



Review of embedding models based on vector geo data

Kamil Raczycki



Table Of Contents



01. Data definition

What data sources can be
used?



02. Embedding approach

How to embed the data?

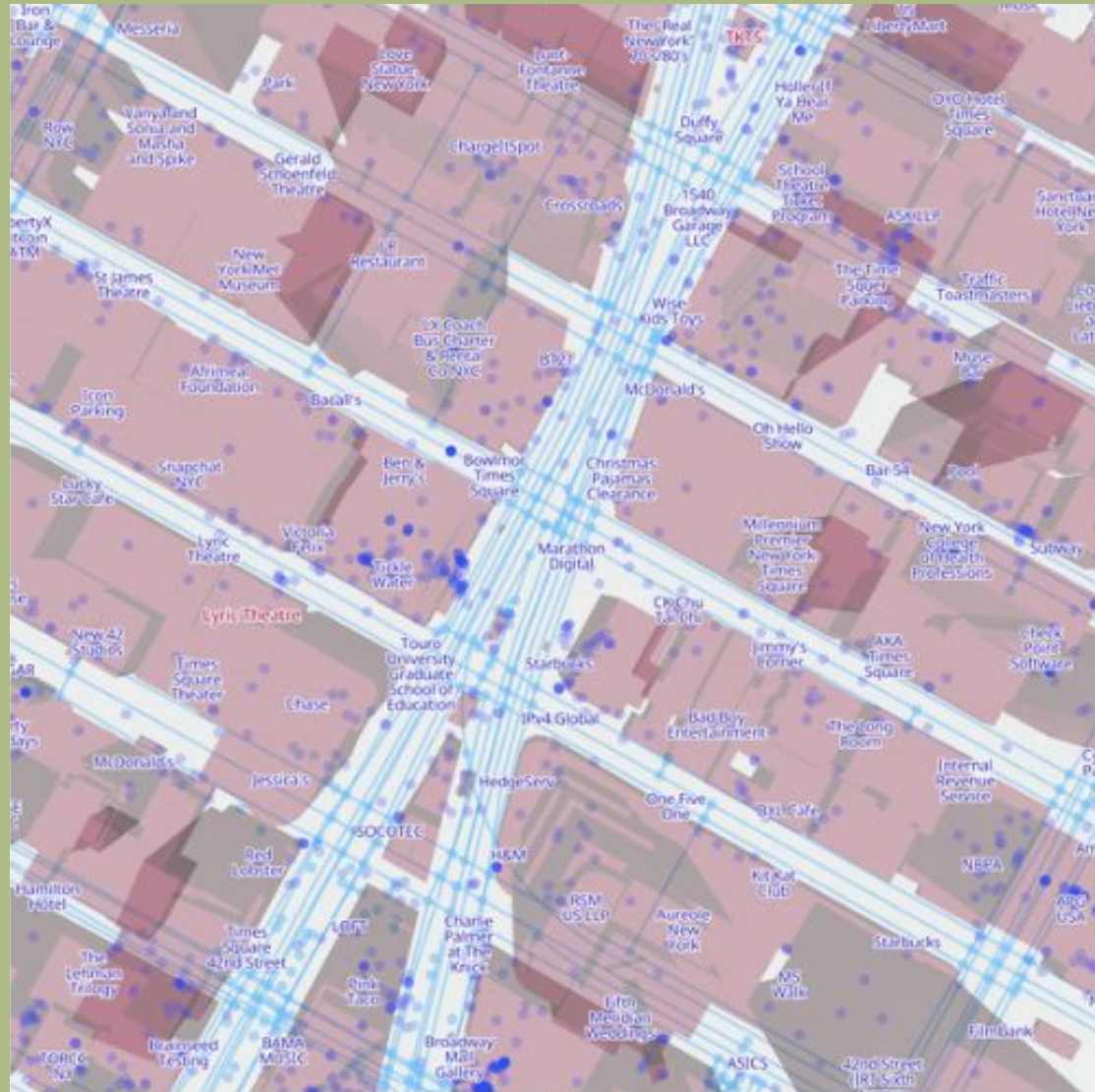


03. Available models

What models were created
and how do they work?



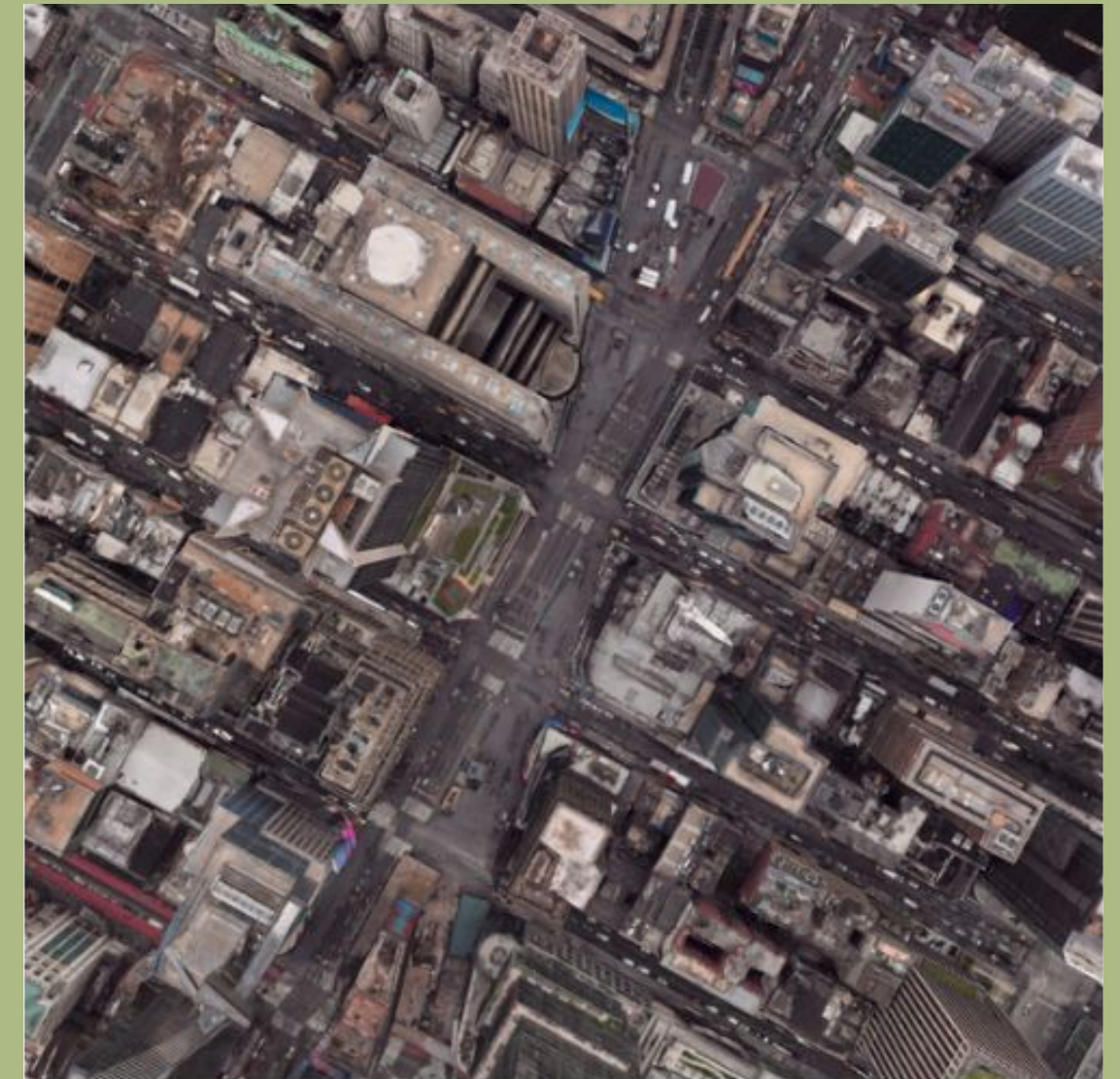
01. Data definition



Vector
detailed
geometries



Raster
pixels



Vector Geo Data

Types of data



Roads and Streets

Describing paths where people move



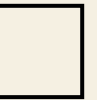
Buildings & Infrastructure

What is built around us?



Landmarks and Points of Interest

Traffic attractors



Mobility data

GTFS schedules,
GPS trajectories



Sources



OpenStreetMap

Public, crowdsourced,
unstructured



Overture Maps

Public, governed,
structured with schema



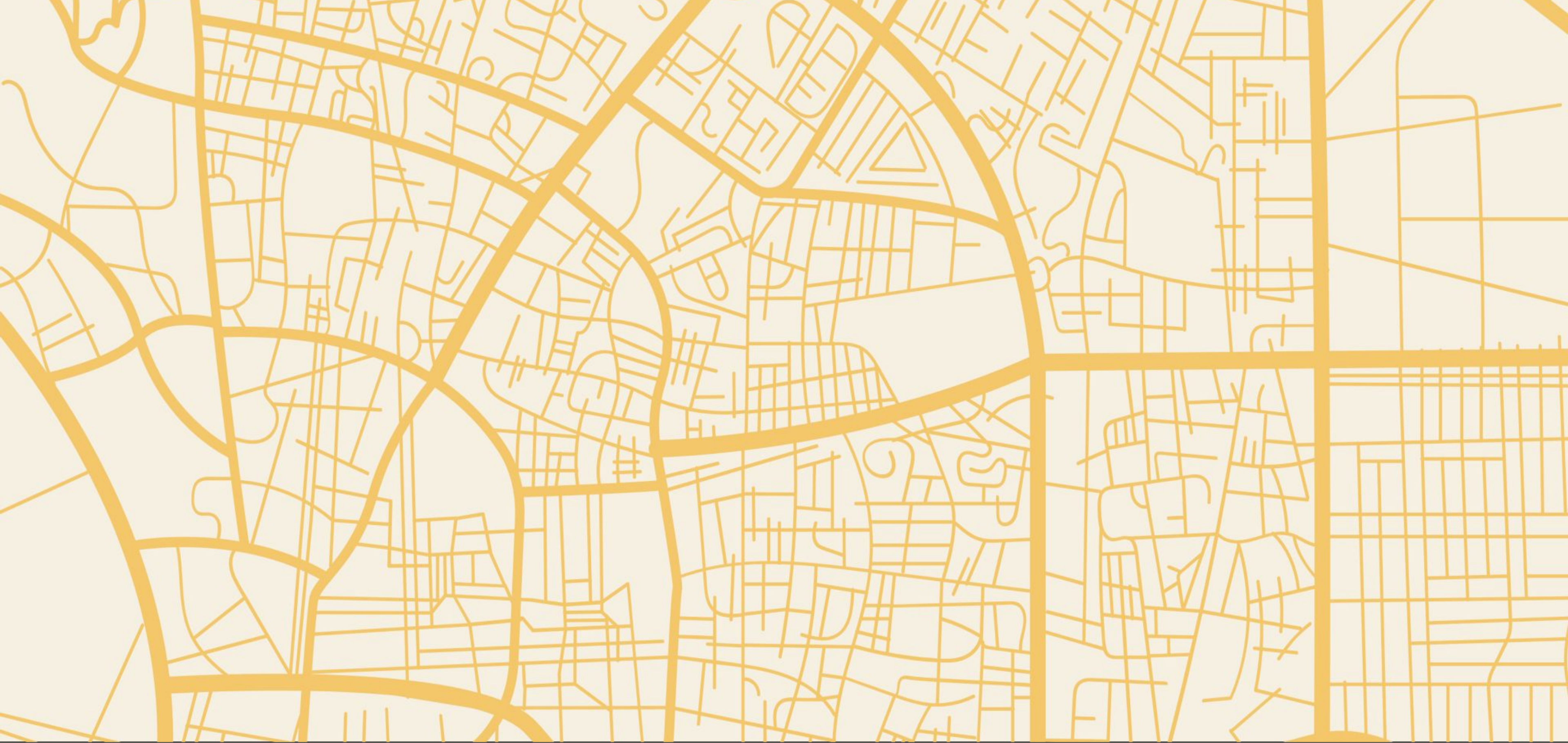
Public datasets

Wikipedia articles, open
data portals



Private datasets

GPS trajectories, locations
reviews, traffic & sales data



02. Embedding approach

How to embed the data?

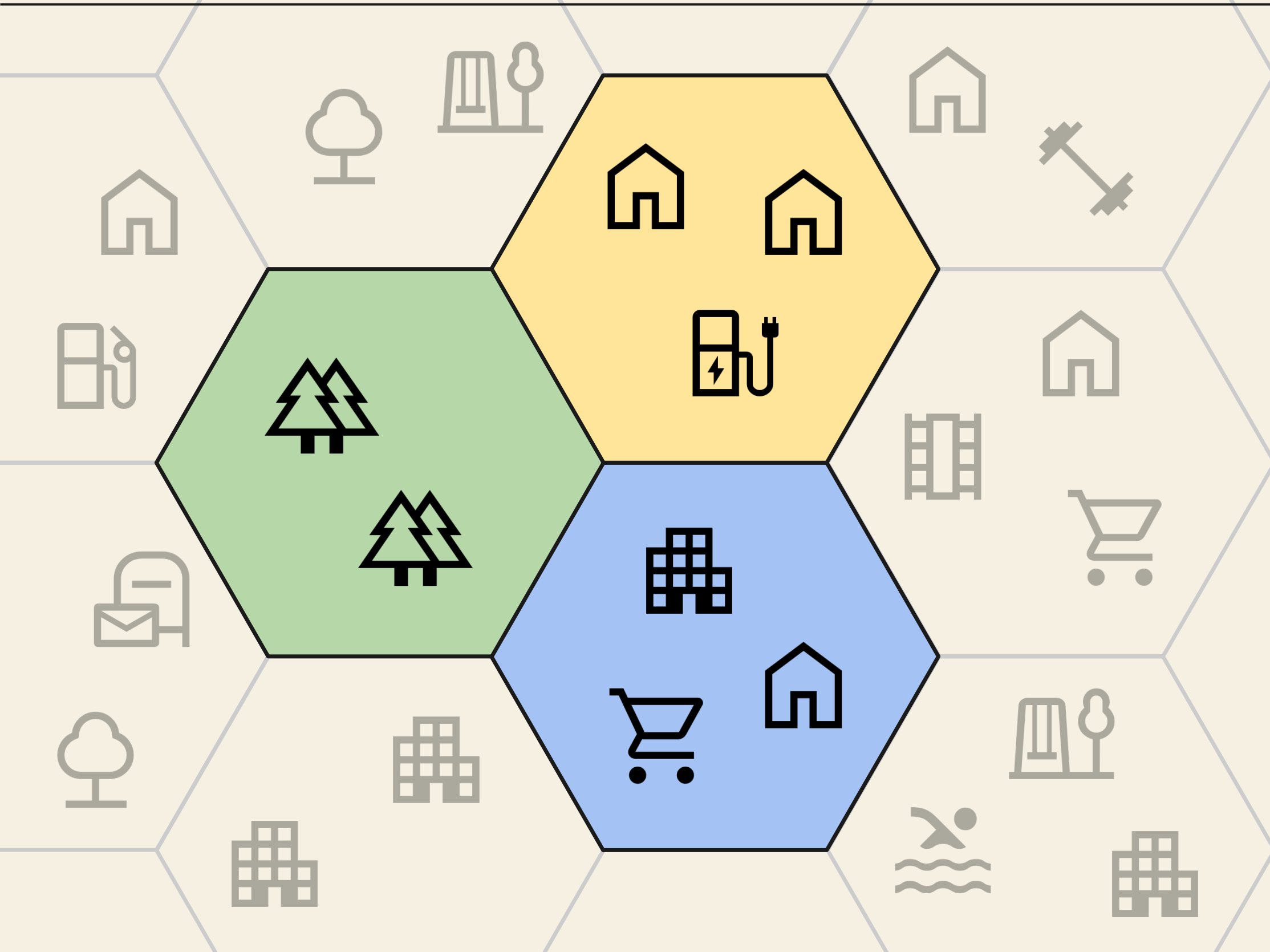


How to embed the data?

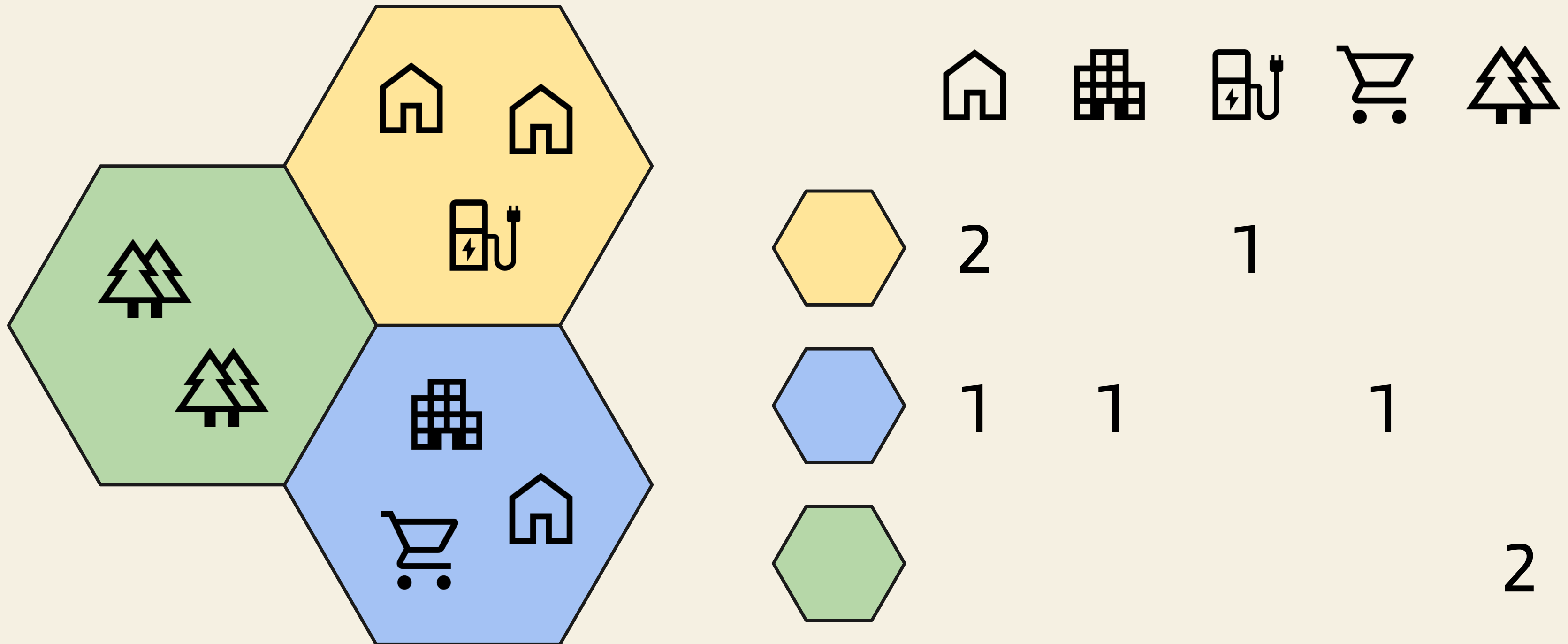


How to embed the data?

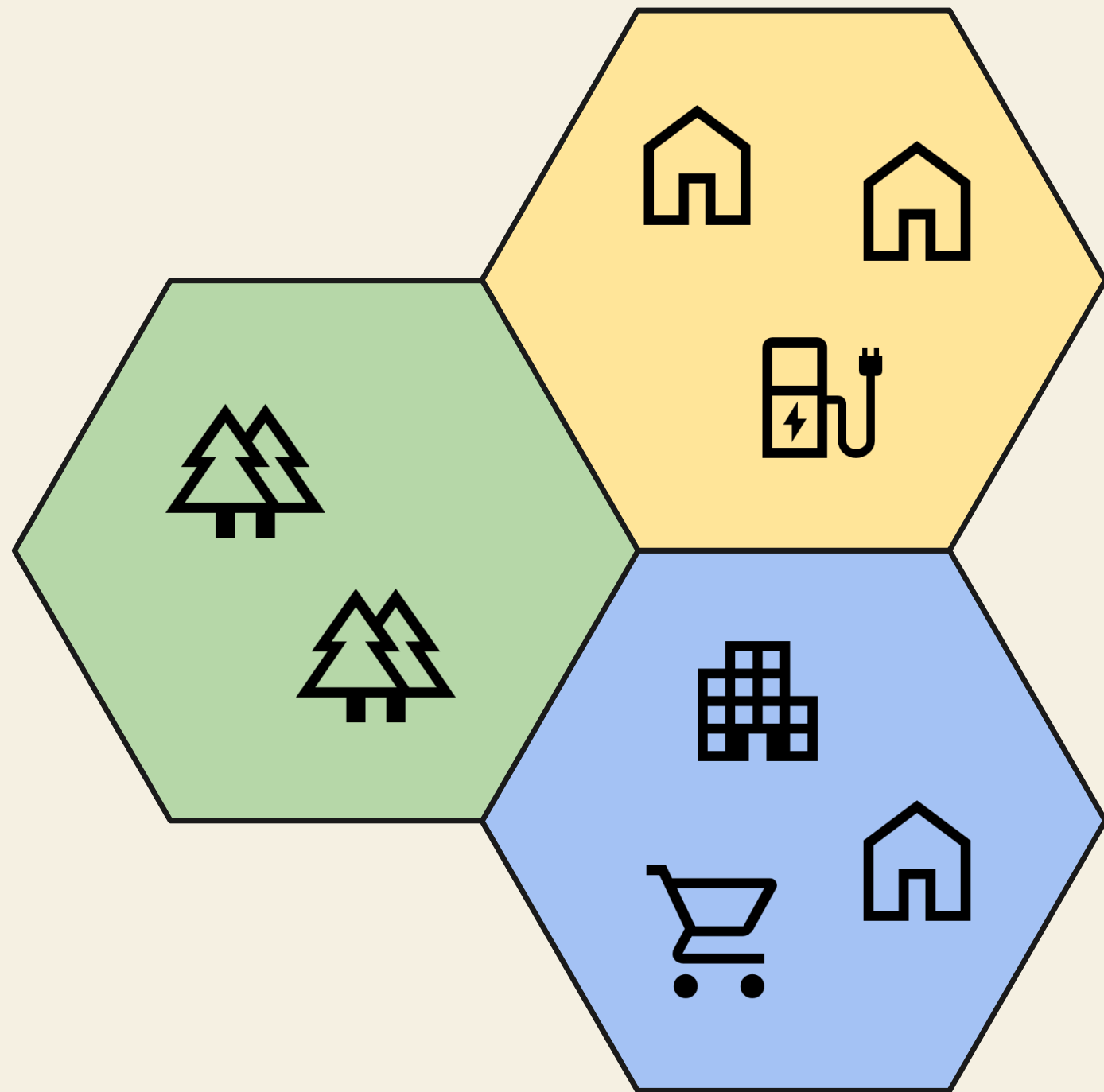
How to embed the data?

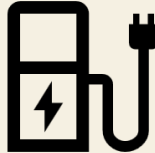
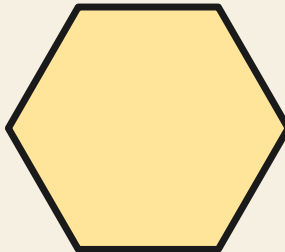
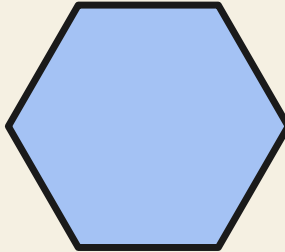
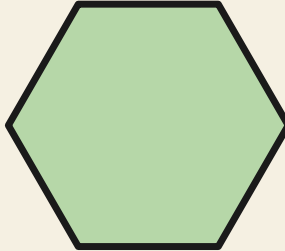


How to embed the data?

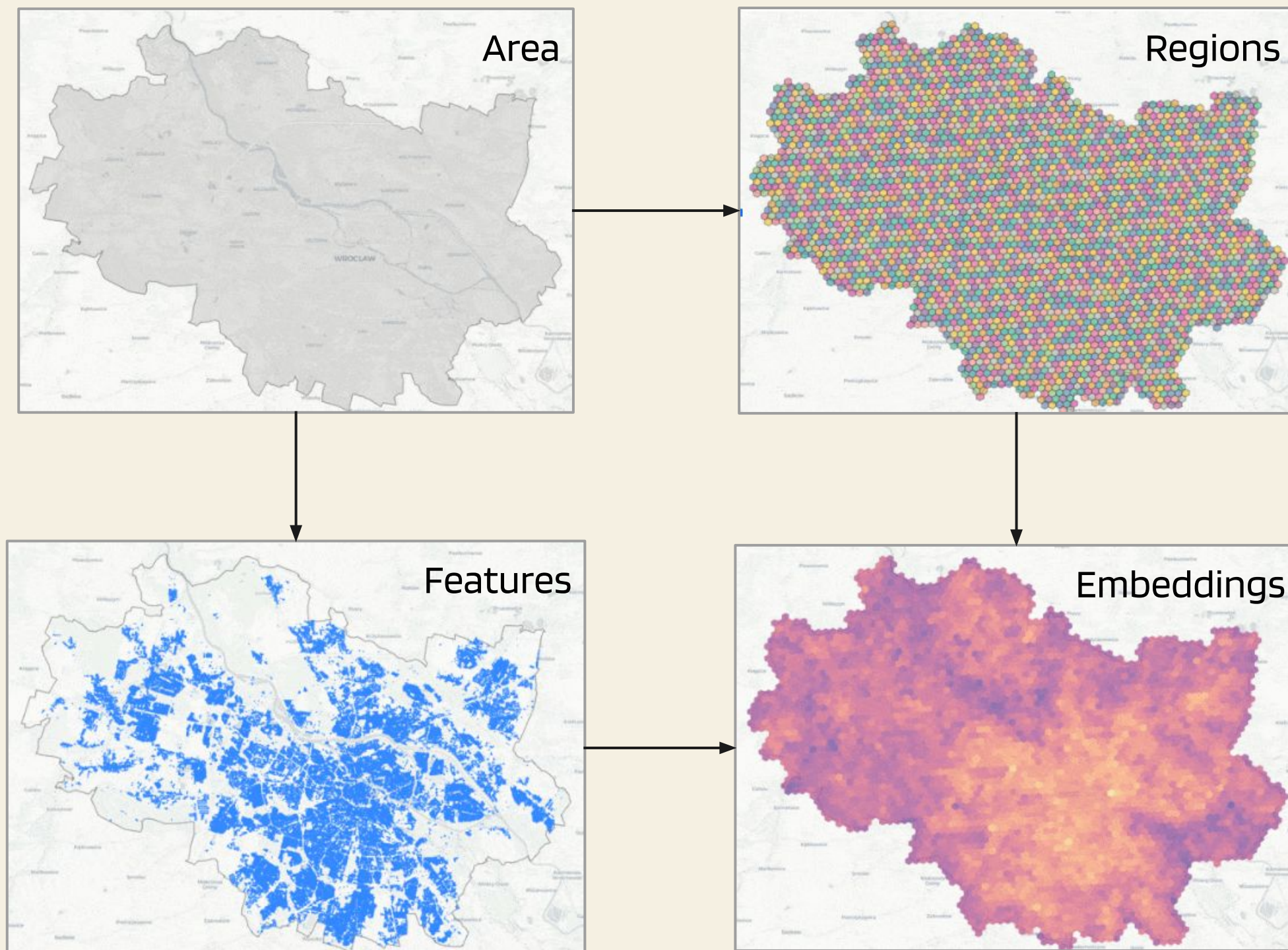


How to embed the data?



					
	2	0	1	0	0
	1	1	0	1	0
	0	0	0	0	2

How to embed the data?



srai

kraina-ai.github.io/srai
github.com/kraina-ai/srai

How to embed the data?

[CountEmbedder](#) & [ContextualCountEmbedder](#)



	leisure_adult_gaming_centre	leisure_playground	amenity_pub
region_id			
881e2040c3ffff	0	11	1
881e204081ffff	0	1	1
881e2040c7ffff	0	1	2
881e20408dffff	0	5	0
881e204089ffff	1	7	33
881e2040d7ffff	0	1	0
881e2040d1ffff	0	2	1
881e2040ddffff	0	0	0
881e2040d5ffff	0	12	10
881e20409dffff	0	4	0
881e20408bffff	0	11	3



03. Available models

Loc2Vec



Sentiance



2018



sentiance.com/loc2vec-learning-location-embeddings-w-triplet-loss-networks

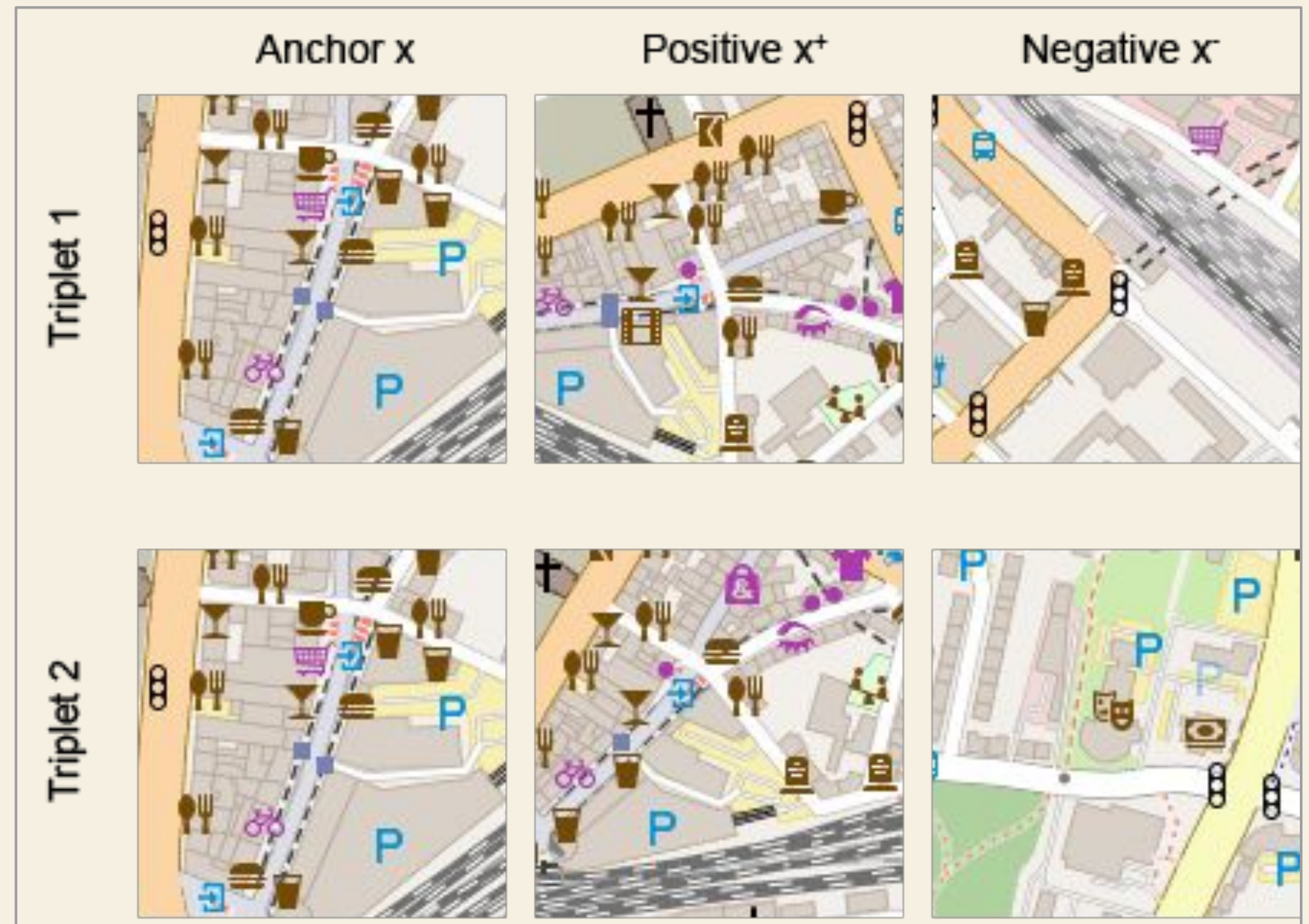
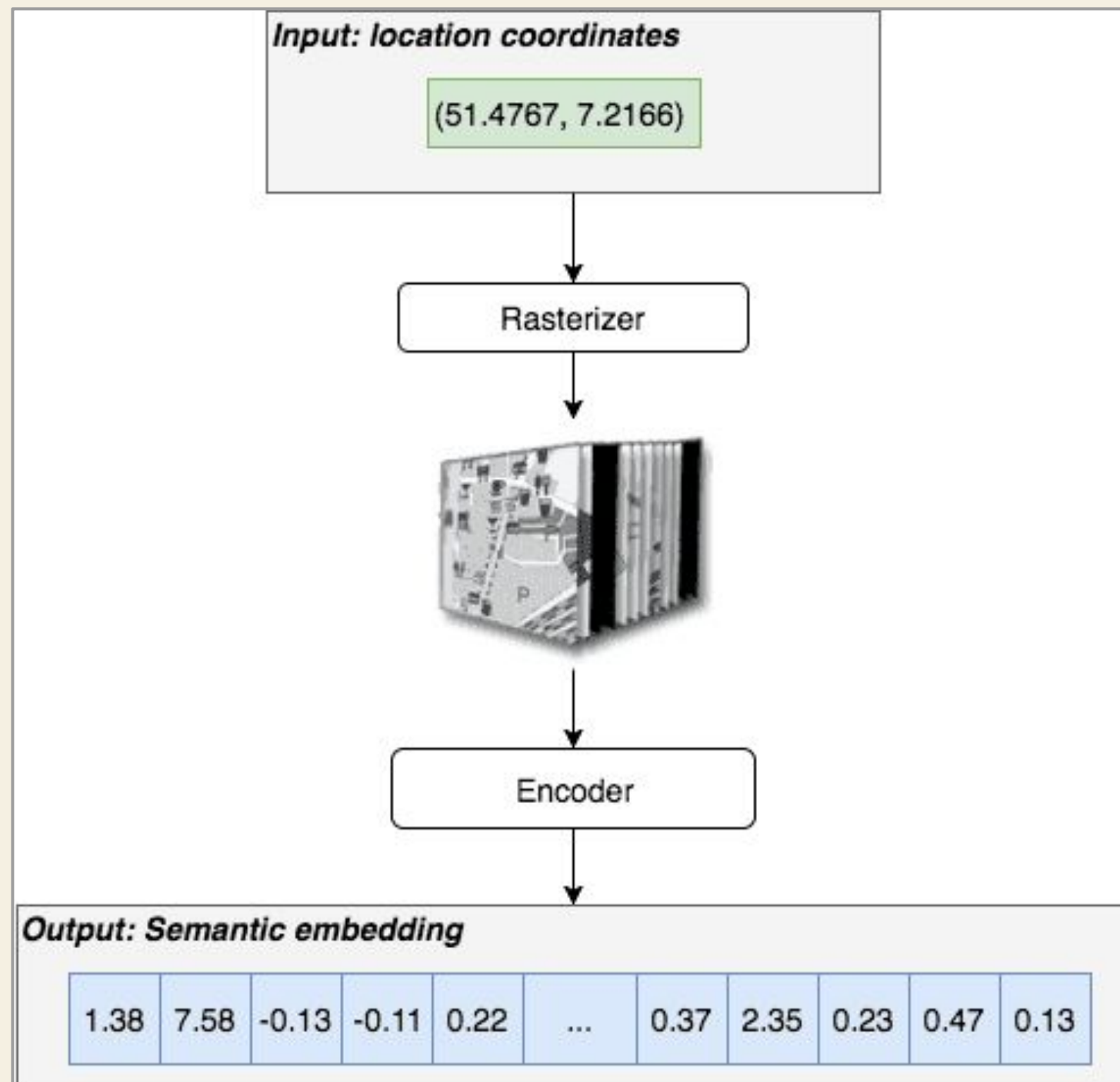


No official implementation

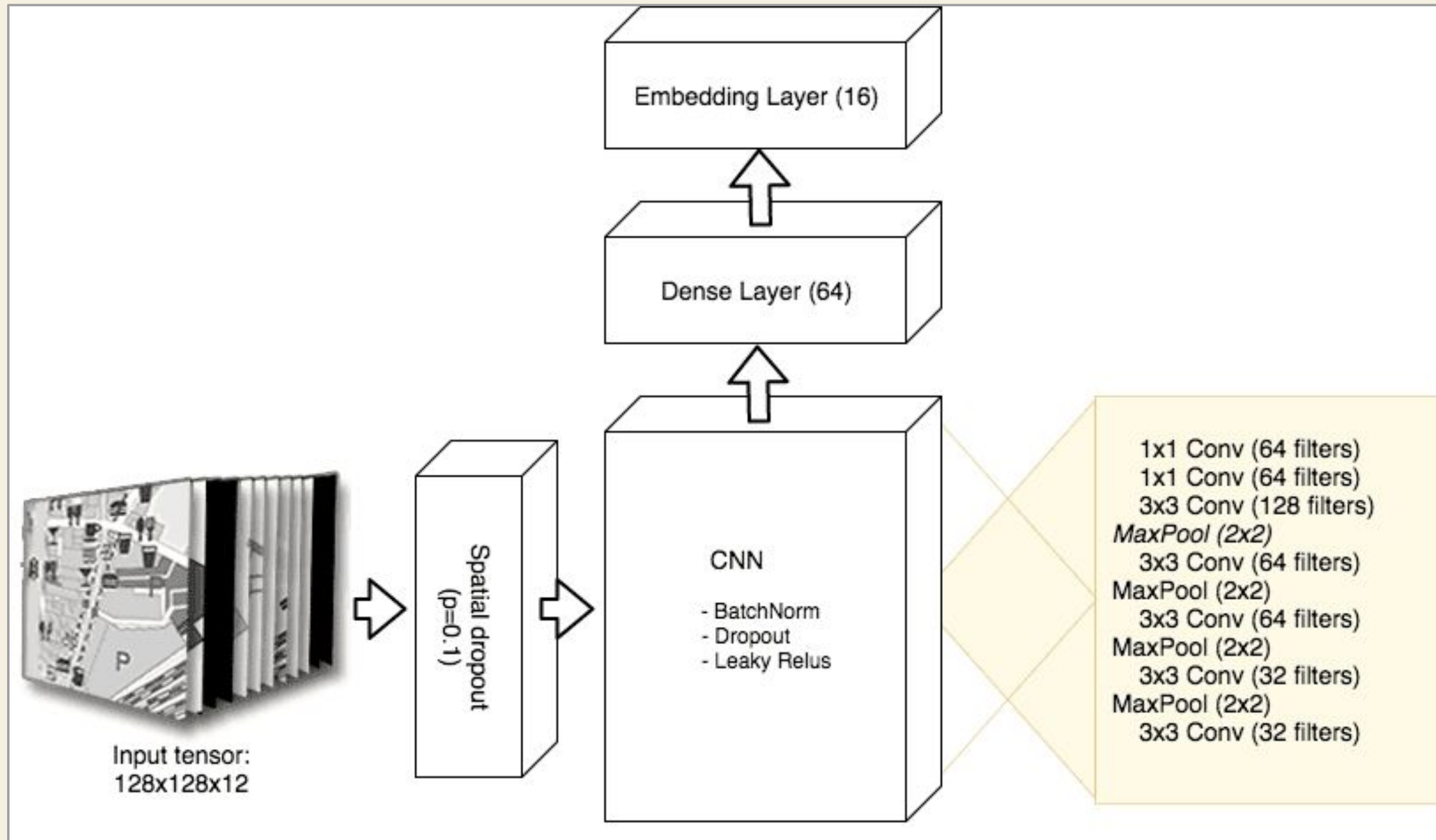
Open Source: [surya501](#), [angus-spence](#), [VasixG](#)



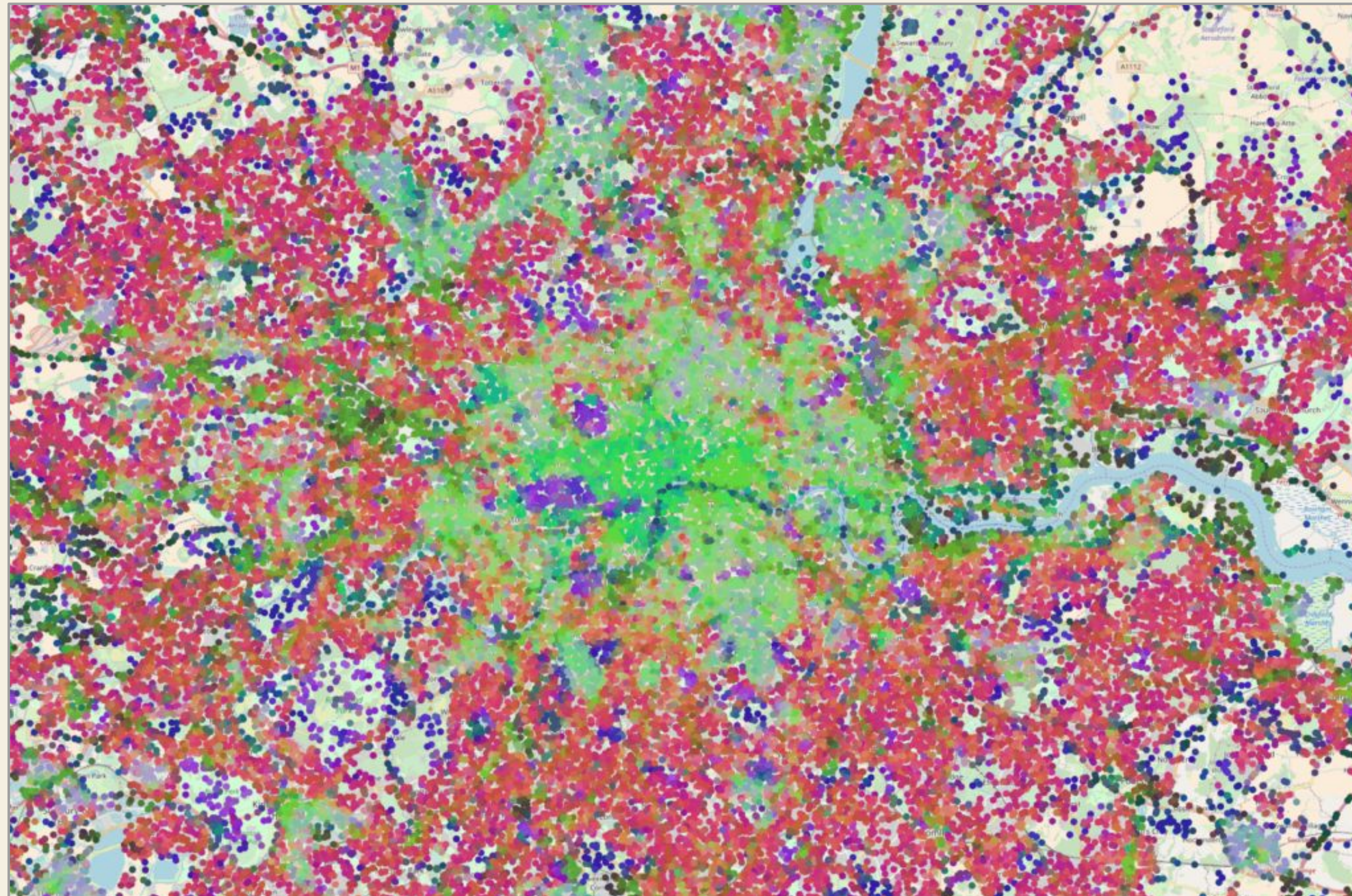
Loc2Vec



Loc2Vec



Loc2Vec



Hex2Vec



Szymon Woźniak, Piotr Szymański



2021



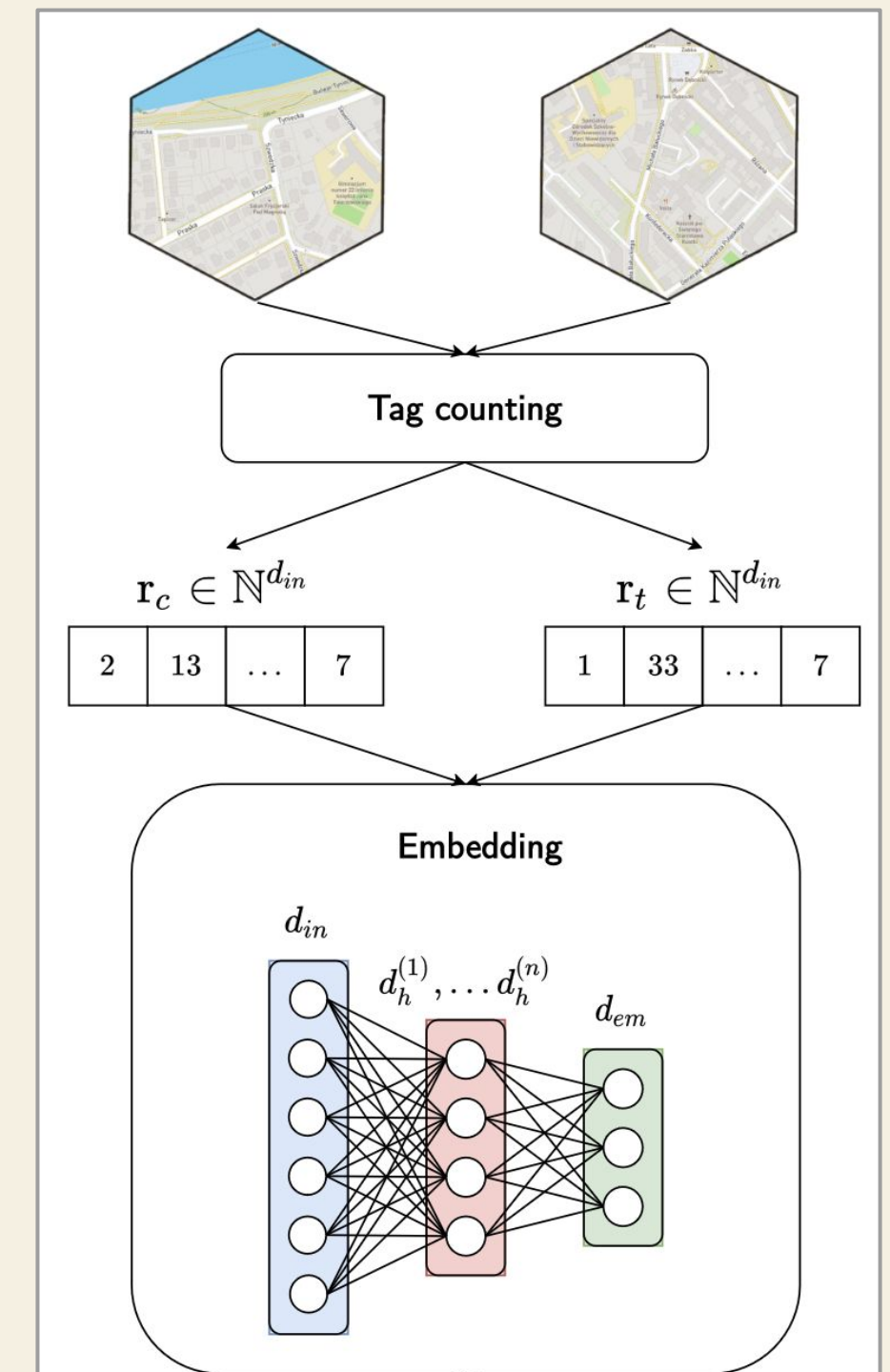
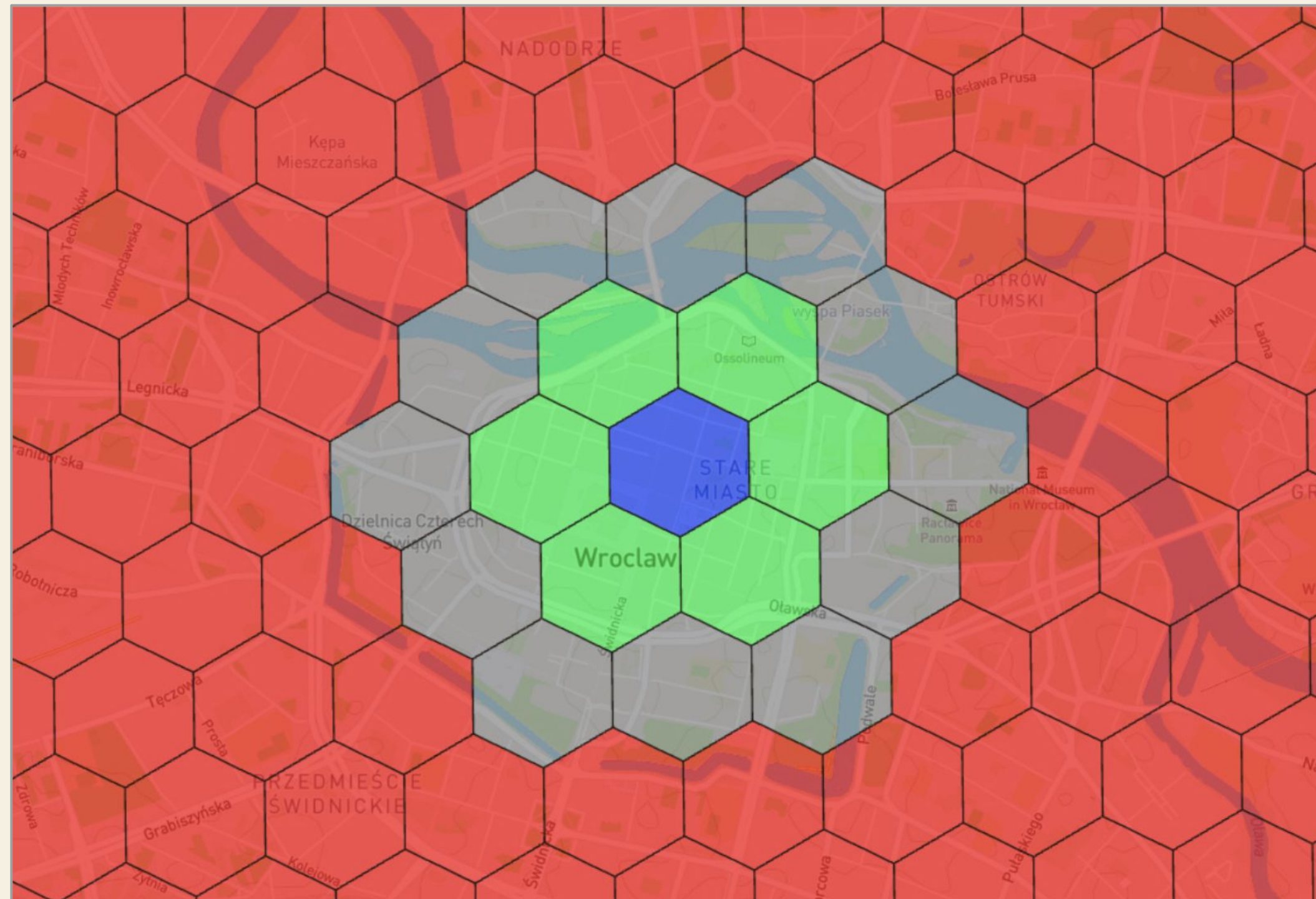
DOI: [10.1145/3486635.3491076](https://doi.org/10.1145/3486635.3491076)



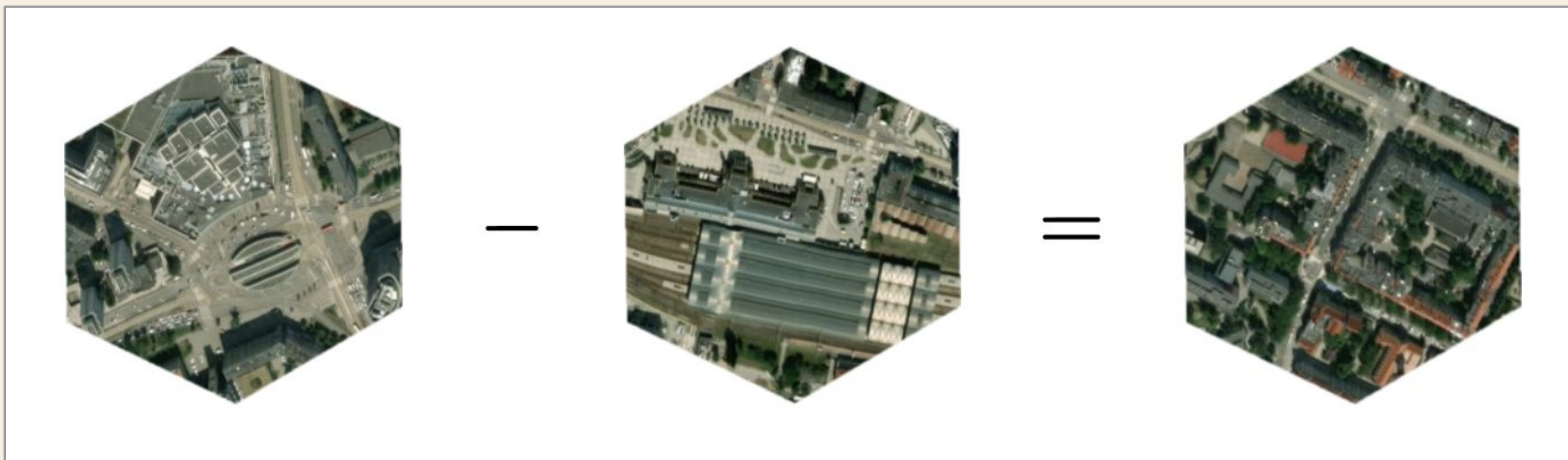
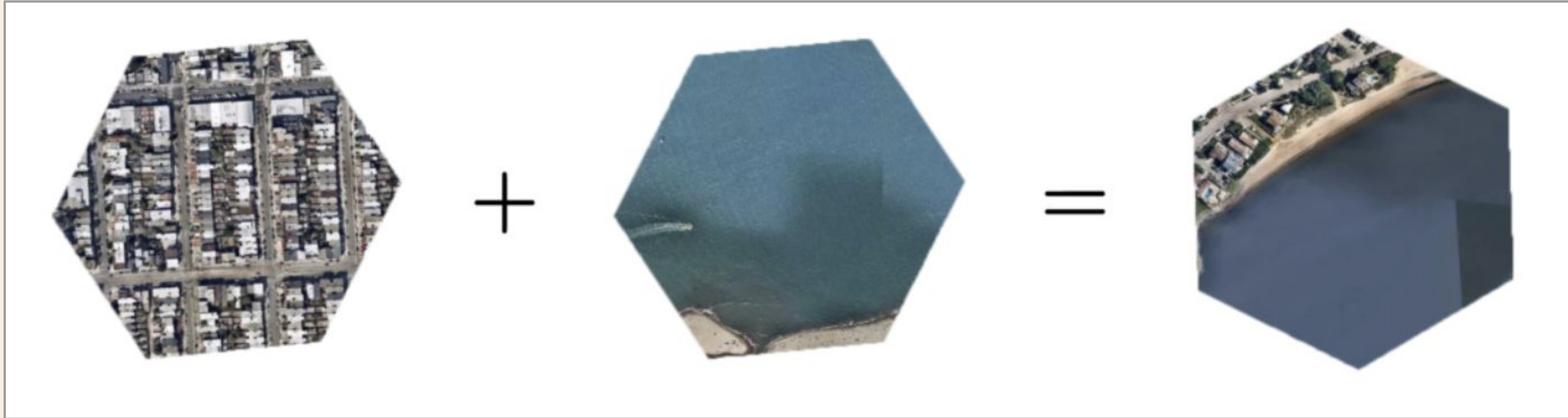
Official implementation: [Hex2VecEmbedder](#) in [SRAI](#)



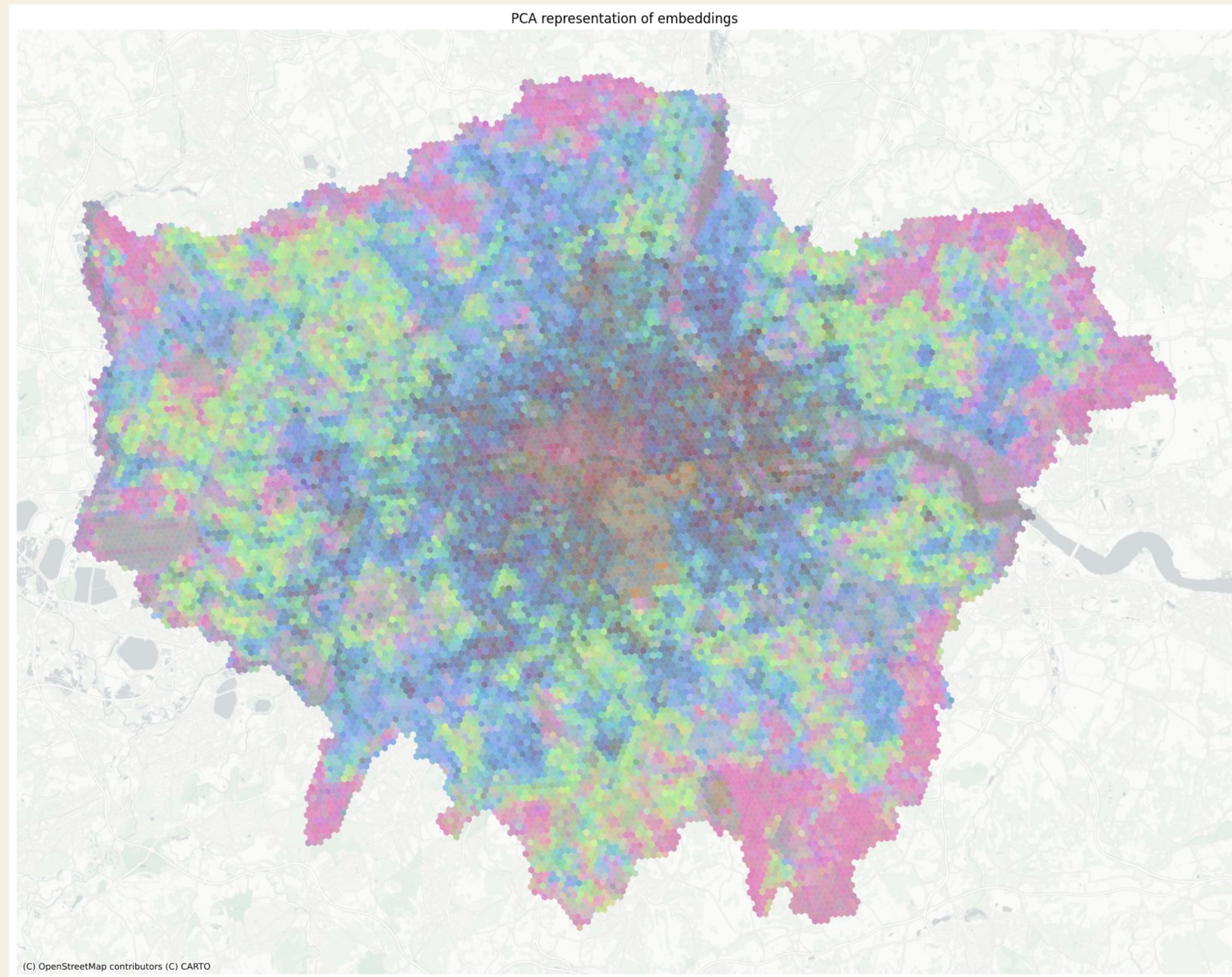
Hex2Vec



Hex2Vec



Hex2Vec



GeoVex



Daniele Donghi, Anne Morvan - Expedia Group



2023



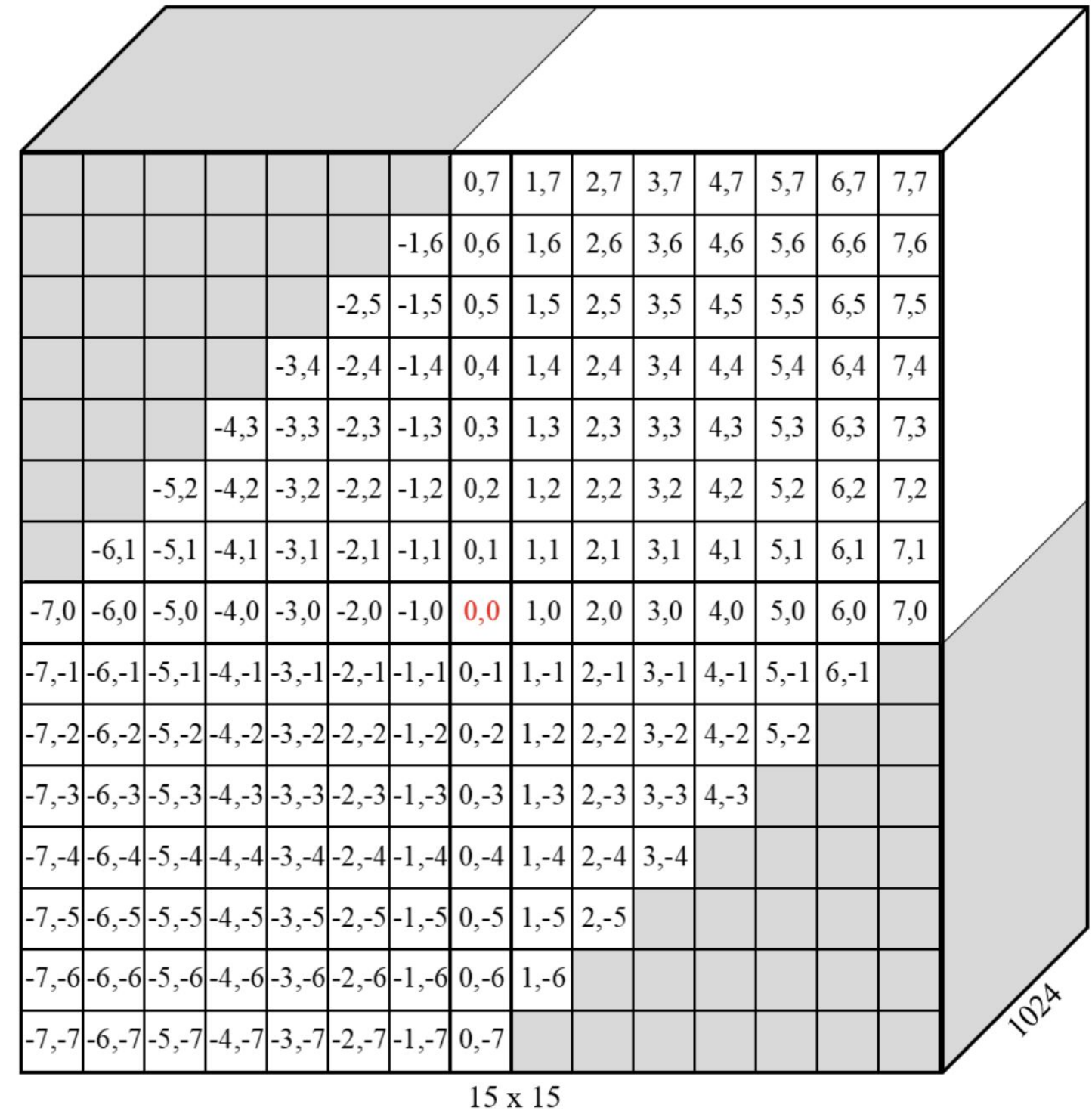
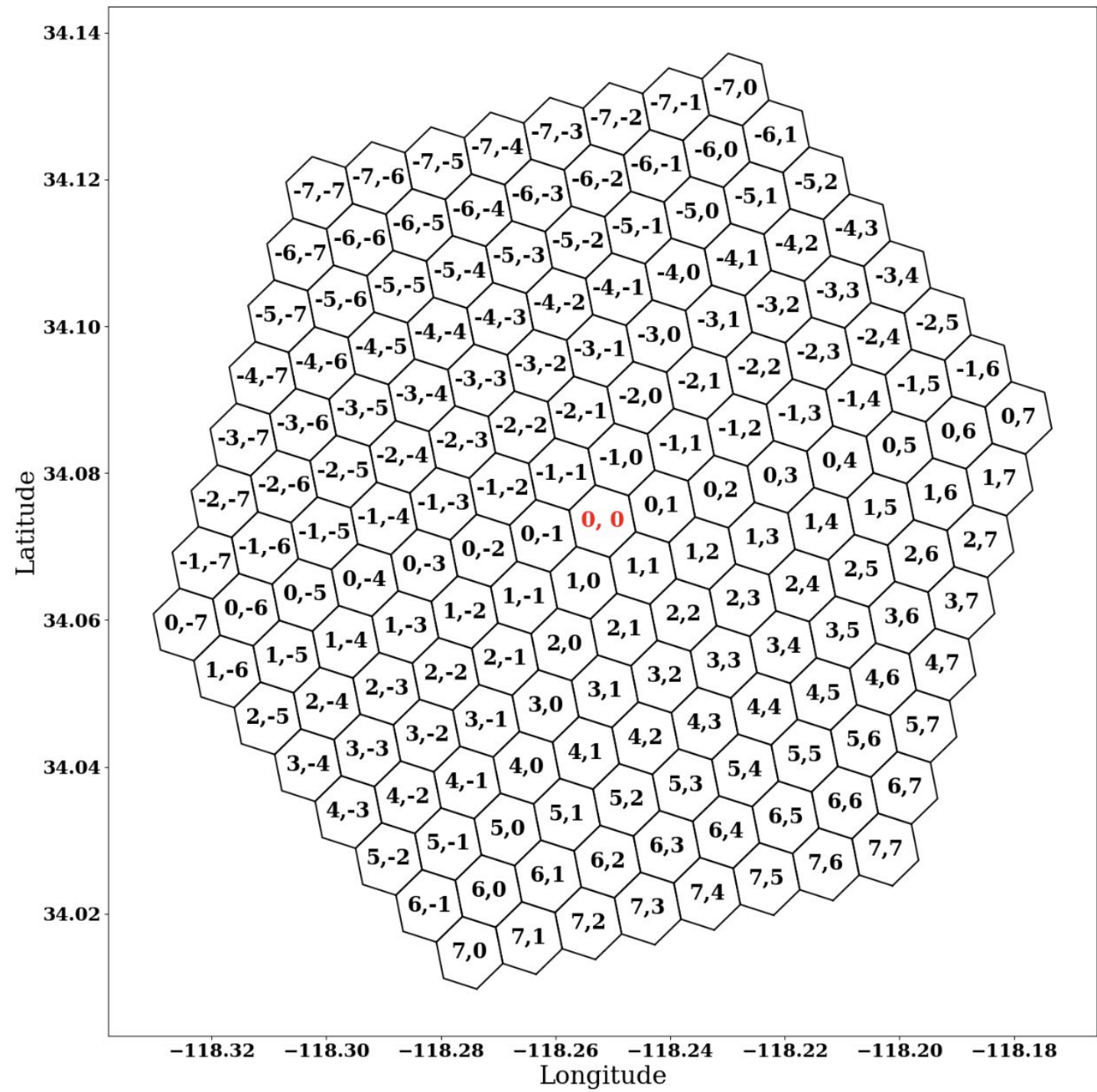
DOI: [10.1145/3615886.3627750](https://doi.org/10.1145/3615886.3627750)



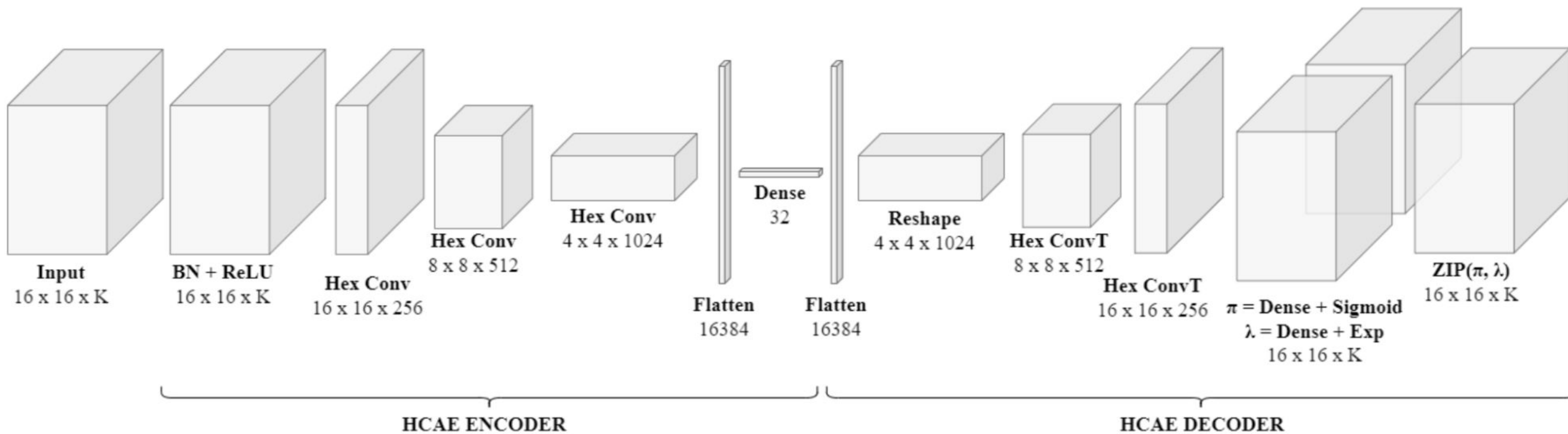
No official implementation

Open Source: [GeoVexEmbedder](#) in [SRAI](#) by [Maxwell Schrader](#)

GeoVex

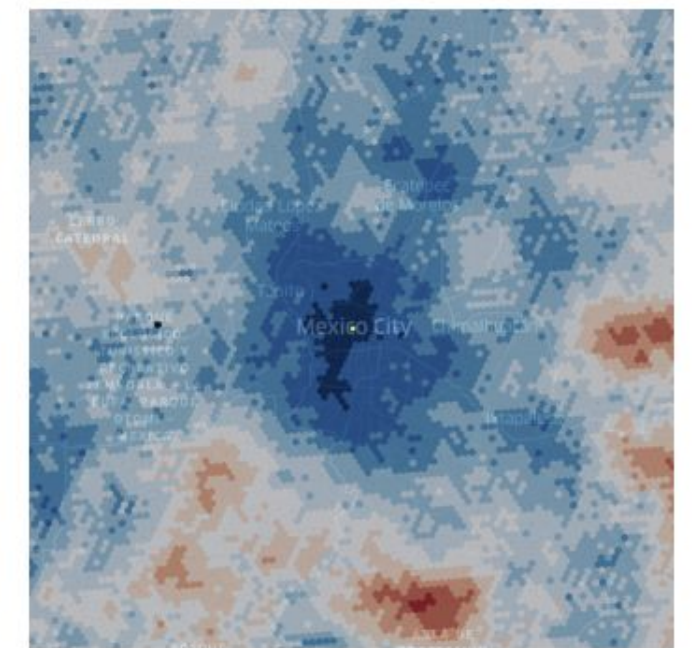
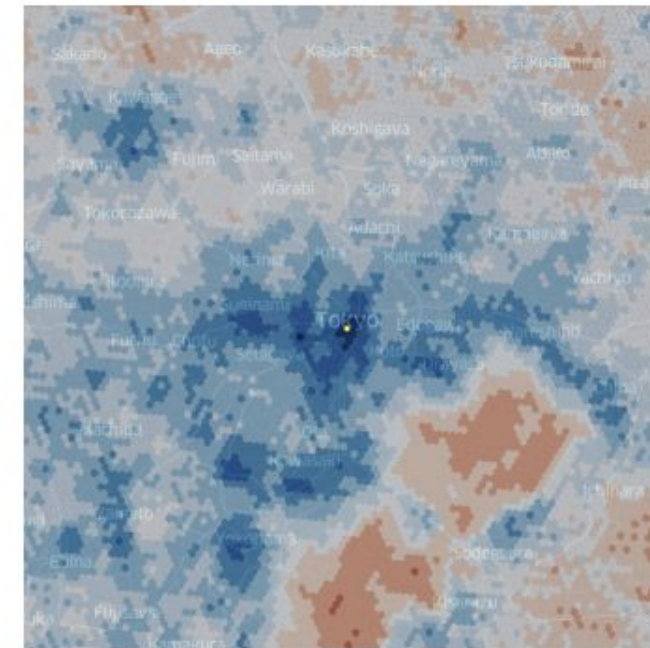
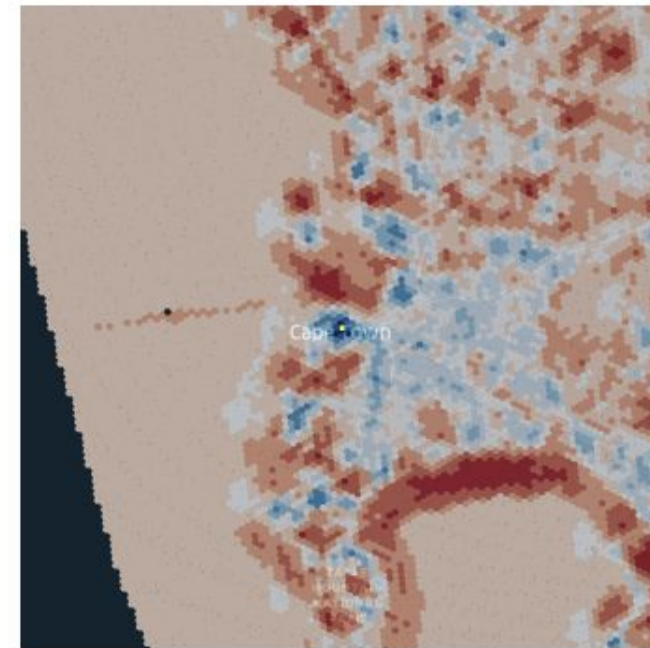
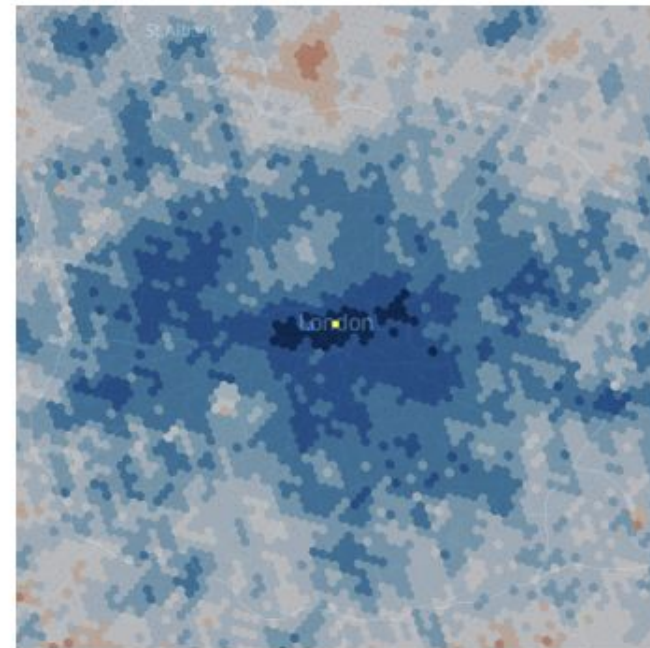
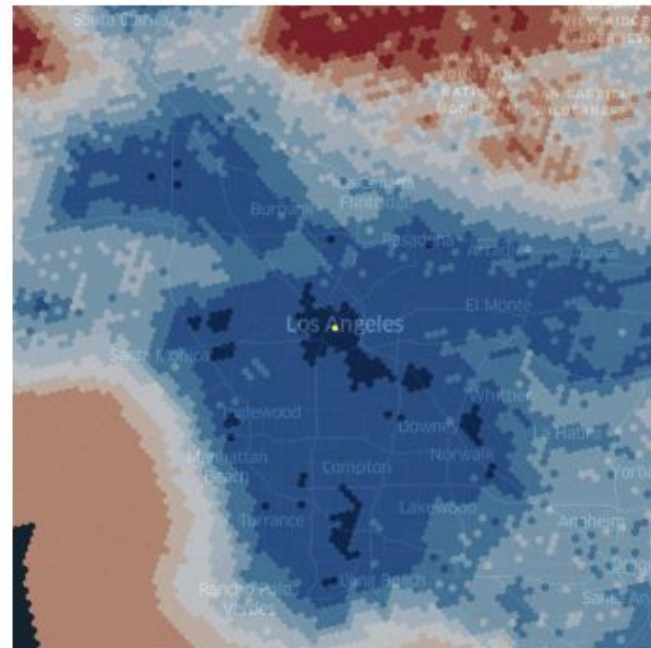


GeoVex

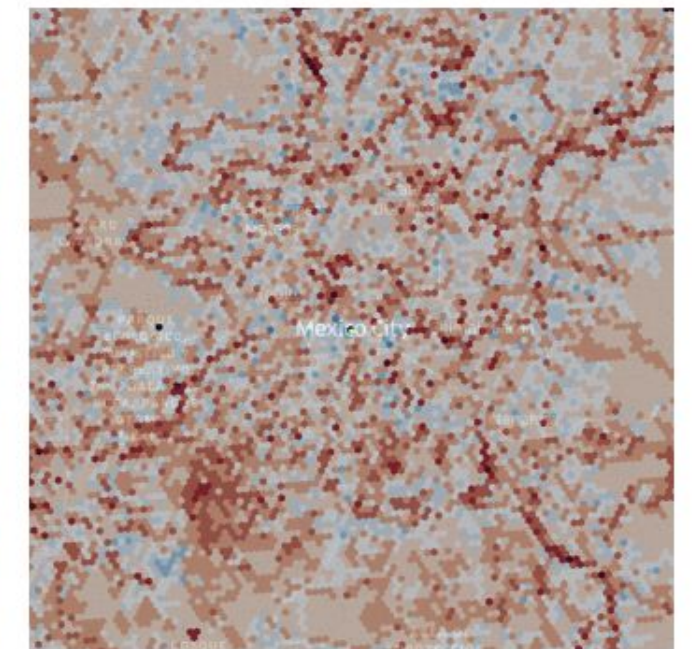
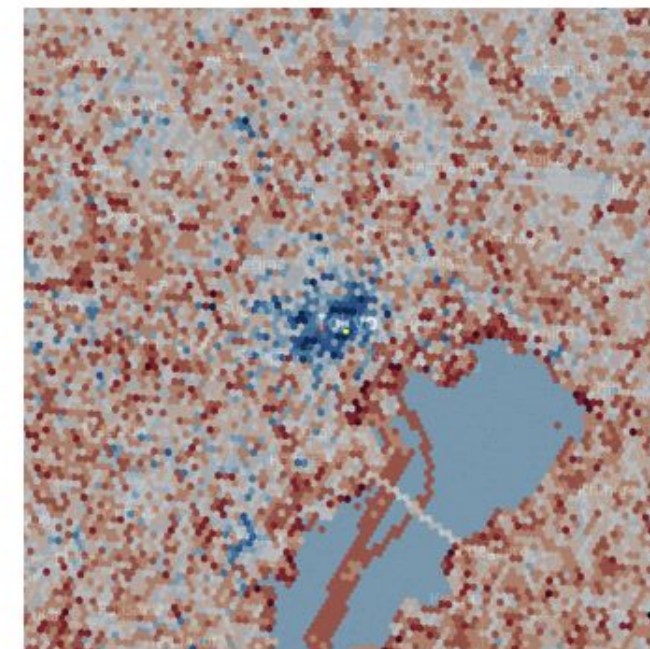
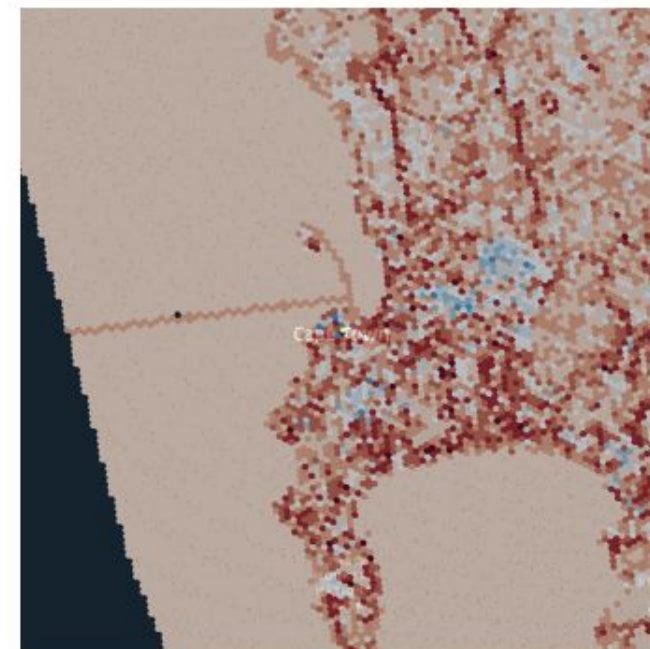
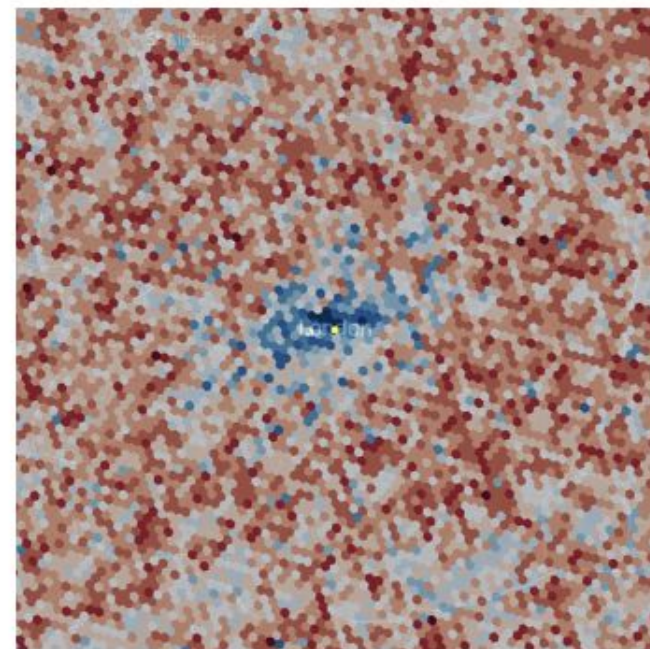
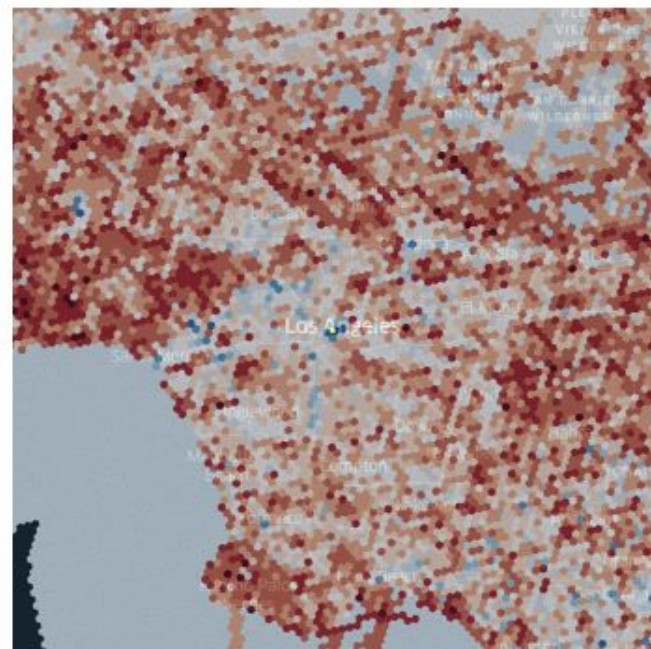


GeoVex

GeoVex



Hex2Vec



(a) Los Angeles

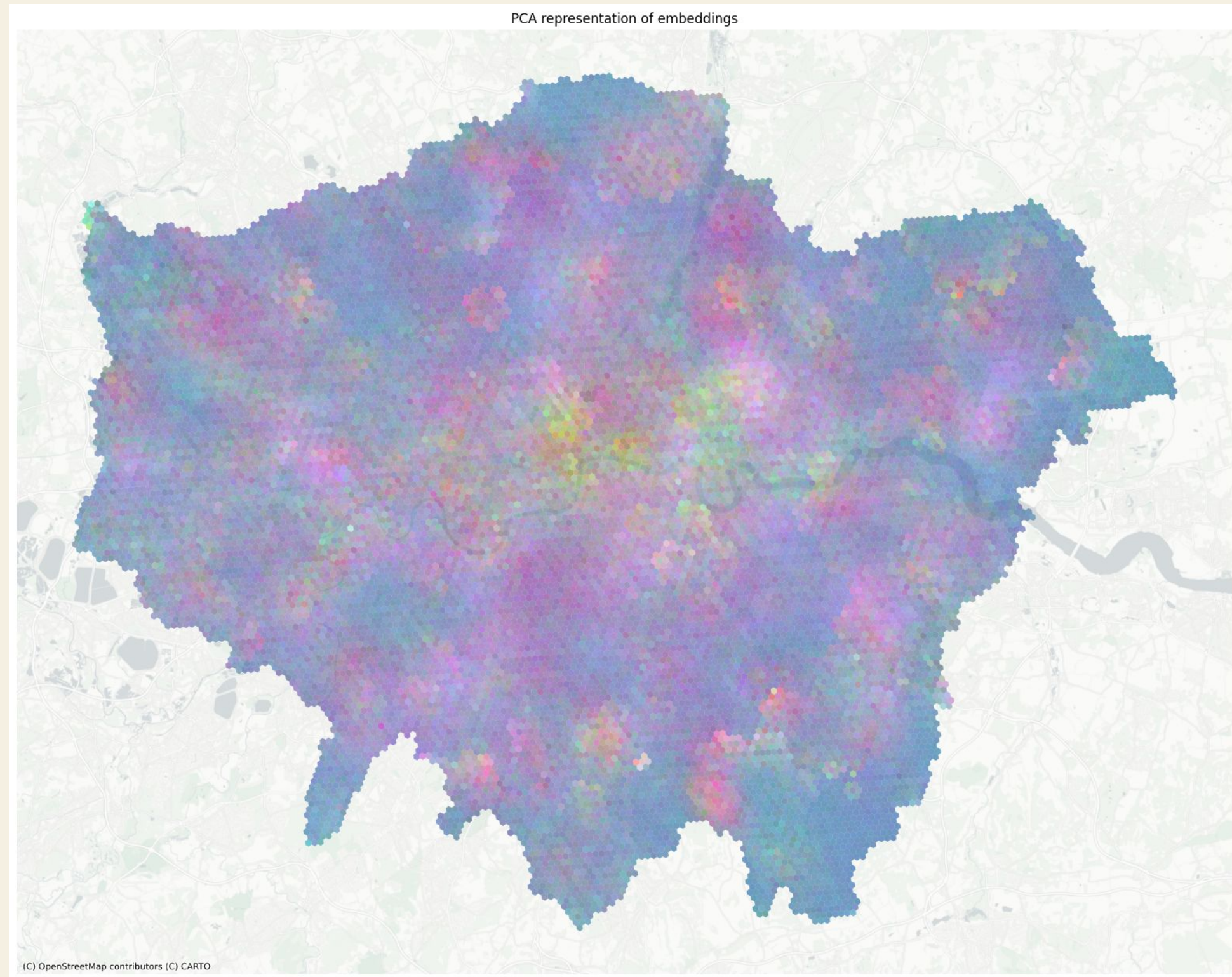
(b) London

(c) Cape Town

(d) Tokyo

(e) Mexico City

GeoVex



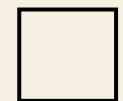
S2Vec



Shushman Choudhury et al - Google



2025

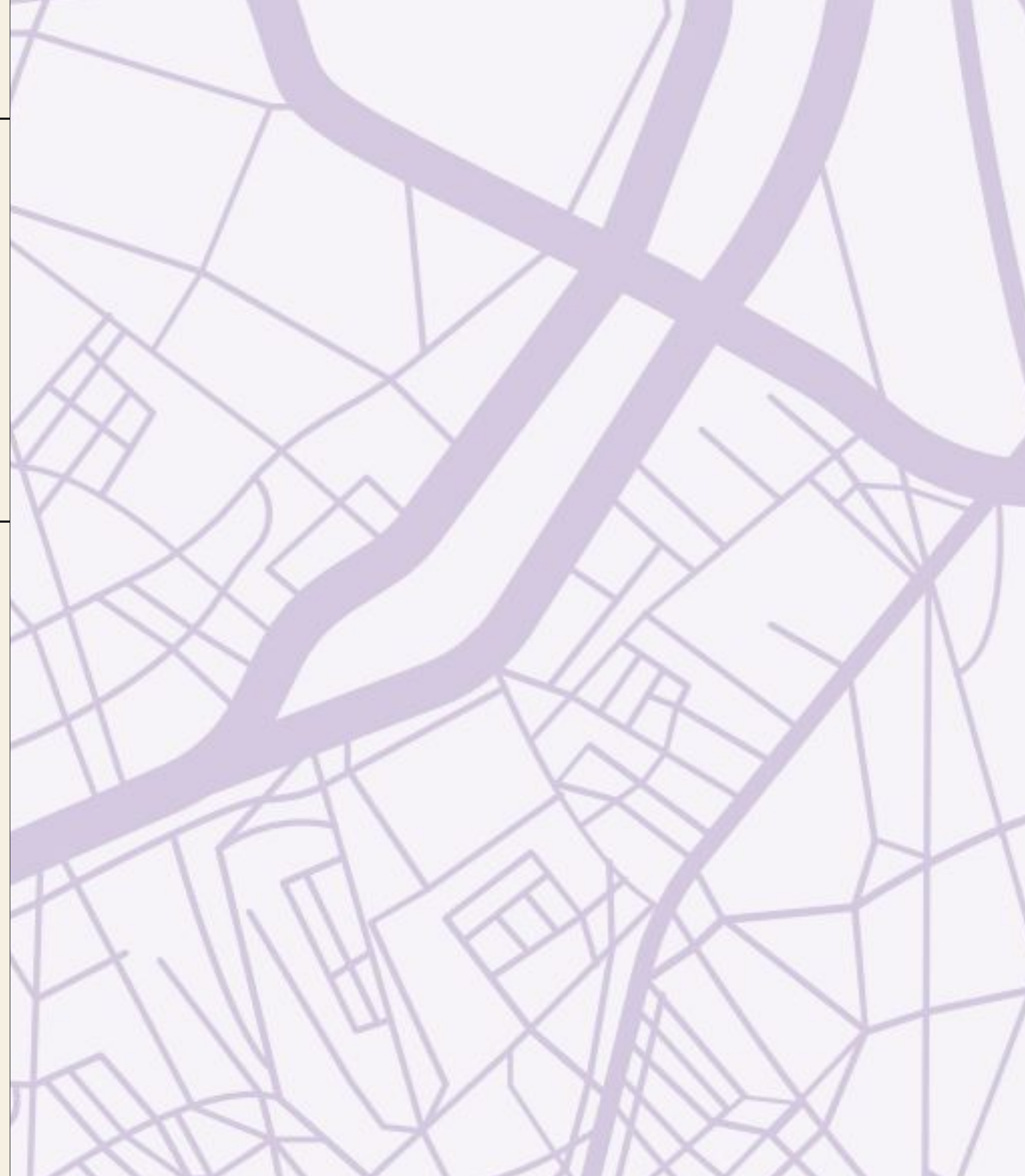


DOI: [10.1145/3787217](https://doi.org/10.1145/3787217)

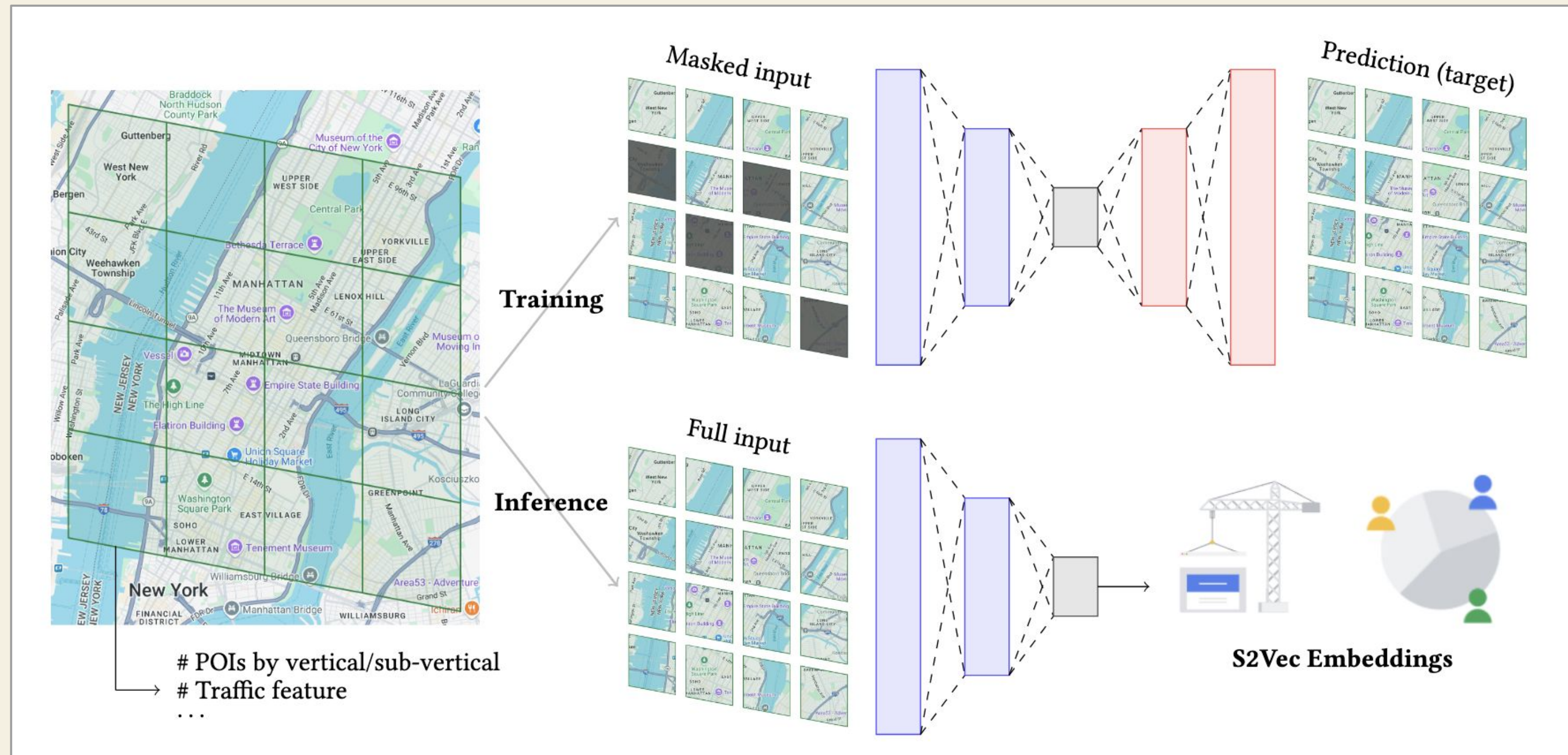


No official implementation

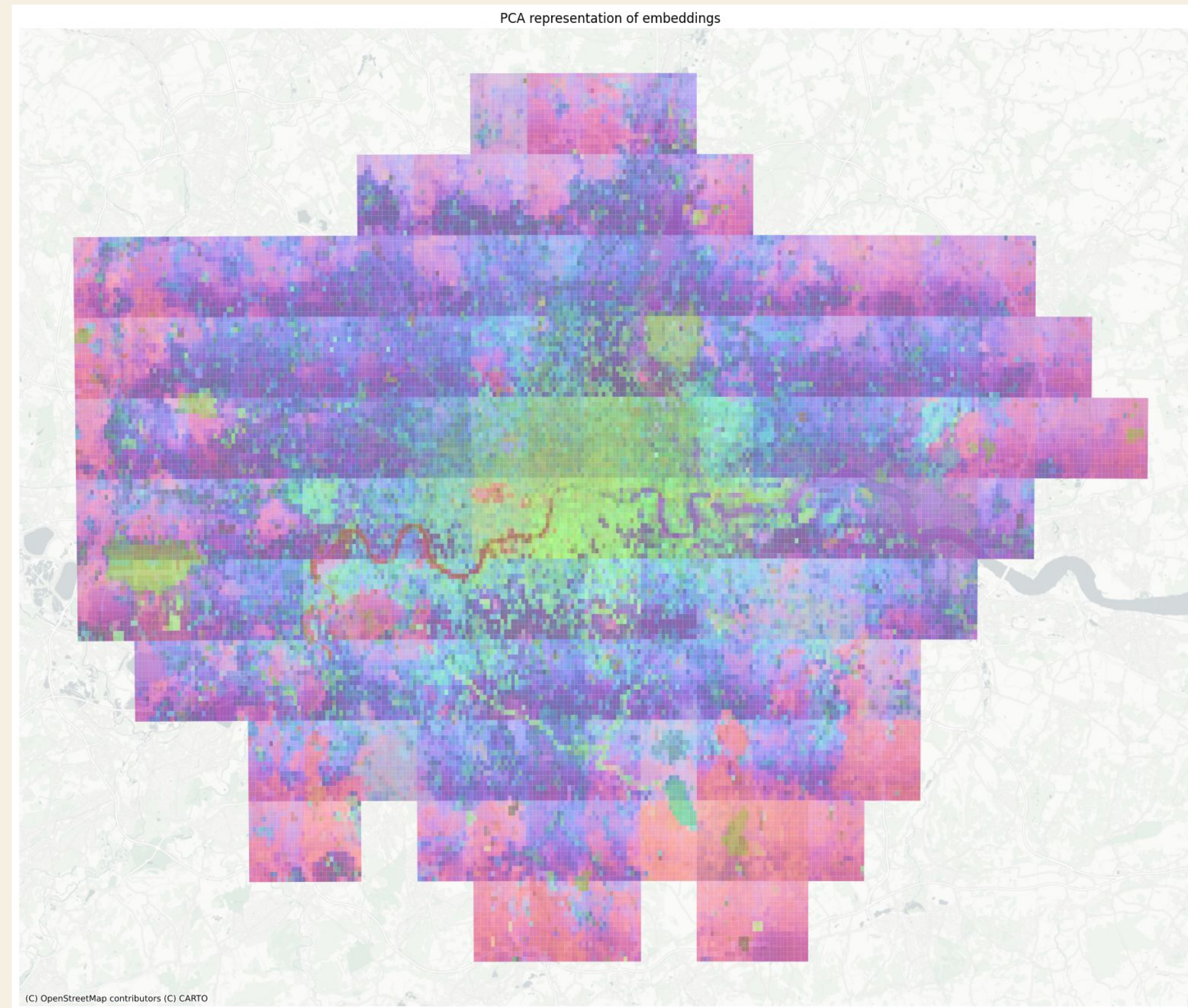
Open Source: [S2VecEmbedder](#) in [SRAI](#) by [Hubert Kriebitzsch](#)



S2Vec



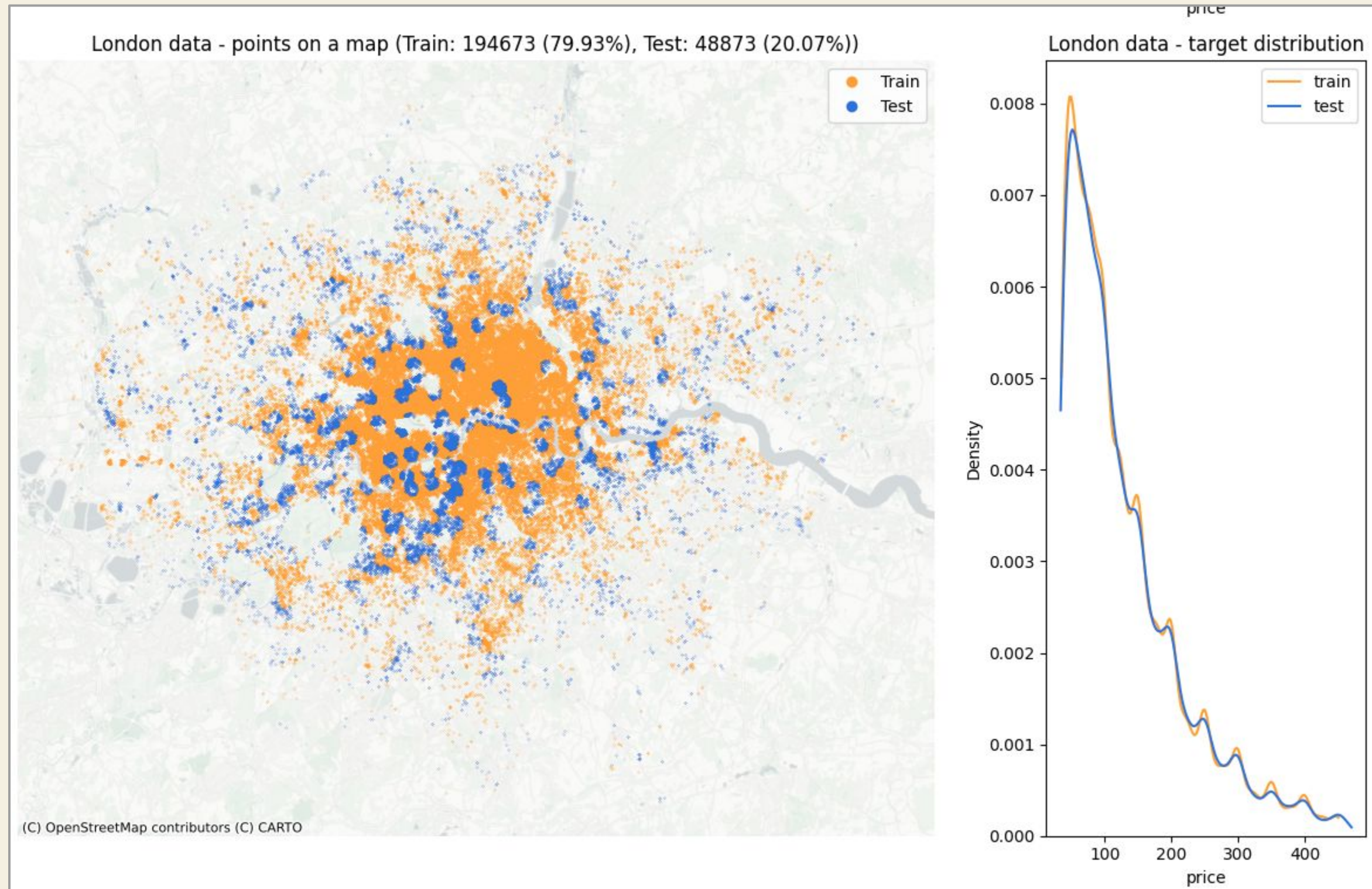
S2Vec



How to check if
those embeddings
are useful?



OBSR: Open Benchmark for Spatial Representations



OBSR: Open Benchmark for Spatial Representations

Table 2: Performance of different embedders on STRPP task on Airbnb dataset at various resolutions. (H2V - Hex2Vec, GV - GeoVex, CE - CountEmbedder, CCE - ContextualCountEmbedder)

Metric	H2V	Res 8			H2V	Res 9			H2V	Res 10		
		GV	CCE	CE		GV	CCE	CE		GV	CCE	CE
MSE $\times 10^2$	40.10	48.33	40.21	38.30	45.97	50.57	45.65	42.53	59.48	63.80	44.96	66.27
RMSE	63.33	69.53	63.41	61.89	67.80	71.11	67.50	65.22	77.12	79.87	81.83	81.40
MAE	43.03	48.58	39.69	39.61	46.97	50.75	42.98	43.21	54.55	56.42	56.96	57.08
MAPE	51.38	47.47	57.26	58.16	54.17	53.90	60.09	60.90	53.02	51.76	53.53	54.48
sMAPE	44.49	41.39	48.09	47.88	45.85	44.67	48.67	49.05	46.76	46.76	49.13	49.03

Table 3: Performance of different embedders on HPP task on House Sales in King County dataset at various resolutions. (H2V - Hex2Vec, GV - GeoVex, CE - CountEmbedder, CCE - ContextualCountEmbedder)

Metric	H2V	Res 8			H2V	Res 9			H2V	Res 10		
		GV	CCE	CE		GV	CCE	CE		GV	CCE	CE
MSE $\times 10^9$	27.51	26.98	31.69	58.06	23.18	26.82	28.16	26.83	25.28	28.19	22.22	25.22
RMSE $\times 10^4$	16.577	16.42	17.80	24.09	15.22	16.37	16.77	16.37	15.90	16.78	14.90	15.83
MAE $\times 10^4$	12.62	13.41	12.95	18.10	11.90	13.38	12.53	12.69	12.56	13.23	11.83	12.62
MAPE	33.91	30.42	40.26	43.35	33.28	31.64	37.51	33.75	32.22	31.40	34.64	32.68
sMAPE	33.73	29.23	38.95	46.99	31.67	29.81	36.25	32.22	30.57	30.69	31.81	30.68



Future Directions

- Combine modalities - **secured financial grant**
- Utilize traffic / social context
- Incorporate historic changes in the data
- Prepare more benchmarks for urban-related tasks

Thank you for listening



Kamil Raczycki

[linkedin.com/in/raczyckikamil](https://www.linkedin.com/in/raczyckikamil)

github.com/RaczeQ



Piotr Szymański

[linkedin.com/in/niedakh](https://www.linkedin.com/in/niedakh)

piotr.szymanski@pwr.edu.pl

This presentation template is free for everyone to use
thanks to the following:



for the presentation template

Pexels, Pixabay

for the photos

Geospatial Foundation Models Workshop, 2026-02-04

Thank you for listening



Kamil Raczycki

[linkedin.com/in/raczyckikamil](https://www.linkedin.com/in/raczyckikamil)

github.com/RaczeQ



srai

kraina-ai.github.io/srai

github.com/kraina-ai/srai

This presentation template is free for everyone to use
thanks to the following:



for the presentation template

Pexels, Pixabay

for the photos

Geospatial Foundation Models Workshop, 2026-02-04