

SLING

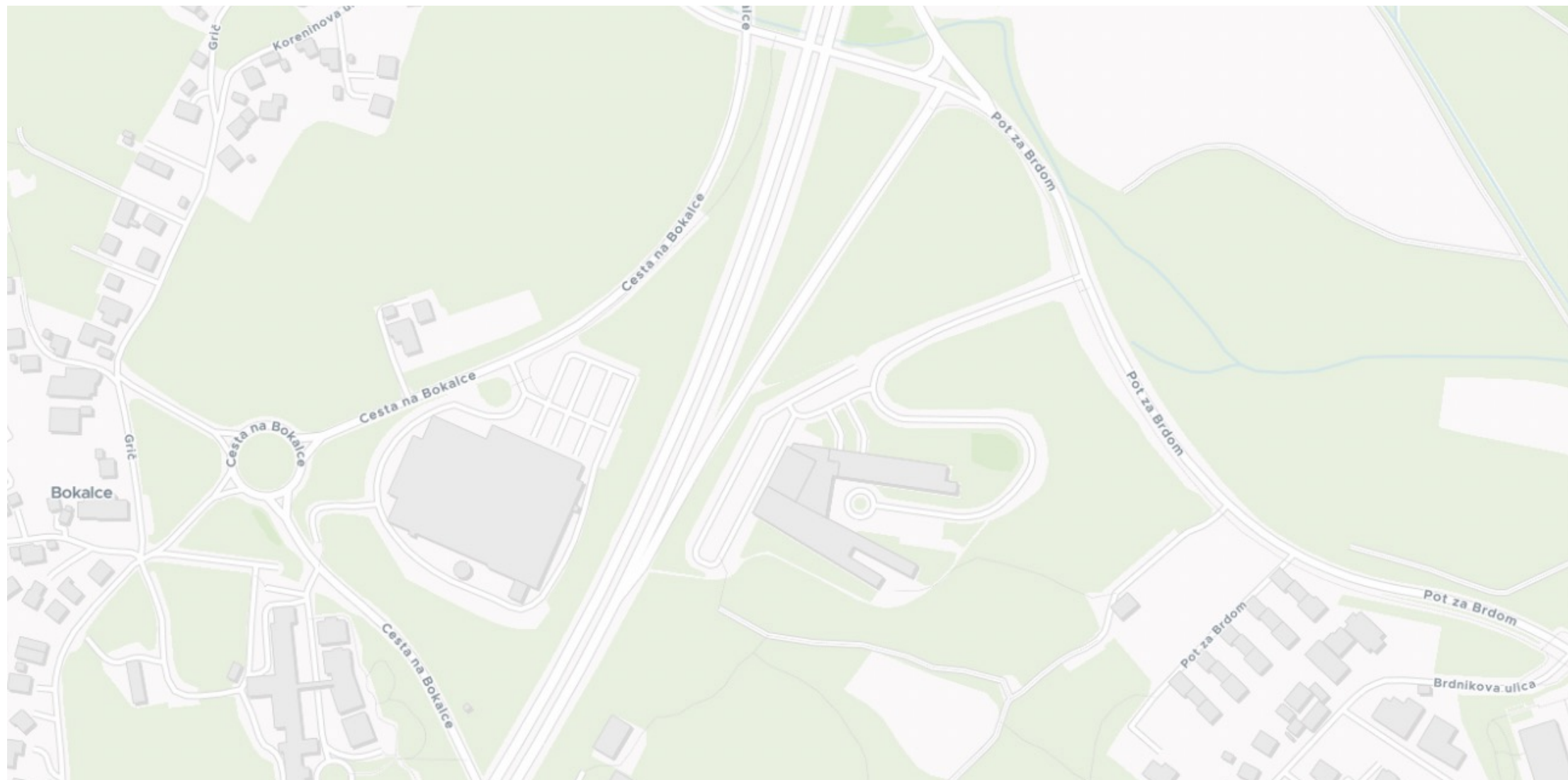


Dnevi SLING

Globoko učenje za razumevanje Lidar oblakov točk in uporaba super računalniške infrastrukture

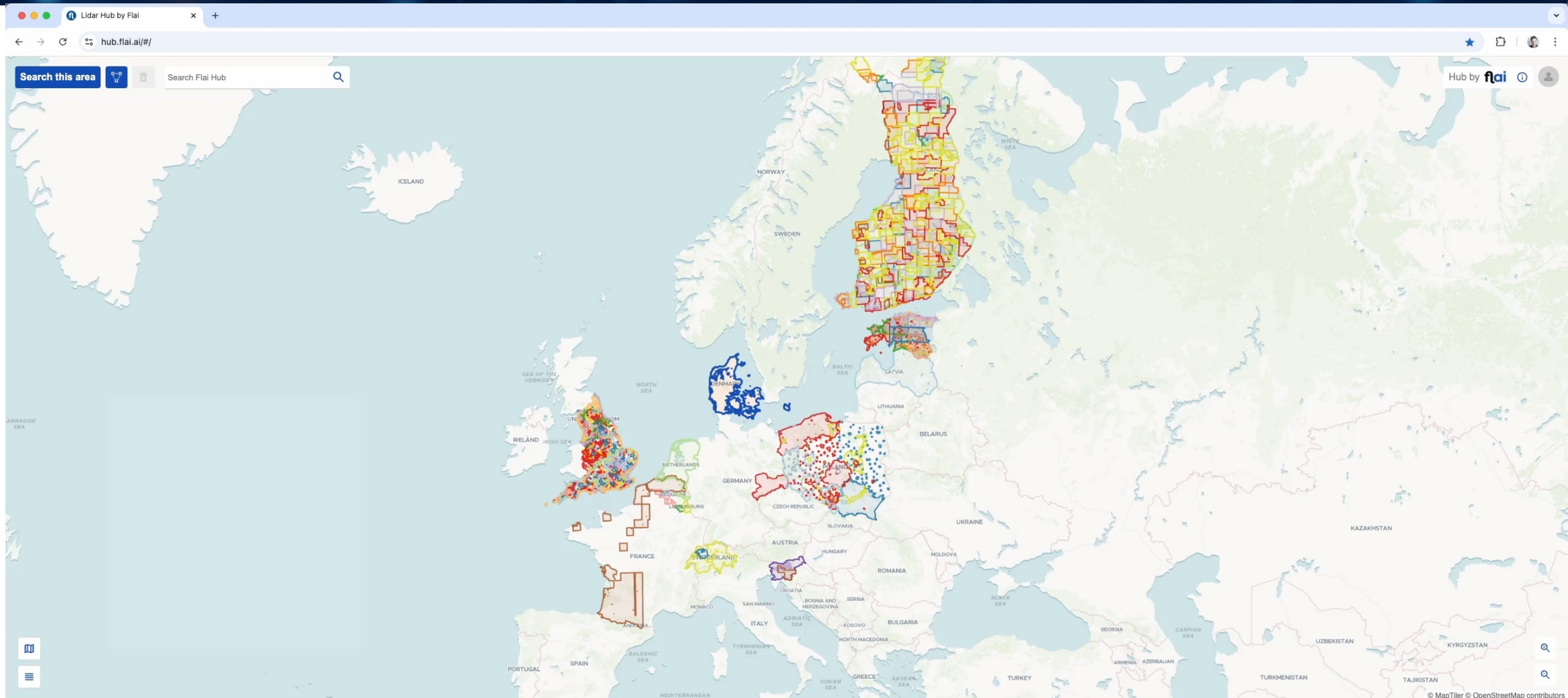
3.-5. december 2024

- Zakaj potrebujemo Lidar oblake točk?
- Kaj so Lidar oblaki točk?
- Kako lahko uporabimo globoko učenje za razumevanje Lidar oblakov točk?
- Proces učenja.
- Kako uporabljamo različno infrastrukturo - (super) računalnike?



3.-5. december 2024

Topograska kartiranje



Name	Time frame	Area	Point density	Reference System	License
Estonia (mets)	2015	8459.03 km ²	0.17 pts/m ²	EPSG:3301	Estonian Land Board, Elevation data 2017-2020 License
Estonia (madal)	2015	512.3 km ²	2.64 pts/m ²	EPSG:3301	Estonian Land Board, Elevation data 2017-2020 License
Germany	2015 - 2023	18992.41 km ²	13.25 pts/m ²	EPSG:25833	GeoSN, dl-de/by-2-0
Finland	2015	44552.68 km ²	1.54 pts/m ²	EPSG:3067	License
United Kingdom	2015	14394.29 km ²	3.15 pts/m ²	EPSG:27700	© Environment Agency License

Infrastruktura



Upravljanje z vegetacijo

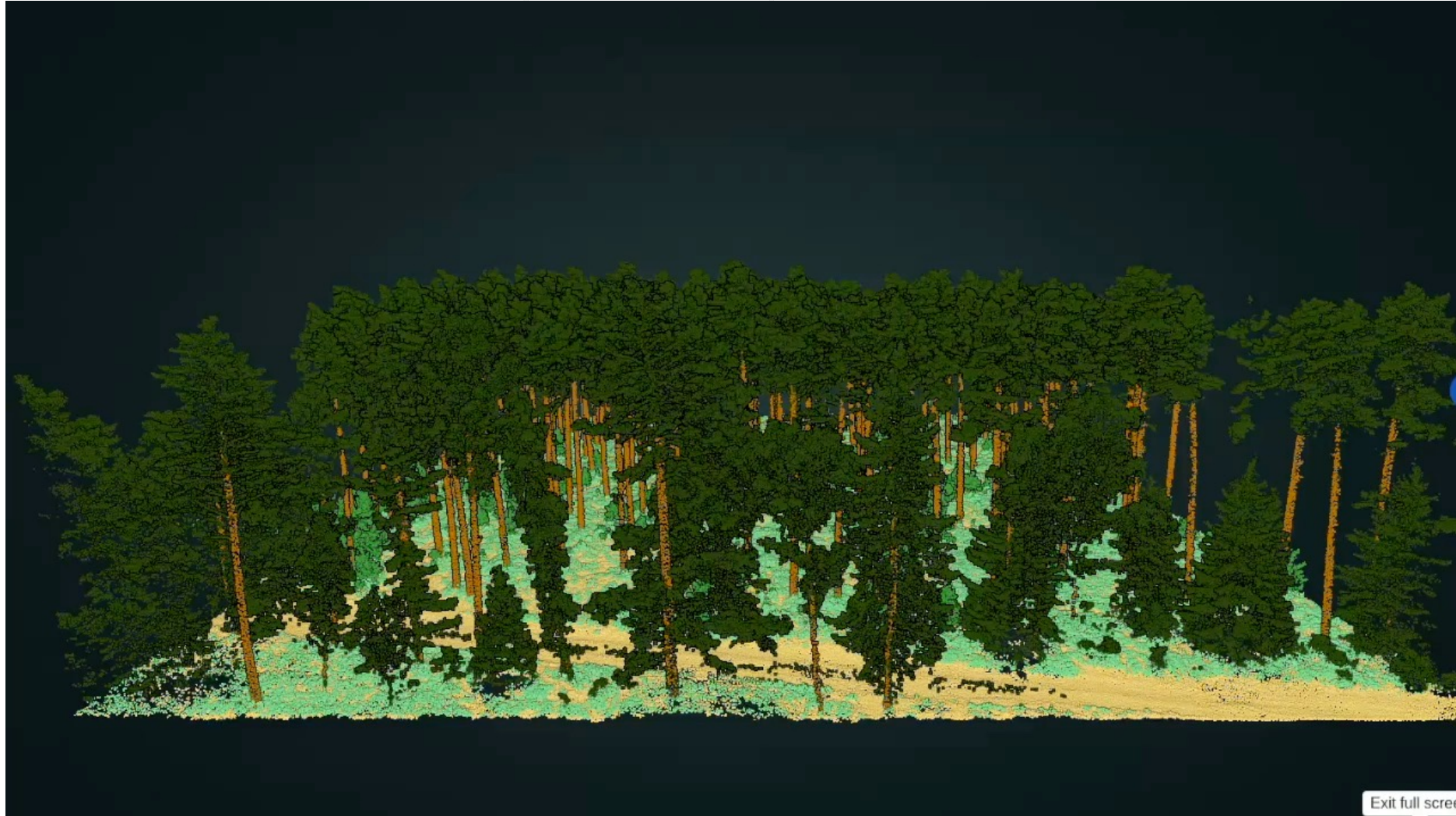
Pregled infrastrukture

Popis sredstev

Ocena stanja sredstev

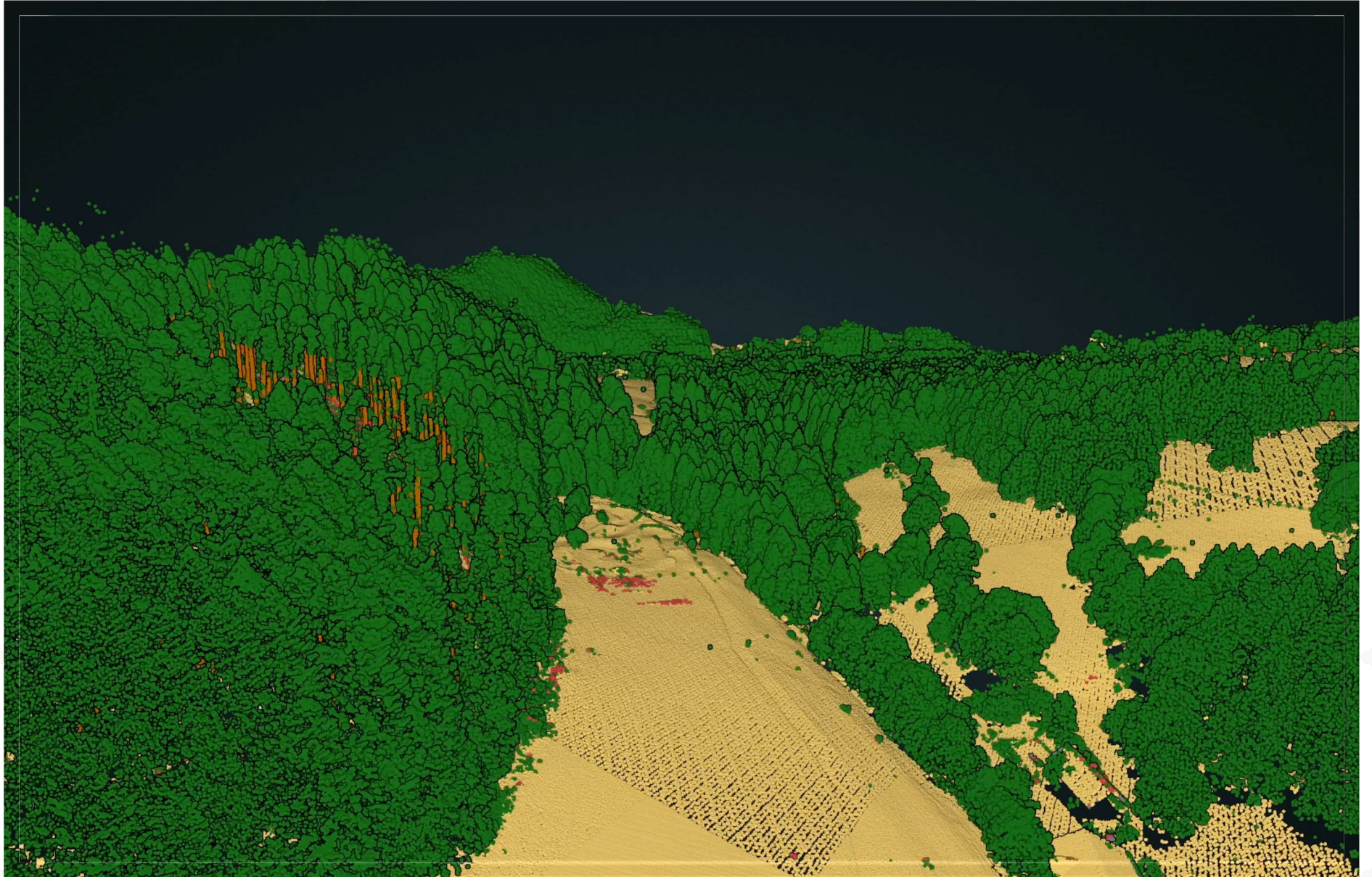
Odkrivanje napak

Gozdno gospodarstvo in okoljevarstvo

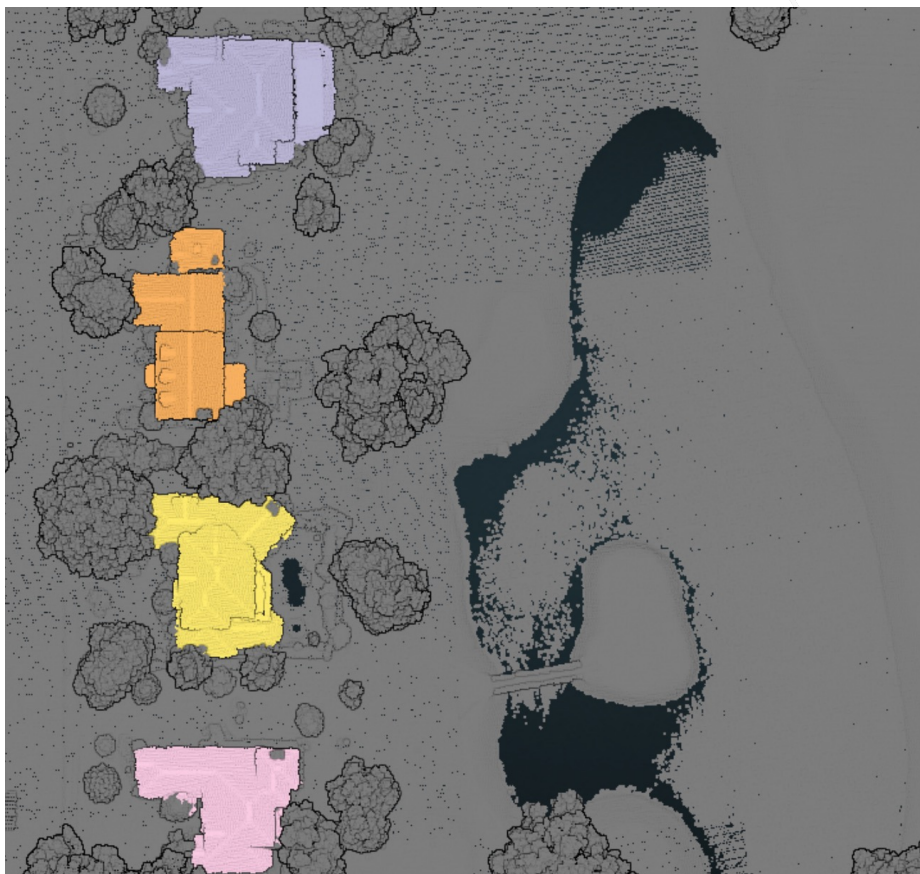


Obramba in varnost

3.-5. december 2024



Katere probleme?



- Klasifikacija
- Detekcija objektov
- Semantična segmentacija
- Segmentacija primerkov

Kako?

- Globoke nevronske mreže.
- Neposredno na oblaku točk.

arXiv:2012.09164v2 [cs.CV] 26 Sep 2021

Point Transformer

Hengshuang Zhao^{1,2} Li Jiang³ Jiaya Jia³ Philip Torr¹ Vladlen Koltun¹
¹University of Oxford ²The University of Hong Kong
³The Chinese University of Hong Kong ⁴Intel Labs

Abstract

Self-attention networks have revolutionized natural language processing and are making impressive strides in image analysis tasks such as image classification and object detection. Inspired by this success, we investigate the application of self-attention networks to 3D point cloud processing. We design self-attention layers for point clouds and use these to construct self-attention networks for tasks such as semantic scene segmentation, object part segmentation, and object classification. Our Point Transformer design improves upon prior work across domains and tasks. For example, on the challenging S3DIS dataset for large-scale semantic scene segmentation, the Point Transformer attains an mIoU of 70.4% on Area 5, outperforming the strongest prior model by 3.3 absolute percentage points and crossing the 70% mIoU threshold for the first time.



Figure 1. The Point Transformer can serve as the backbone for various 3D point cloud understanding tasks such as object classification, object part segmentation, and semantic scene segmentation.

1. Introduction

3D data arises in many application areas such as autonomous driving, augmented reality, and robotics. Unlike images, which are arranged on regular pixel grids, 3D point clouds are sets embedded in continuous space. This makes 3D point clouds structurally different from images and precludes immediate application of deep network designs that have become standard in computer vision, such as networks based on the discrete convolution operator.

A variety of approaches to deep learning on 3D point clouds have arisen in response to this challenge. Some voxelize the 3D space to enable the application of 3D discrete convolutions [23, 32]. This induces massive computational and memory costs and underutilizes the sparsity of point sets in 3D. Sparse convolutional networks relieve these limitations by operating only on voxels that are not empty [9, 3]. Other designs operate directly on points and propagate information via pooling operators [25, 27] or continuous convolutions [42, 37]. Another family of approaches connect the point set into a graph for message passing [44, 19].

In this work, we develop an approach to deep learning on point clouds that is inspired by the success of transformers

in natural language processing [39, 45, 5, 4, 51] and image analysis [10, 28, 54]. The transformer family of models is particularly appropriate for point cloud processing because the self-attention operator, which is at the core of transformer networks, is in essence a set operator: it is invariant to permutation and cardinality of the input elements. The application of self-attention to 3D point clouds is therefore quite natural, since point clouds are essentially sets embedded in 3D space.

We flesh out this intuition and develop a self-attention layer for 3D point cloud processing. Based on this layer, we construct Point Transformer networks for a variety of 3D understanding tasks. We investigate the form of the self-attention operator, the application of self-attention to local neighborhoods around each point, and the encoding of positional information in the network. The resulting networks are based purely on self-attention and pointwise operations.

We show that Point Transformers are remarkably effective in 3D deep learning tasks, both at the level of detailed object analysis and large-scale parsing of massive scenes. In particular, Point Transformers set the new state of the art on large-scale semantic segmentation on the S3DIS dataset (70.4% mIoU on Area 5), shape classification on ModelNet40 (93.7% overall accuracy), and object part segmenta-

modern computer mental building local neighborhoods on a 2D grid. Thanks to this regular structure, it can be computed with high efficiency on modern hardware, but when deprived of this regular structure, the convolution operation has yet to be defined properly, with the same efficiency as on 2D grids.

Many applications relying on such irregular data have grown with the rise of 3D scanning technologies. For example, 3D point cloud segmentation or 3D simultaneous localization and mapping rely on non-grid structured data: point clouds. A point cloud is a set of points in 3D (or higher-dimensional) space. In many applications, the points

¹Project page: <https://github.com/HuguesTHOMAS/KPCConv>

PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation

Charles R. Qi* Hao Su* Kaichun Mo Leonidas J. Guibas
Stanford University

Abstract

Point cloud is an important type of geometric data structure. Due to its irregular format, most researchers transform such data to regular 3D voxel grids or collections of images. This, however, renders data unnecessarily

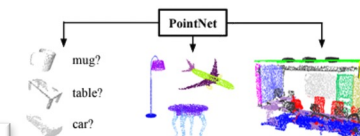


Figure 1. Applications of PointNet. We propose a novel deep net architecture that consumes raw point cloud (set of points) without voxelization or rendering. It is a unified architecture that learns both global and local point features, providing a simple, efficient and effective approach for a number of 3D recognition tasks.

and Deformable Convolution for Point Clouds

R. Qi² Jean-Emmanuel Deschaud¹ Beatriz Marcotequi¹
Boris Goulette¹ Leonidas J. Guibas^{2,3}

²Facebook AI Research ³Stanford University

KPCConv), a new erates on point tion. The convolution space by is close to them. points gives KPCConv. Further- ice and can be can be extended up kernel points sampling strategy densities, complex tasks, or rks outperform tion studies and what has been scription power

are coupled with corresponding features like colors. In this work, we will always consider a point cloud as those two elements: the points $P \in \mathbb{R}^{N \times 3}$ and the features $F \in \mathbb{R}^{N \times D}$. Such a point cloud is a sparse structure that has the property to be unordered, which makes it very different from a grid. However, it shares a common property with a grid which is essential to the definition of convolutions: it is spatially localized. In a grid, the features are localized by their index in a matrix, while in a point cloud, they are localized by their corresponding point coordinates. Thus, the points are to be considered as structural elements, and the features as the real data.

Various approaches have been proposed to handle such data, and can be grouped into different categories that we will develop in the related work section. Several methods fall into the grid-based category, whose principle is to project the sparse 3D data on a regular structure where a convolution operation can be defined more easily [24, 29, 34]. Other approaches use multilayer perceptrons (MLP) to process point clouds directly, following the idea proposed by [49, 26].

More recently, some attempts have been made to design a convolution that operates directly on points [2, 45, 20, 14, 13]. These methods use the spatial localization property of

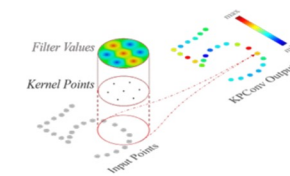
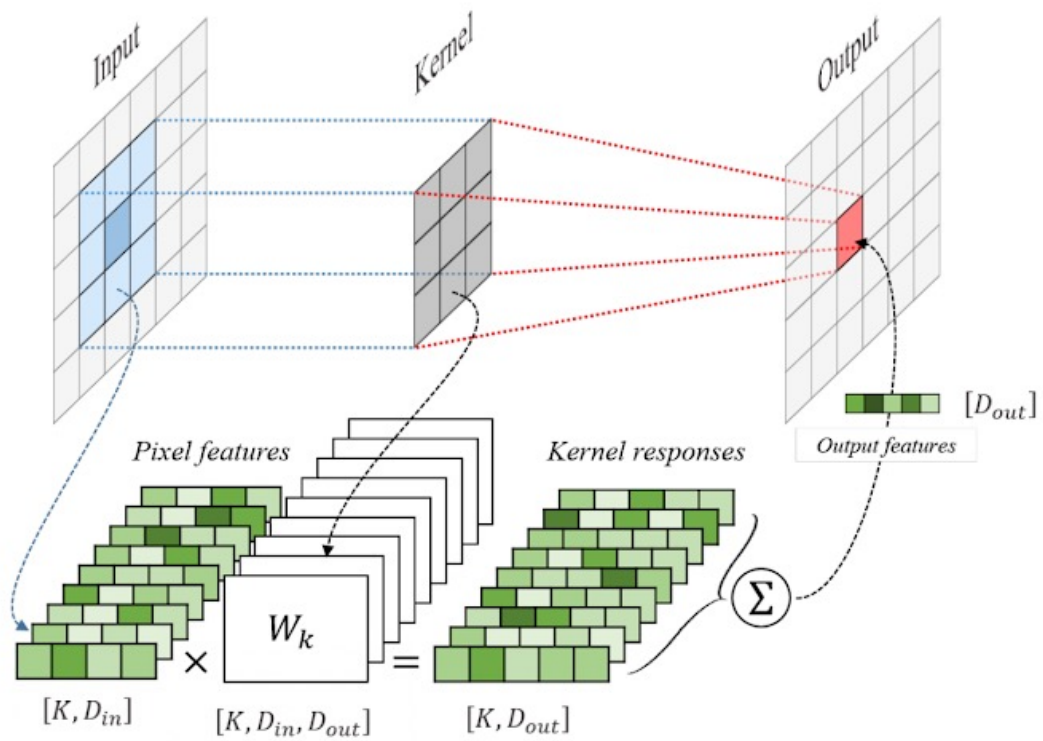


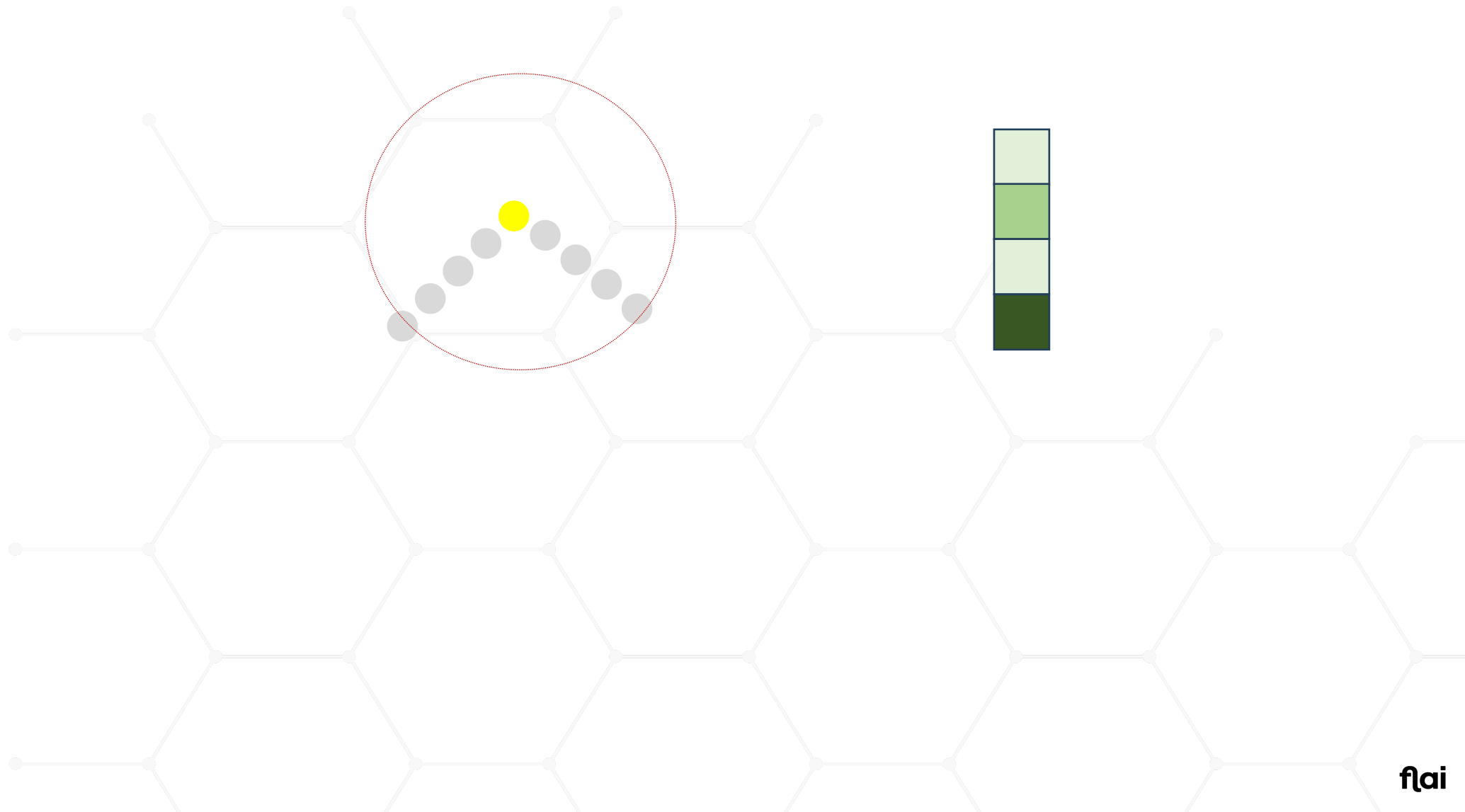
Figure 1. KPCConv illustrated on 2D points. Input points with a constant scalar feature (in grey) are convolved through a KPCConv that is defined by a set of kernel points (in black) with filter weights on each point.



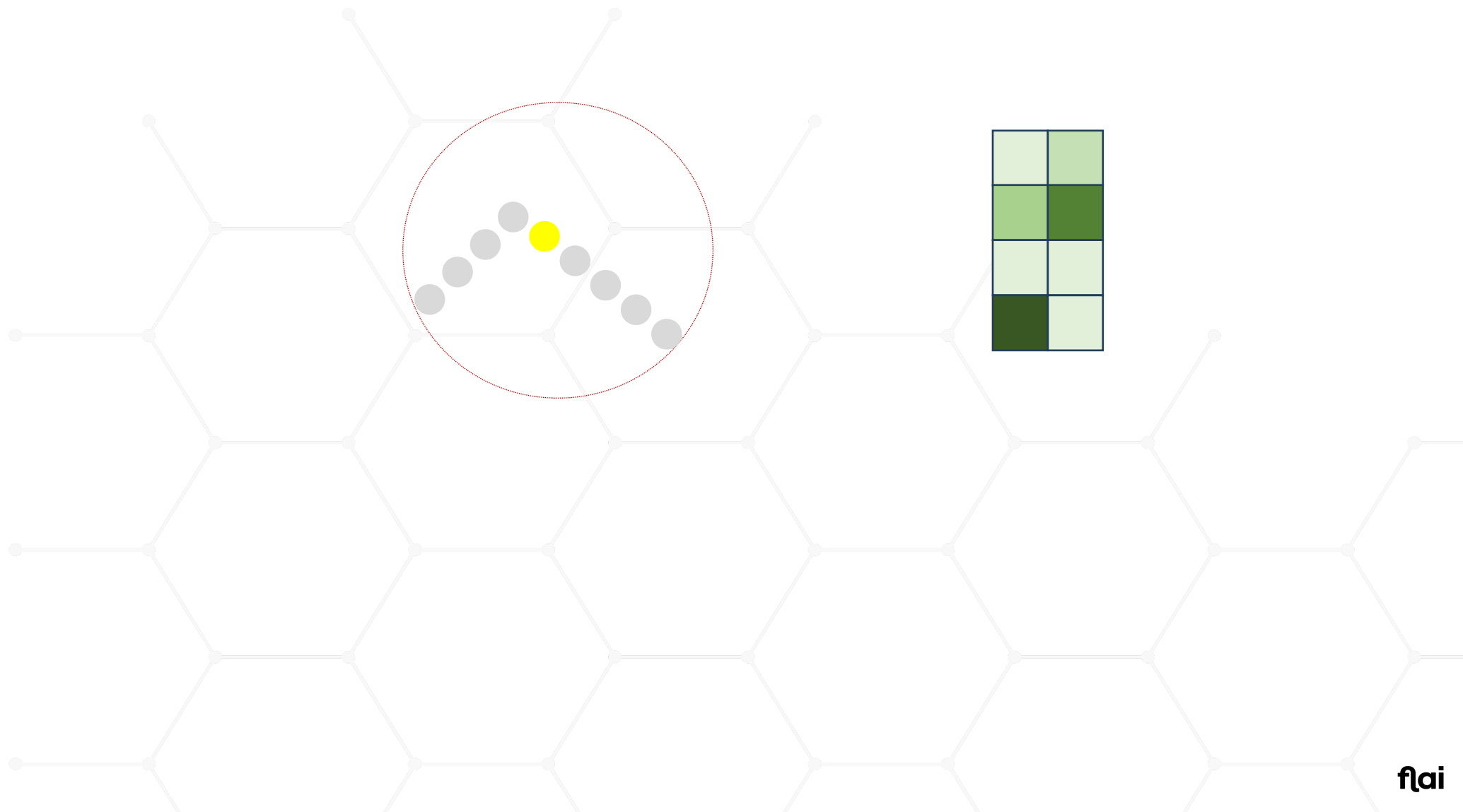


3.-5. december 2024

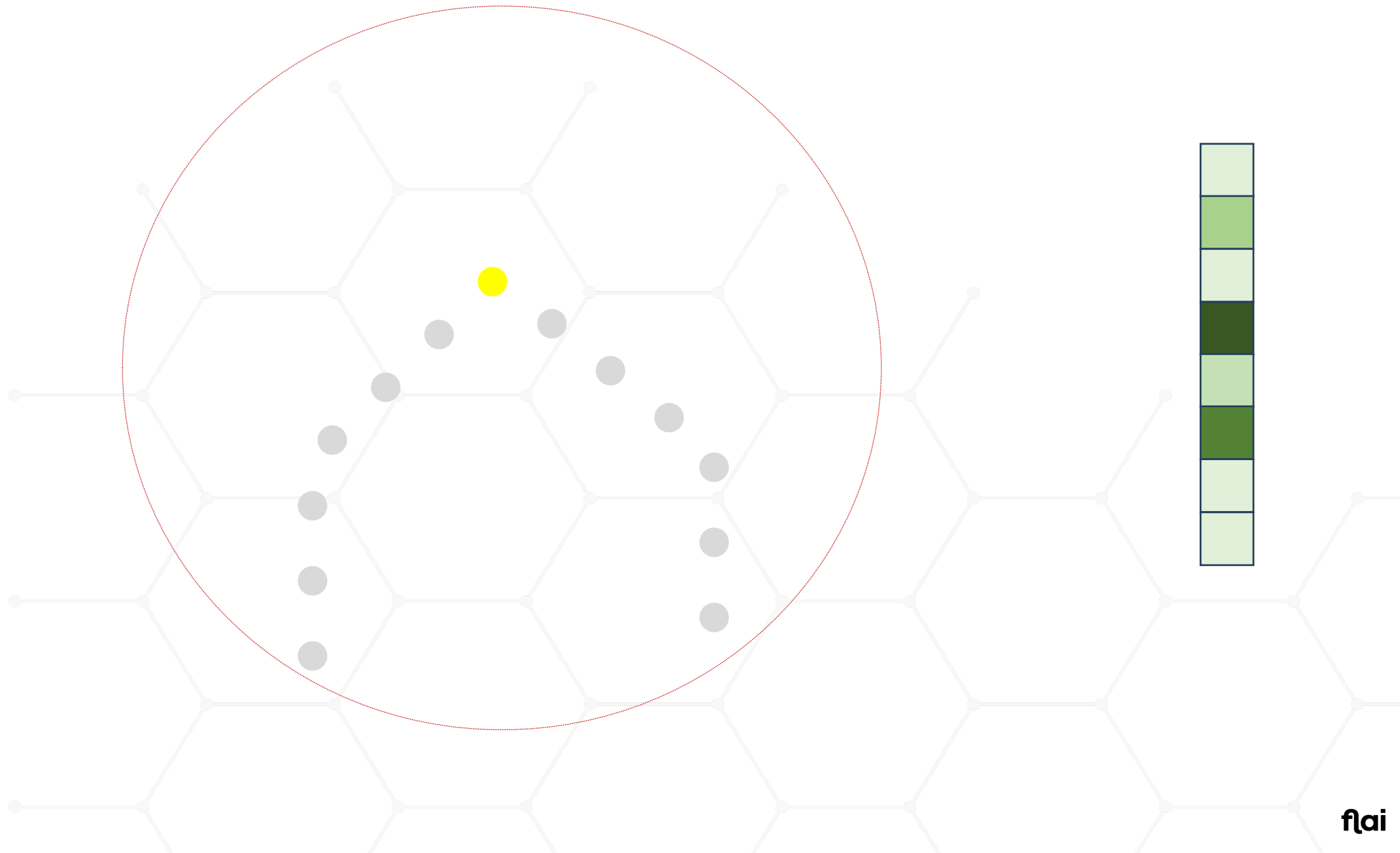
fai



3.-5. december 2024



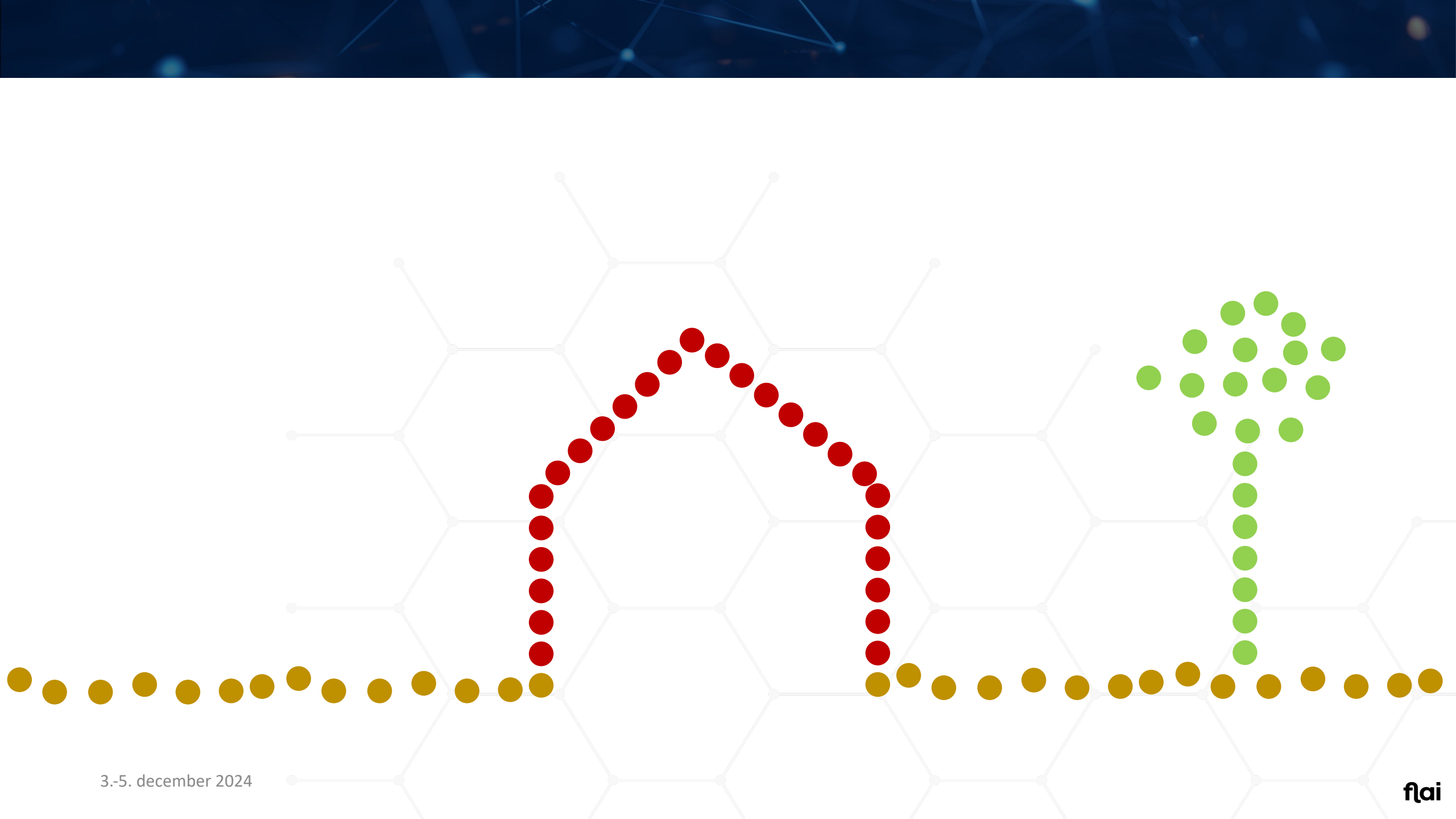
3.-5. december 2024



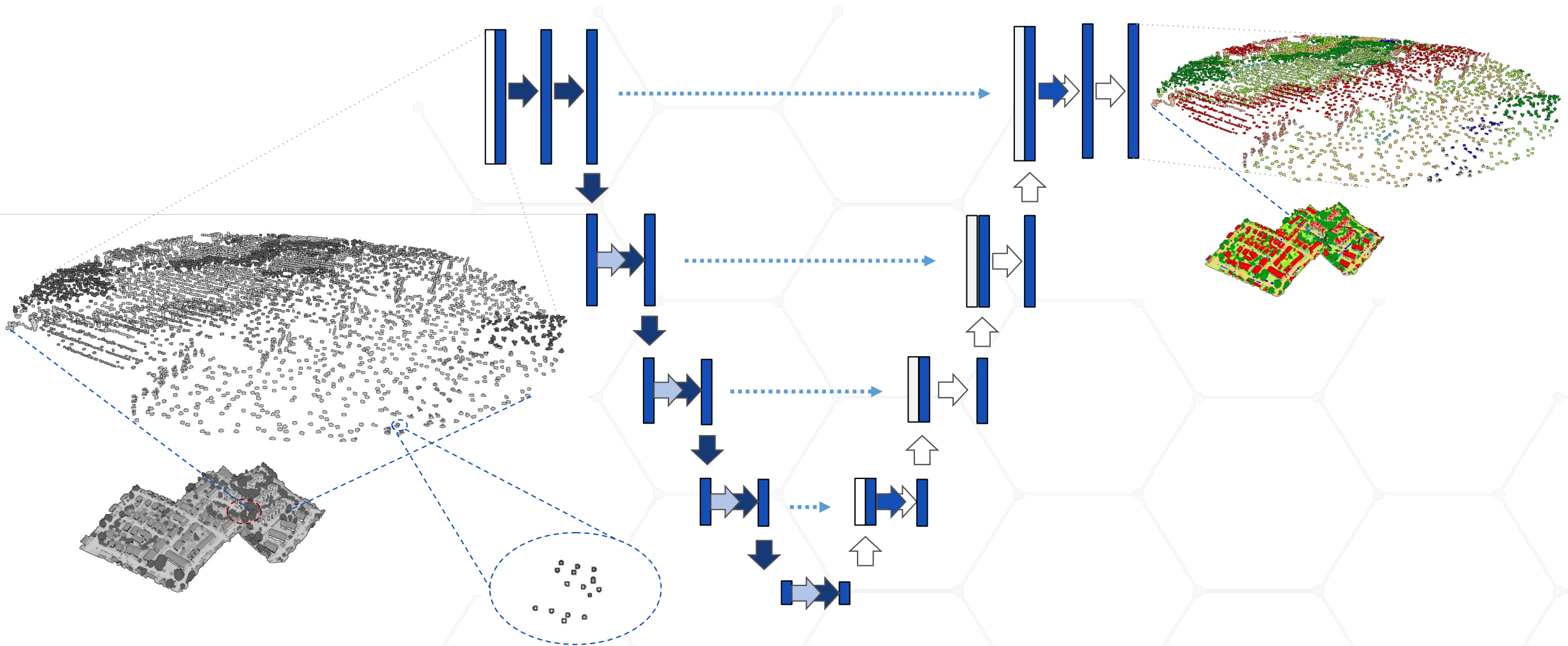
3.-5. december 2024



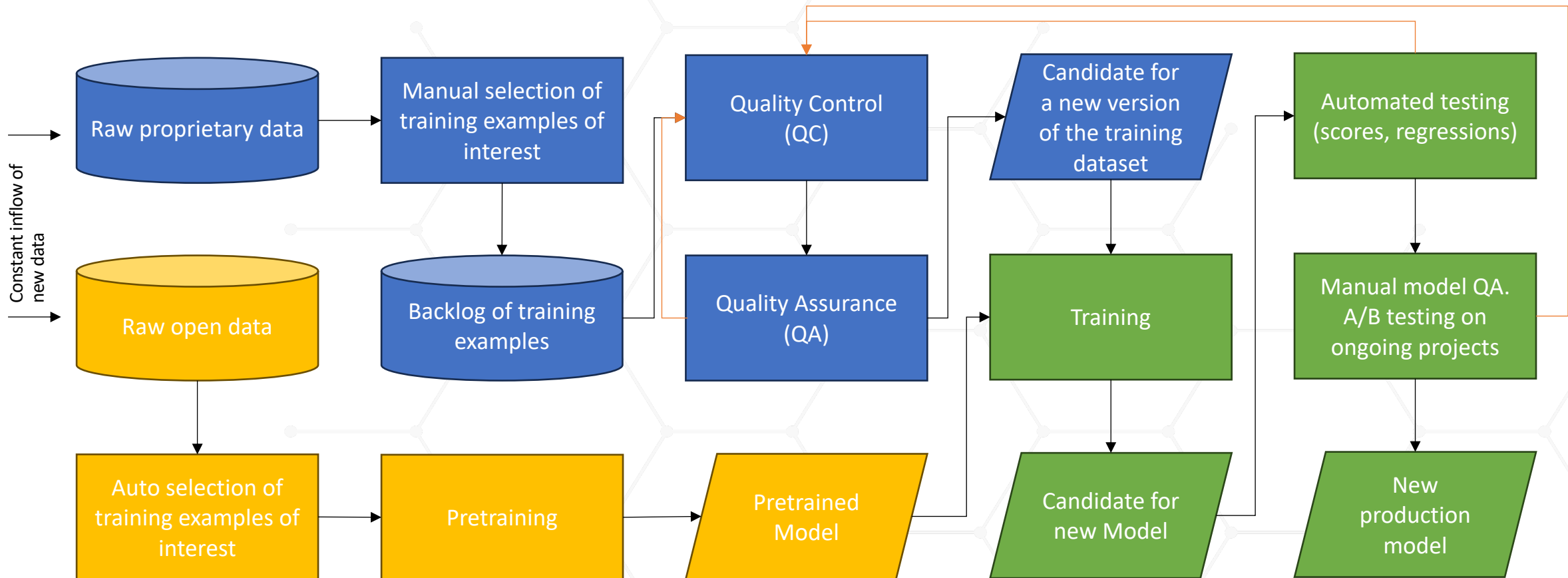
3.-5. december 2024



3.-5. december 2024



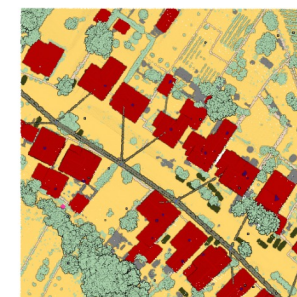
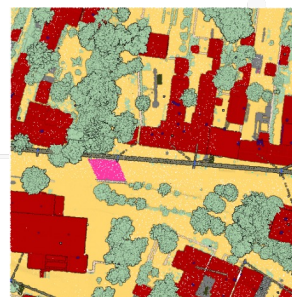
Neprekinjen proces učenja novih modelov



Kakšne količine podatkov?

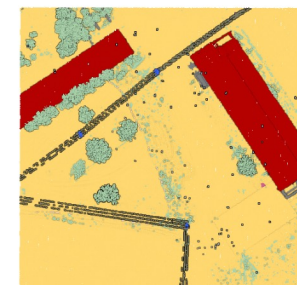
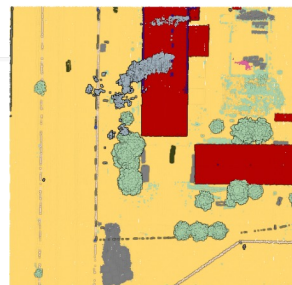
Proprietary data

115,5 TB



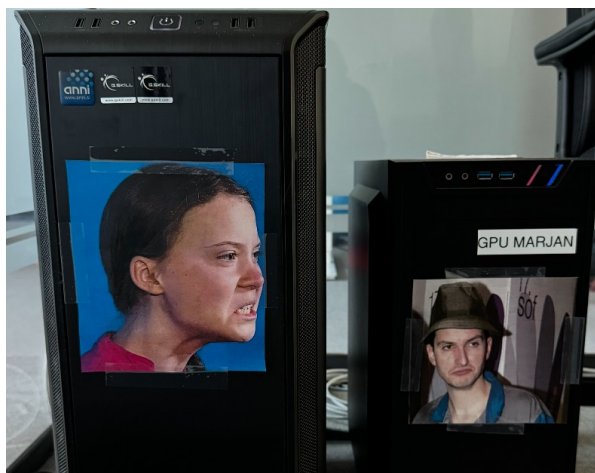
Open data

194,4 TB



Kje učimo in kje izvajamo inferenco?

Onprem



večinoma A6000

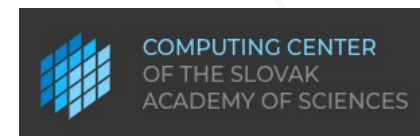
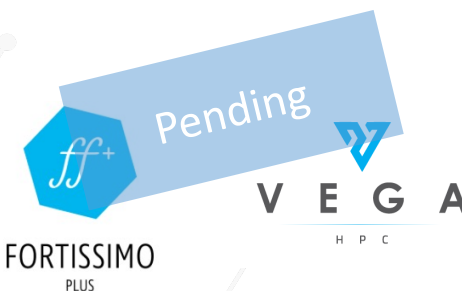
Cloud



Google Cloud Platform

večinoma A100

HPC



SLING



Hvala!

Globoko učenje za razumevanje Lidar oblakov točk in uporaba super računalniške infrastrukture



EuroHPC
Joint Undertaking



REPUBLIKA SLOVENIJA
**MINISTRSTVO ZA VISOKO ŠOLSTVO,
ZNANOST IN INOVACIJE**

Projekt EuroCC 2 financira Evropska unija. Financiran je s sredstvi Skupnega podjetja za evropsko visokozmogljivo računalništvo (EuroHPC JU) ter Nemčije, Bolgarije, Avstrije, Hrvaške, Cipra, Češke republike, Danske, Estonije, Finske, Grčije, Madžarske, Irske, Italije, Litve, Latvije, Poljske, Portugalske, Romunije, Slovenije, Španije, Švedske, Francije, Nizozemske, Belgije, Luksemburga, Slovaške, Norveške, Turčije, Republike Severne Makedonije, Islandije, Črne gore in Srbije v okviru sporazuma o dodelitvi sredstev št. 101101903.