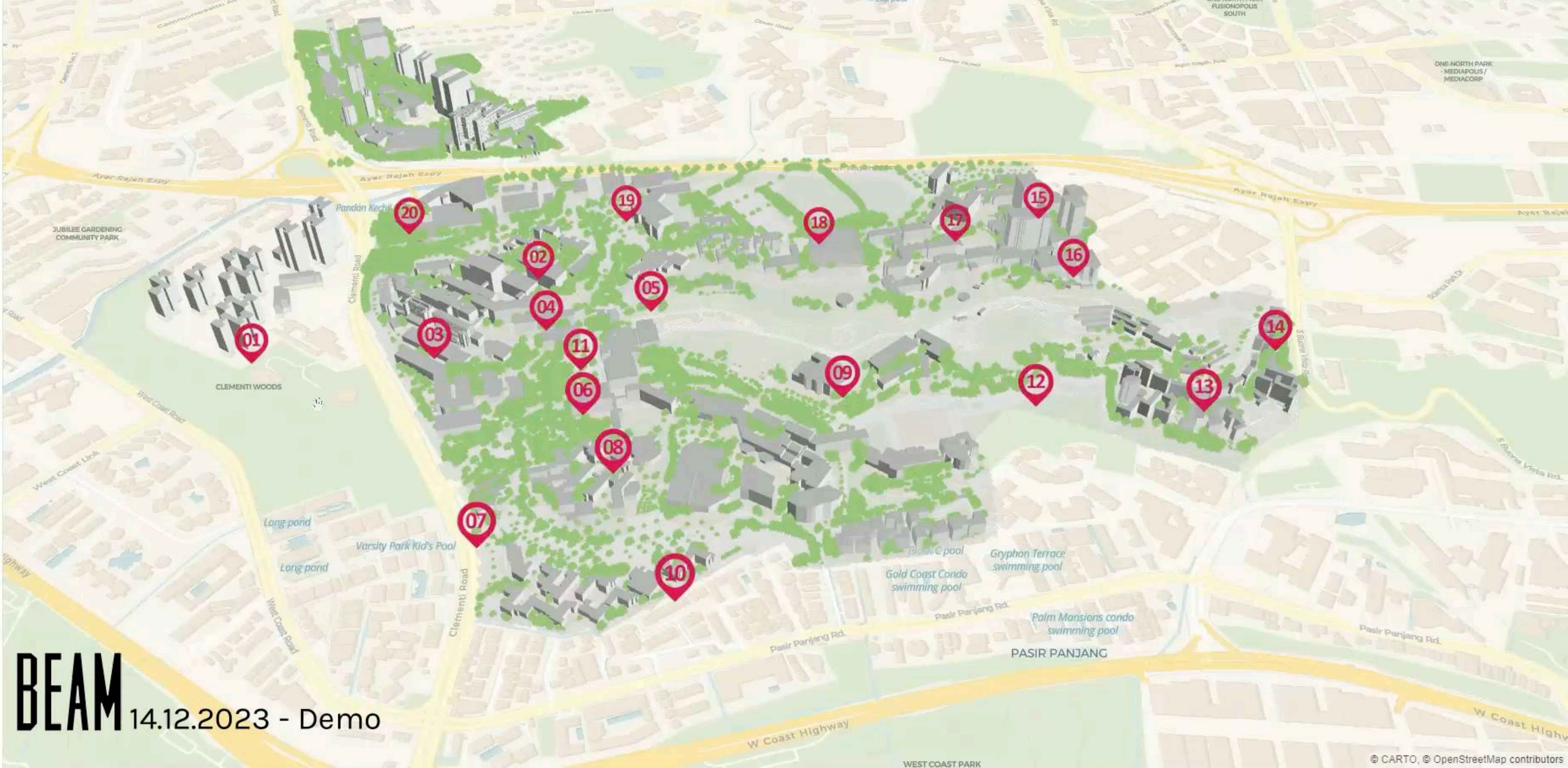


Multi-scale integration in a digital twin

RESET VIEW

[1] Historical Weather Data

20 Stations (Feb 2019 - Apr 2020)



BEAM 14.12.2023 - Demo

Conclusions, takeaways, and future directions

- **Emerging data sources** can help us gather further insights in the built environment. **Crowdsourced data** will grow in importance and availability in the future
- With the growth of AI techniques and the above data, it’s important to set up means to benchmark and assess **quality** of data. Not a new topic, but much more can be done. Urban data/analytics scientists tend to take data for granted
- **Integration** of data across multiple scales
- Human-centric AI — data on people and personalised models

Urban Analytics Lab | Singapore

Connect with us

https://ual.sg @urbanalyticslab

We welcome collaborations and research visits in Singapore