



Interim presentation 04.12.2020

PG - Bicycle Data

Structure

- Introduction
- SmartHelm
- Database
- BITS



Introduction

■ BITS

- Unification of different data sources
 - Counting data
 - Near-accidents
 - Bicycle parking
 - Snifferbike


 Output of open cycle data and analysis results



■ SmartHelm

- Development of an intelligent bicycle helmet as an assistant for courier drivers
- Data Warehouse
 - Storage, analysis and processing of data
 - Integration of various data formats



 The aim is to analyse bicycle routes

- Using EEG data



SmartHelm

Review

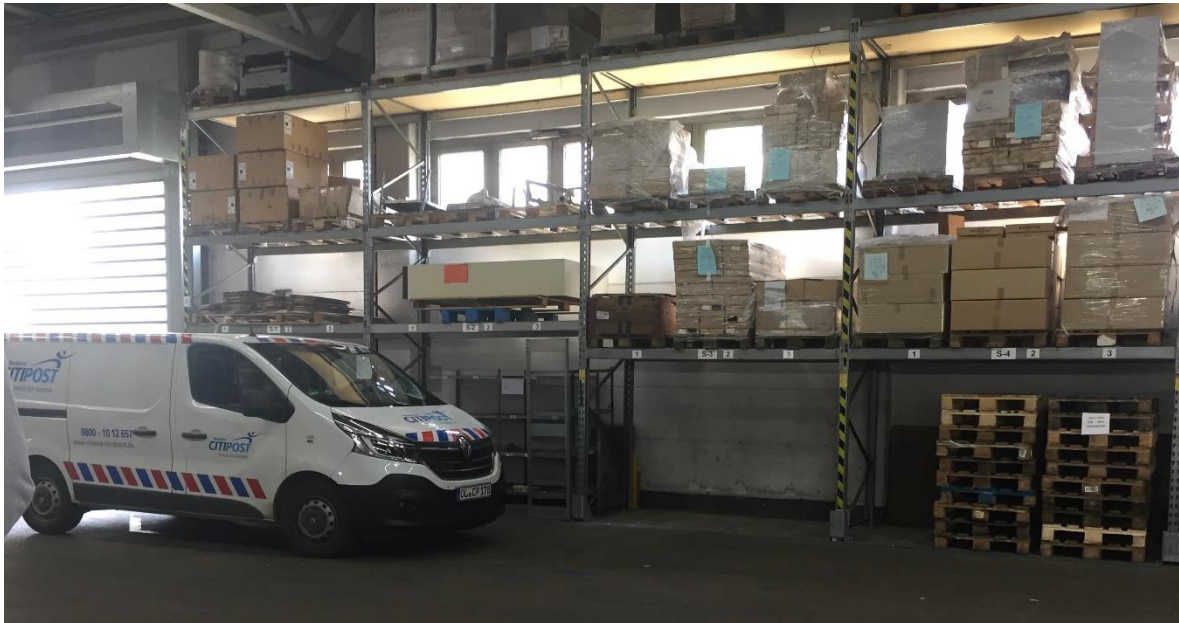
- Objective: To analyse bicycle routes using SmartHelm data
- Development of a data warehouse system
- Visualisation of the distraction factors via heat maps



Quelle: UVEX SPORTS GROUP

Data availability

- No hall study could be carried out so far, only a laboratory study was carried out -> No complete data basis for analyses available



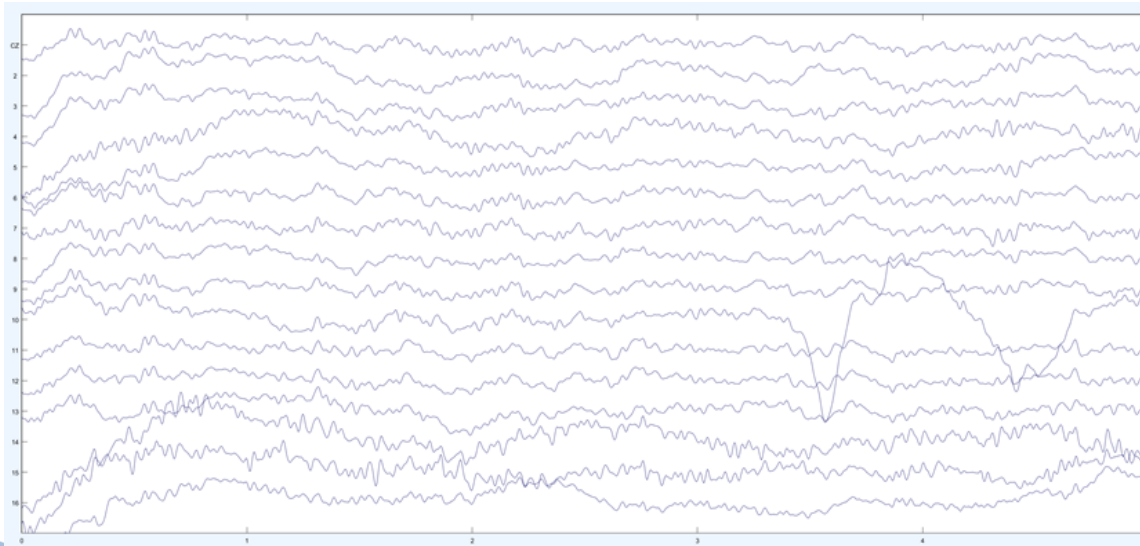
Picture: Dr. Dennis Küster

- Alternative approach: synthesise EEG data based on laboratory data provided by the Cognitive Systems Lab (CSL) of the University of Bremen

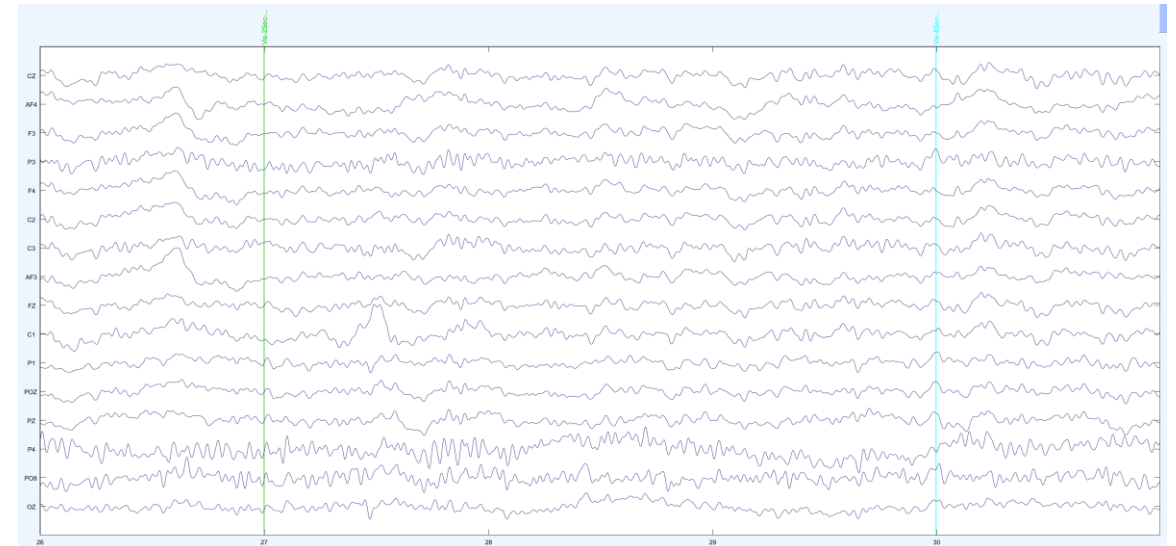
Preprocessing of EEG data

Mapping

- .xdf files from the experiment were imported into EEGLAB
- Mapping of the measuring points (electrodes) used in the experiments
- Mapping of the events to the EEG data
- Noise data removal



Raw (imported) data

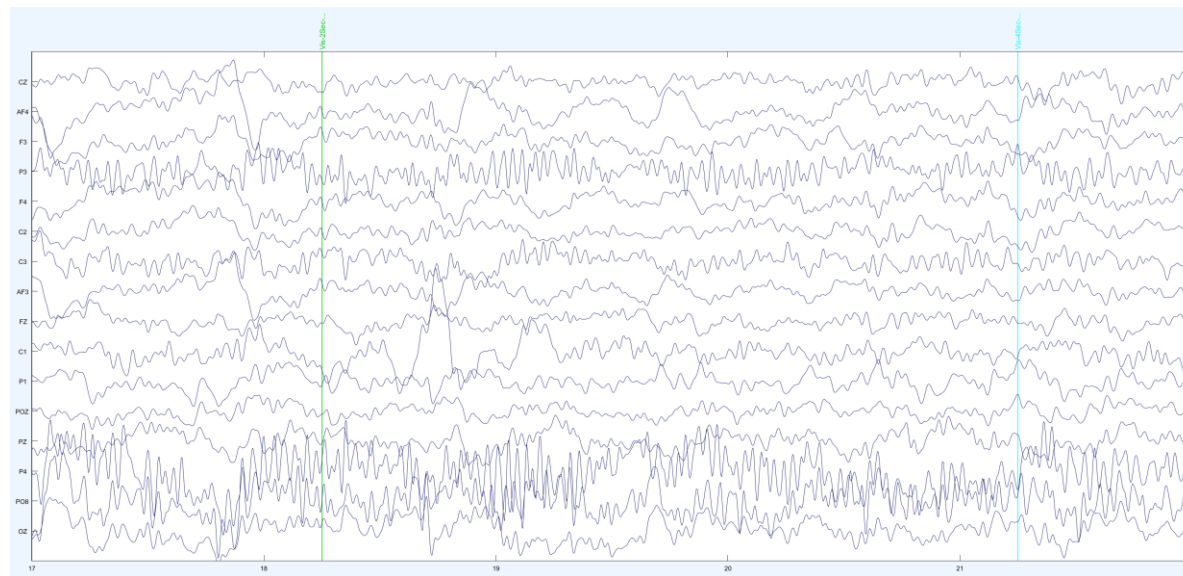


Mapped data

Preprocessing of EEG data

Preprocessing without deletion (ICA)

- In preparation for ICA, the data was re-referenced and filtered
- Application to recognise and analyse individual components in data

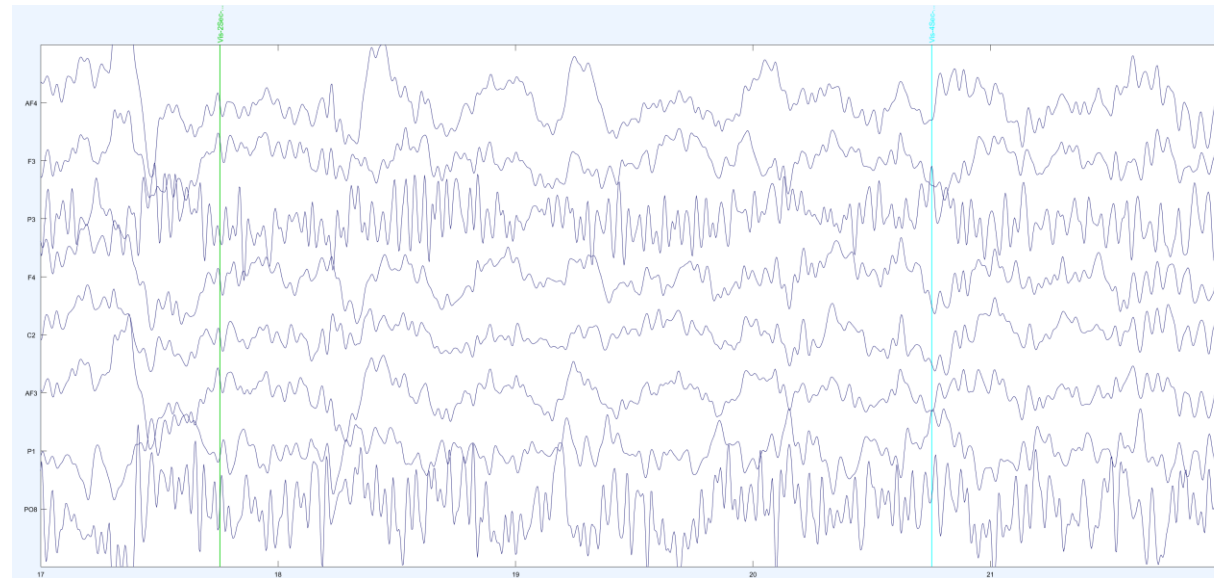


Pre-processed data

Preprocessing of EEG data

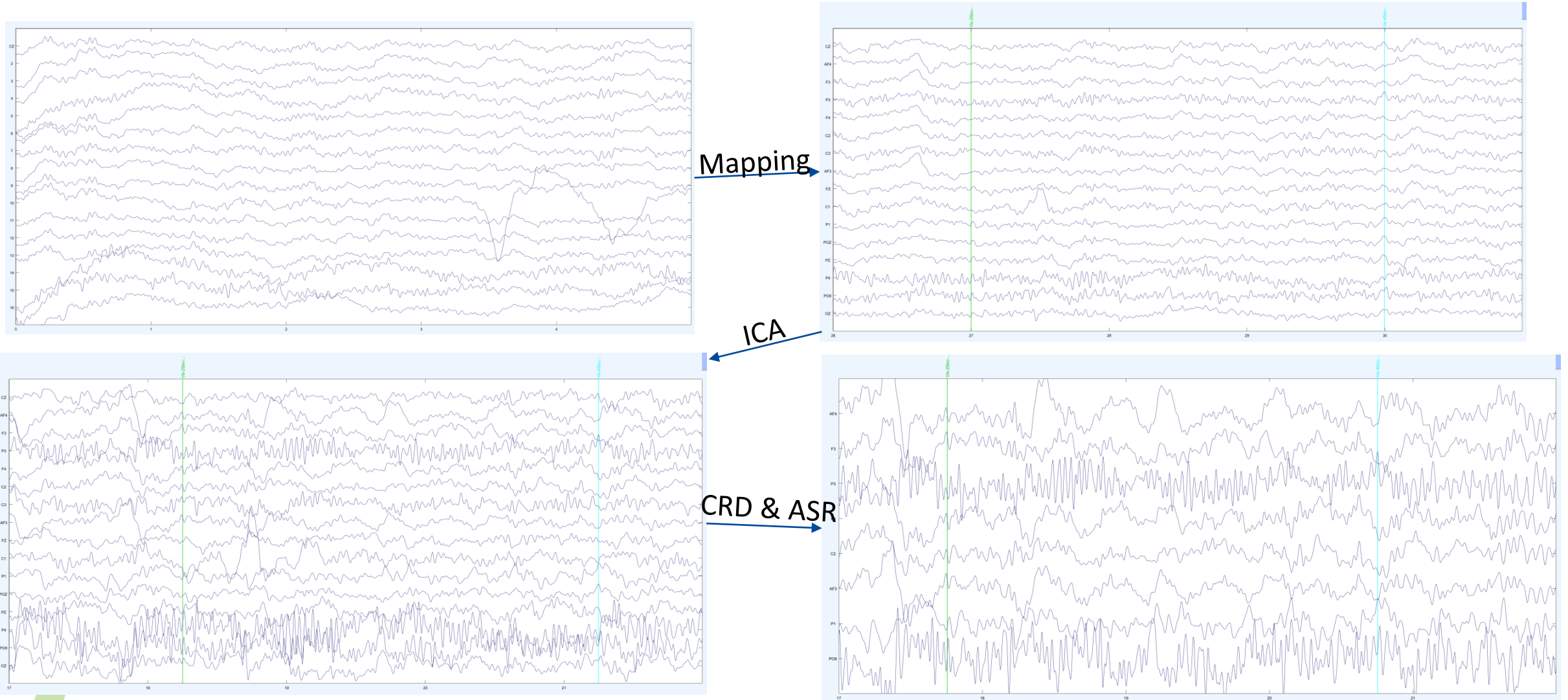
Preprocessing with deletion (Clear Raw Data, ASR)

- Resampling of the data
- Methods were applied with default settings



Preprocessed data (with deletion)

Preprocessing of EEG data



Production of synthetic data

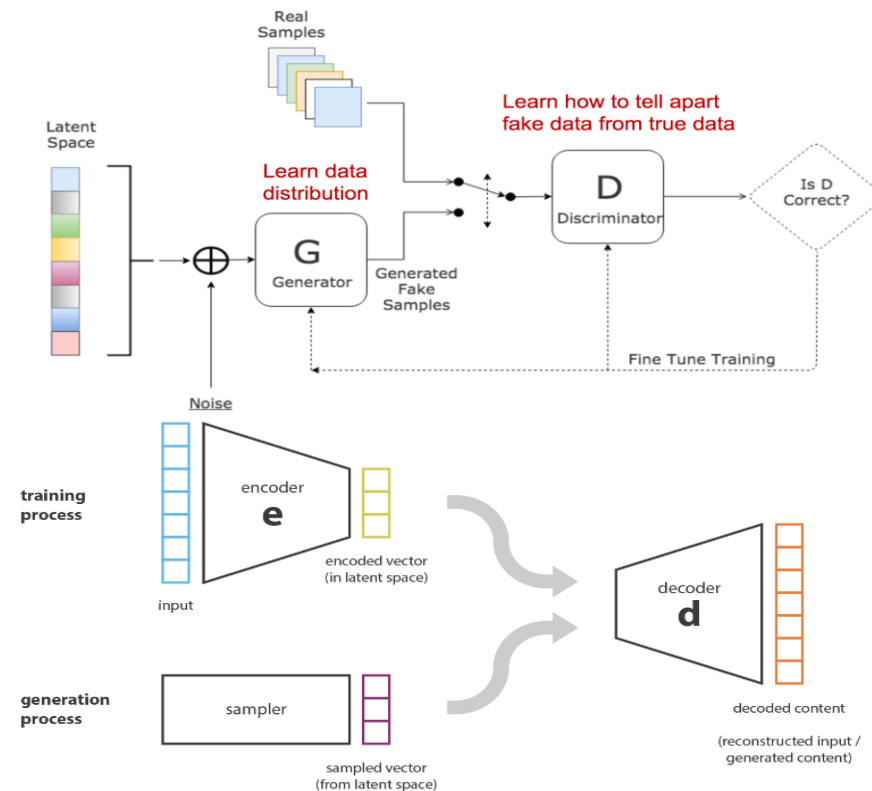
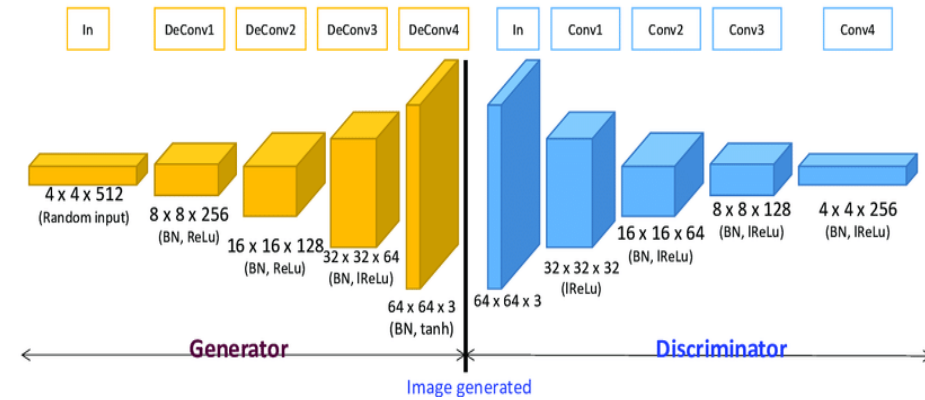
- Result of some tests: Standard libraries not suitable
- Research result: Treatment of the problem with Machine Learning



Create Synthetic Data

Successfully tested methods

- Deep Convolutional Generative Adversarial Network (DCGAN)
 - generate data that is improved until it is considered to be real data
- Wasserstein Generative Adversarial Network (WGANM)
 - Similar to the DCGAN, with differences in the input of the generator and the operation of the discriminator
- Variational Autoencoder (VAE)
 - The data is controlled and automatically coded within a certain variation



Outlook

- Structure of our data is fundamentally different, therefore a simple use of the given methods is not possible
- Conclusion: Synthesis of EEG data very complex
- New main task: research and implementation of data synthesis with specially developed machine learning algorithms
 - The performance of the analyses is postponed





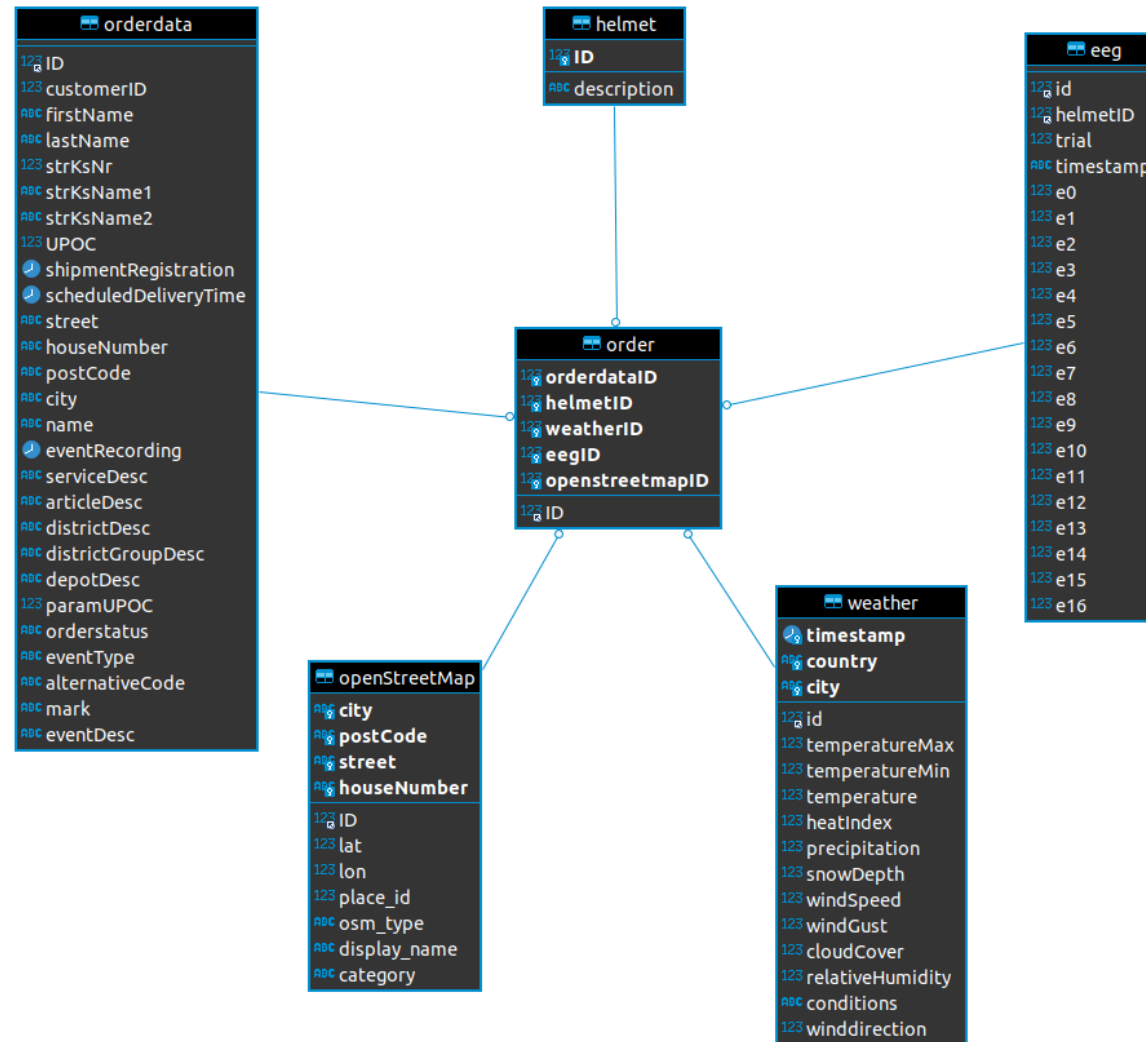
Database

Intentions

- Unification of all CSV files from different partners
- Easy connection to other projects (e.g. CyclingDataHub)
- Automated algorithmic analyses
- MariaDB
 - Relational database
 - State of the art performance
 - Open Source (developed from MySQL)
 - Widespread



Star schema in the Smarthelm database



Different structures

	A	B	C	D	E	F	G
1	date	time start	time end	counting station	forward	reverse	total
2	2020.04.01.	00:00	00:15	Berchem TOT	0	0	0
3	2020.04.01.	00:15	00:30	Berchem TOT	0	2	2
4	2020.04.01.	00:30	00:45	Berchem TOT	0	0	0
5	2020.04.01.	00:45	01:00	Berchem TOT	0	0	0
6	2020.04.01.	01:00	01:15	Berchem TOT	0	0	0
7	2020.04.01.	01:15	01:30	Berchem TOT	0	0	0
8	2020.04.01.	01:30	01:45	Berchem TOT	0	0	0



Different structures

	A	B	C	D	E	F	G										
1	date	time start	time end	counting station	forward	reverse	total										
2	2020.04.01.	00:00	00:15	Berchem TOT	0	0	0										
3	2020.04.01.	00:15	00:30	Berchem TOT	0	2	2										
4	2020.04.01.	00:30	00:45	Berchem TOT				A	B	C	D	E	F	G	H	I	J
5	2020.04.01.	00:45	01:00	Berchem TOT	1	Zeitraum		1. April 2018 -> 30. April 2018									
6	2020.04.01.	01:00	01:15	Berchem TOT	2												
7	2020.04.01.	01:15	01:30	Berchem TOT	3	Time		50101 Damm	50601_Nado	50201_Garte	heilig	50701_Pferd	50301_Roons	50401 Ofene	50501 Alexar	50801_Amal	
8	2020.04.01.	01:30	01:45	Berchem TOT	4	Sonntag, 1. April 2018		3	0	6	6	6	1	9	2	C	
					5	Sonntag, 1. April 2018		6	3	4	7	3	0	10		C	
					6	Sonntag, 1. April 2018		4	0	3	3	5	0	11	0	C	
					7	Sonntag, 1. April 2018		10	0	2	2	1	1	12	2	C	
					8	Sonntag, 1. April 2018		4	0	4	4	0	1	6	4	C	
					9	Sonntag, 1. April 2018		4	9	5	14	3	0	8		C	
					10	Sonntag, 1. April 2018		9	0	6	6	5	0	6	0	C	



Different structures

	A	B	C	D	E	F	G
1	date	time start	time end	counting station	forward	reverse	total
2	2020.04.01.	00:00	00:15	Berchem TOT	0	0	0
3	2020.04.01.	00:15	00:30	Berchem TOT	0	2	2
4	2020.04.01.	00:30	00:45	Berchem TOT			
5	2020.04.01.	00:45	01:00	Berchem TOT			
6	2020.04.01.	01:00	01:15	Berchem TOT			
7	2020.04.01.	01:15	01:30	Berchem TOT			
8	2020.04.01.	01:30	01:45	Berchem TOT			

	A	B	C	D	E	F	G	H	I	J
1	Zeitraum	1. April 2018 -> 30. April 2018								
2										
3	Time	50101 Damm	50601_Nado	50201_Garte	heilig	50701_Pferd	50301_Roons	50401 Ofene	50501 Alexar	50801_Amal
4	Sonntag, 1. April 2018	3	0	6	6	6	1	9	2	C
5	Sonntag, 1. April 2018	6	3	4	7	3	0	10		C
6	Sonntag, 1. April 2018	4	0	3	3	5	0	11	0	C
7	Sonntag, 1. April 2018	10	0	2	2	1	1	12	2	C
8	Sonntag, 1. April 2018	4	0	4	4	0	1	6	4	C
9	Sonntag, 1. April 2018	4	9	5	14	3	0	8		C
10	Sonntag, 1. April 2018	9	0	6	6	5	0	6	0	C

	A	B	C	D	E	F
1	Datum	K15 Berliner Stra�ve	K61 Christophorusweg	K51 Nikolausberger Weg	K63 Robert-Koch-Stra�ve	K13 Wender Landstra�ve
2	Montag, 13. April 2020	19	8	6	15	13
3	Montag, 13. April 2020	21	9	14	13	15
4	Montag, 13. April 2020	12	2	4	24	11
5	Montag, 13. April 2020	7	3	15	10	11
6	Montag, 13. April 2020	16	2	4	8	11
7	Montag, 13. April 2020	31	8	22	10	9
8	Montag, 13. April 2020	16	8	12	23	8
9	Montag, 13. April 2020	7	14	9	17	10
10	Montag, 13. April 2020	16	8	17	18	9



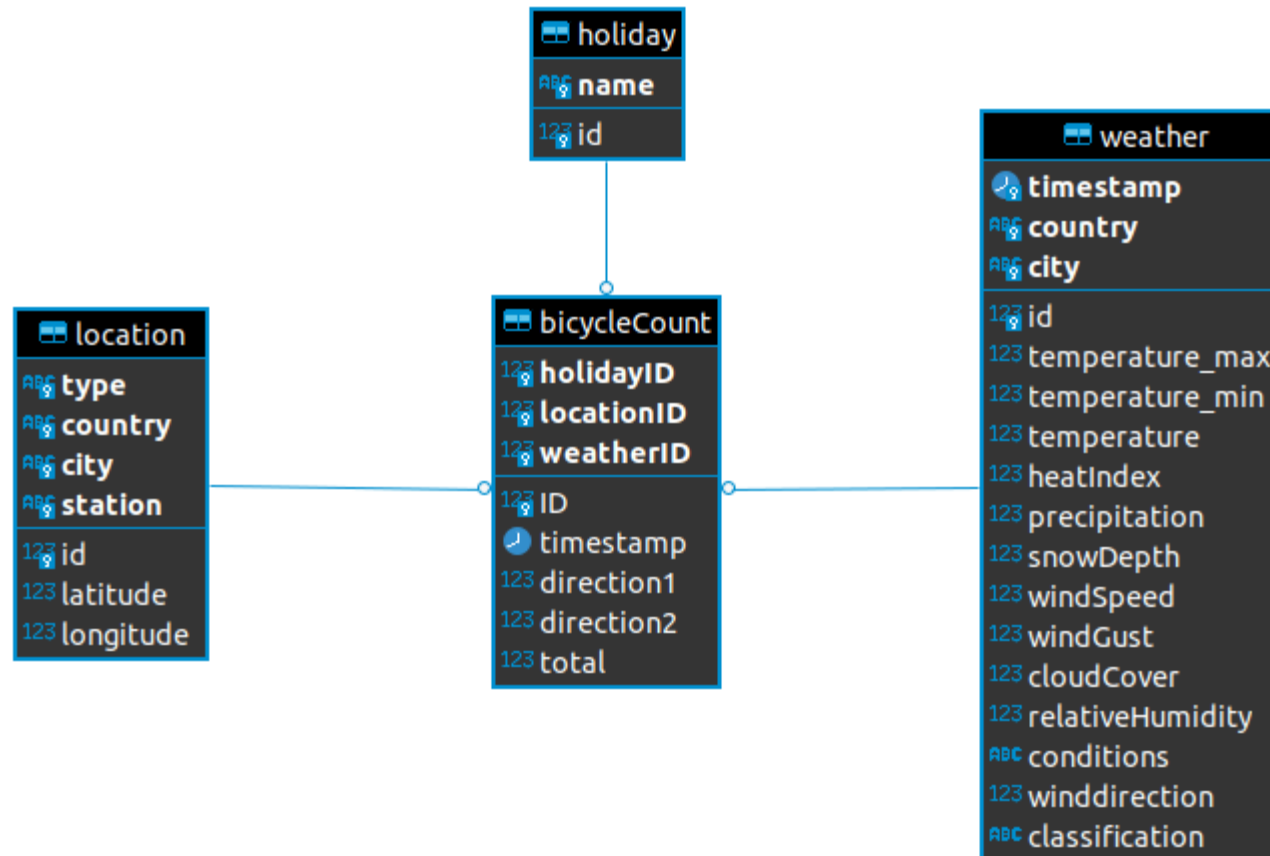
Different structures

	A	B	C	D	E	F	G
1	date	time start	time end	counting station	forward	reverse	total
2	2020.04.01.	00:00	00:15	Berchem TOT	0	0	0
3	2020.04.01.	00:15	00:30	Berchem TOT	0	2	2
4	2020.04.01.	00:30	00:45	Berchem TOT			
5	2020.04.01.	00:45	01:00	Berchem TOT			
6	2020.04.01.	01:00	01:15	Berchem TOT			
7	2020.04.01.	01:15	01:30	Berchem TOT			
8	2020.04.01.	01:30	01:45	Berchem TOT			

	A	B	C	D	E	F	G	H	I	J
1	Zeitraum	1. April 2018 -> 30. April 2018								
2										
3	Time	50101 Damm	50601_Nado	50201_Garte	heilig	50701_Pferd	50301_Roons	50401 Ofene	50501 Alexar	50801_Amal
4	Sonntag, 1. April 2018	3	0	6	6	6	1	9	2	C
5	Sonntag, 1. April 2018	6	3	4	7	3	0	10		C
6	Sonntag, 1. April 2018	4	0	3	3	5	0	11	0	C
7	Sonntag, 1. April 2018	10	0	2	2	1	1	12	2	C
8	Sonntag, 1. April 2018	4	0	4	4	0	1	6	4	C
9	Sonntag, 1. April 2018	4	9	5	14	3	0	8		C
10	Sonntag, 1. April 2018	9	0	6	6	5	0	6	0	C

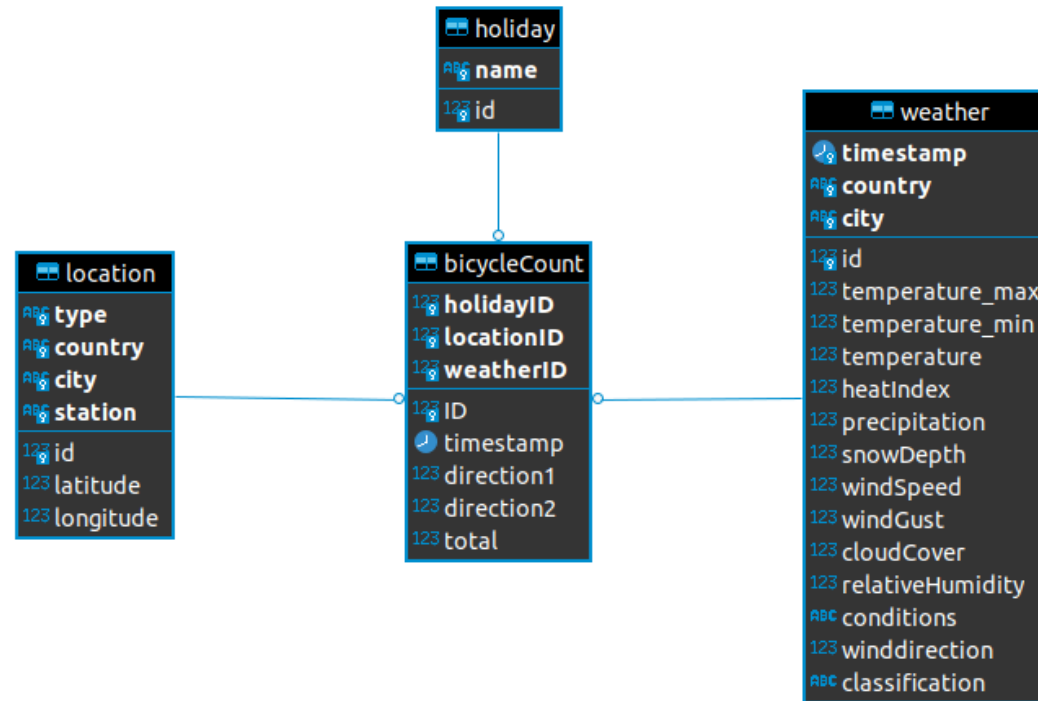
	A	B	C	D	E	F
1	Datum	K15 Berliner Stra�ve	K61 Christophorusweg	K51 Nikolausberger Weg	K63 Robert-Koch-Stra�ve	K13 Wender Landstra�ve
2	Montag, 13. April 2020	19	8	6	15	13
3	Montag, 13. April 2020	21	9	14	13	15
4	Montag, 13. April 2020	12	2	4	24	11
5	Montag, 13. April 2020	7	3	15	10	11
6	Montag, 13. April 2020	16	2	4	8	11
7	Montag, 13. April 2020	31	8	22	10	9
8	Montag, 13. April 2020	16	8	12	23	8
9	Montag, 13. April 2020	7	14	9	17	10
10	Montag, 13. April 2020	16	8	17	18	9





Structure

Extract



bicycledata.bicycleCount: 5.543.500 Zeilen gesamt (ungefähr), limitiert auf 1.000

ID	timestamp	holidayID	locationID	weatherID	direction1	direction2	total
226.068	2018-06-01 00:00:00	1	1	20.165	(NULL)	(NULL)	5
226.243	2018-06-01 00:15:00	1	1	20.181	(NULL)	(NULL)	12
226.424	2018-06-01 00:30:00	1	1	20.195	(NULL)	(NULL)	5
226.608	2018-06-01 00:45:00	1	1	20.210	(NULL)	(NULL)	7
262.430	2018-06-01 01:00:00	1	1	23.089	(NULL)	(NULL)	17
262.453	2018-06-01 01:15:00	1	1	23.091	(NULL)	(NULL)	5



BITS API

Review

- Development of a data warehouse
 - MariaDB as database
 - Structure of a star schema for data storage
- Scripts for
 - Entering the data
 - Provide initial analyses
- Connection of a weather API for additional data
 - Use of visualcrossingweather.com
 - Purchase of a licence to use the data



Current status

- Development of an API interface
 - To download raw data
 - csv files
 - json files
 - runs on the protected server
 - Using a proxy
 - connected to freely accessible website
- Remodel of the infrastructure for inserting data into the database
 - optimised codebase
 - Better folder structure for future expandability



Current process



Problems

- Visualcrossingweather
 - No rights to publish raw weather data
 - Sometimes has bugs that we have to report
 - but processed data can be freely published
- Server interface
 - IT services of the university provided proxy address
 - Time was lost because the problem was not ours
- Corona situation makes communication problems more difficult

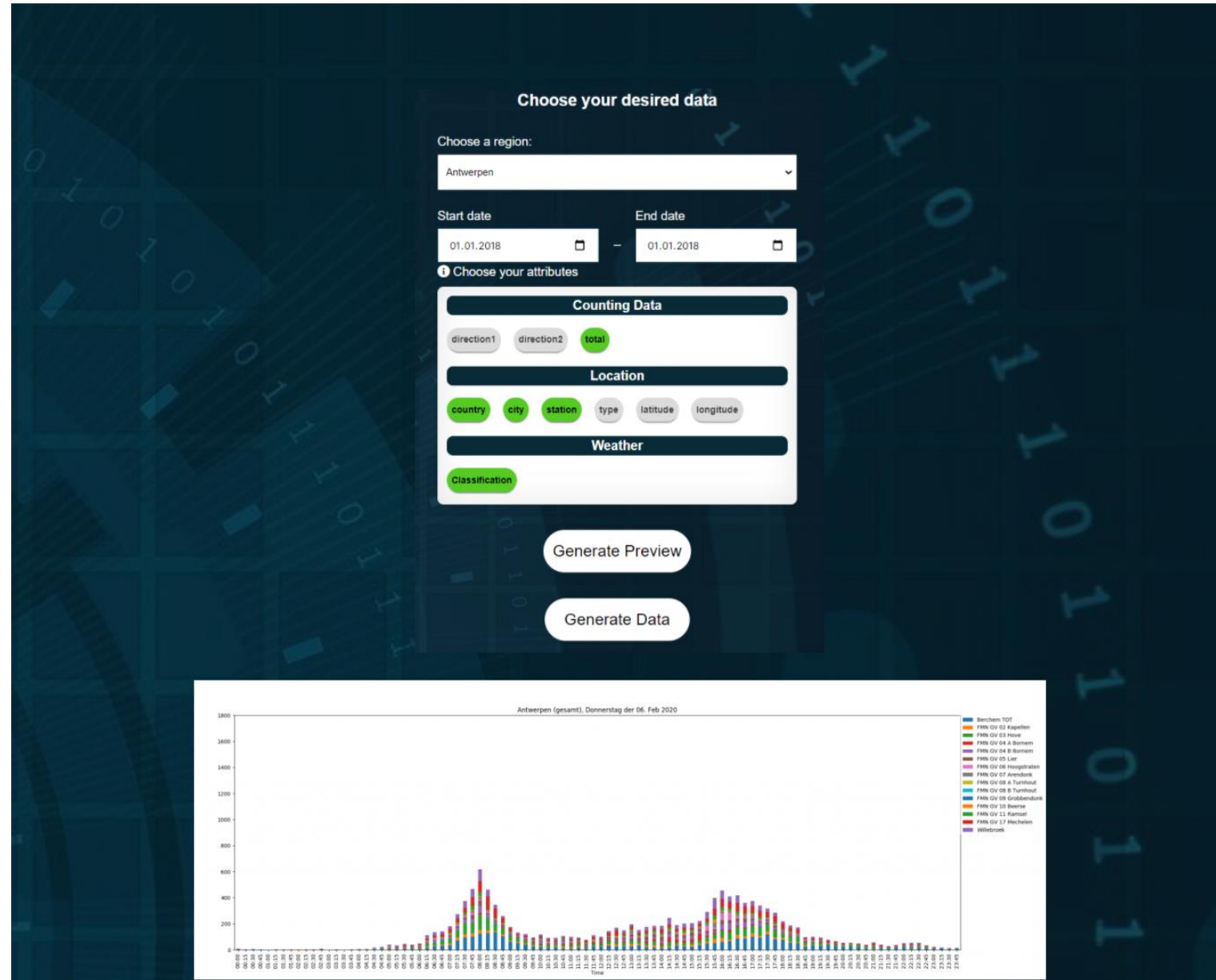


Outlook

- Further processing of raw data by the end of the year
- Analysis with Power BI and ArcGis Pro
- Development of the website
 - Calculation and provision of KPI's
 - Dynamic graphs
 - heat maps
 - Snifferbike
 - Weather data classifications



Private: Bicycle's Data Mockup



Mockups

Choose your desired data

Choose a region:
Snifferbike

Start date: 01.01.2018 — End date: 01.01.2018

Choose your attributes

Snifferbike

recording_time trip_sequence pm1_0 pm2_5 voc
voltage

Location

country city station type latitude longitude

Weather

Classification

Generate Preview

Generate Data

The map displays a heatmap of PM2.5 concentrations over a city area. A legend in the bottom-left corner shows a color scale from blue (Low) to red (High). The map features a grid of streets and several red and orange hotspots, indicating areas with higher PM2.5 concentrations. The Esri logo is visible in the bottom-right corner of the map.

- Integration via Folium / Leaflet
- especially Snifferbike



Thank you for your kind attention.

Are there any questions?

Get in touch

