EUMETNET E-AI initiative
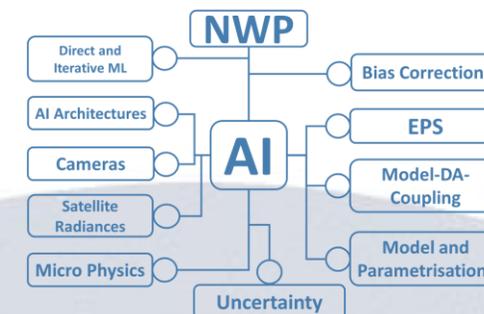
"Artificial Intelligence and Machine Learning for Weather, Climate and Environmental Applications" (E-AI) Optional Programme
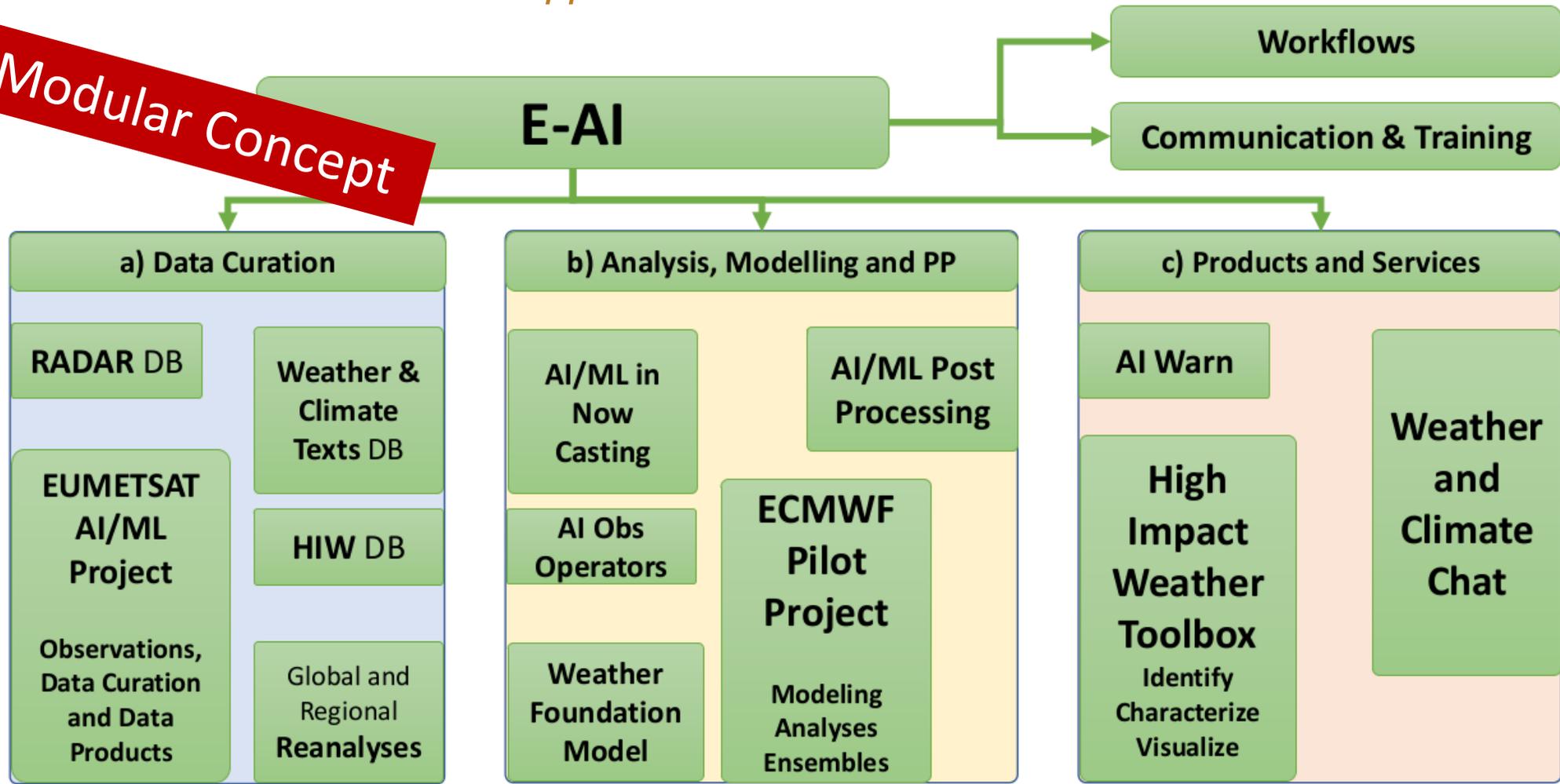
High level objectives:

- To enhance the collaboration of European NMHSs and external partners in the area of AI/ML in weather, climate and environment.

- To share the developments which take place under E-AI using a commonly-used permissive open-source licence.

Coordinating Member for E-AI: DWD (Roland Potthast)

E-AI: *Artificial Intelligence and Machine Learning in Weather, Climate and Environmental Applications*

Modular Concept

E-AI

Workflows

Communication & Training

a) Data Curation

RADAR DB

Weather & Climate Texts DB

EUMETSAT AI/ML Project

HIW DB

Observations, Data Curation and Data Products

Global and Regional **Reanalyses**

b) Analysis, Modelling and PP

AI/ML in Now Casting

AI/ML Post Processing

AI Obs Operators

ECMWF Pilot Project

Weather Foundation Model

Modeling Analyses Ensembles

c) Products and Services

AI Warn

High Impact Weather Toolbox

Identify Characterize Visualize

Weather and Climate Chat

Roland Potthast 2024

EUMETNET

3

E-AI Activity - 2024 Survey (Year 1)

Month	Activity
1	Establishing a structure and a pool of supporting NHMS and corresponding scientists, including EUMETSAT and ECMWF.
2	Establishing the Strategic Expert Team (SET) to coordinate the evolution of the Programme activities.
2-6	Workshops starting the three Modules of the Programme.
4-8	Establishing tools and a framework for communication.
9	Organizing three basic tutorials on AI/ML techniques and work environment.
8-12	Establishing working groups on AI/ML techniques along the whole value chain.
10-12	Raising further support from member states and collaborating institutions.

Participating Weather Services (2024)

18 NHMS, EUMETSAT, ECMWF:

- Spain,
- Slovenia,
- Denmark,
- Germany,
- Finland,
- Hungary,
- Poland,
- Netherlands,
- Ireland,
- France,
- Switzerland,
- Norway,
- United Kingdom,
- Italy,
- Sweden
- Slovenia
- Austria
- Czechia

About 260 participants, dedicated experts or points of contact

E-AI Strategic Expert Team (E-AI SET)

The <mark>E-AI Strategic Expert Team</mark> (E-AI SET) has been established and worked throughout the year. *The E-AI SET currently meets monthly.* The current E-AI SET members are:

1) Charles Ewen (Met Office)

2) Mark Liniger (MeteoSwiss) and Katrin Ehlert/Gabriela Aznar (ECMWF Pilot Project Lead)

3) Jordis Tradowsky (MET Norway) and Jorn Kristiansen (ECMWF Pilot Project Lead)

4) Andy Brown (ECMWF), Deputy Victoria Bennet

5) Matthieu Plu (Météo-France)

6) Paolo Ruti (EUMETSAT)

7) Jan Willem Noteboom (KNMI)

8) Roland Potthast (DWD, Programme Manager)

Strategic Experts!

EUMETNET

Three Initial Workshops in 2024

1. Initial Workshops EUMETSAT

There are initial workshops starting the activities in all three Modules. On initiative of Paolo Ruti we started with the EUMETSAT Workshop on AI/ML focusing on data curation on 19 – 20 February 2024 in Darmstadt. More than 30 experts on AI/ML from NHMSs and research institutes discussed the state of the field and worked out important steps for further development. The outcome of the workshop is summarised in the Annex.

## 2. The ECMWF Pilot Project Kick-Off

The ECMWF Machine Learning Partnership (ML PP) was launched during a 5-day workshop event hosted by ECMWF in Reading in June 2024. The workshop brought together about 50-60 participants from EUMETNET Members, NMHSs from ECMWF Member and Co-operating States, and EUMETSAT.

3. E-AI Workshop AI/ML for Products and Services

The workshop was held on June 5 and 6, 2024, at KNMI, the Netherlands. The workshop on products and services brought together about 120 participants. It suggested 10 working groups to be organized by E-AI as a joint activity of the weather services and their partners.
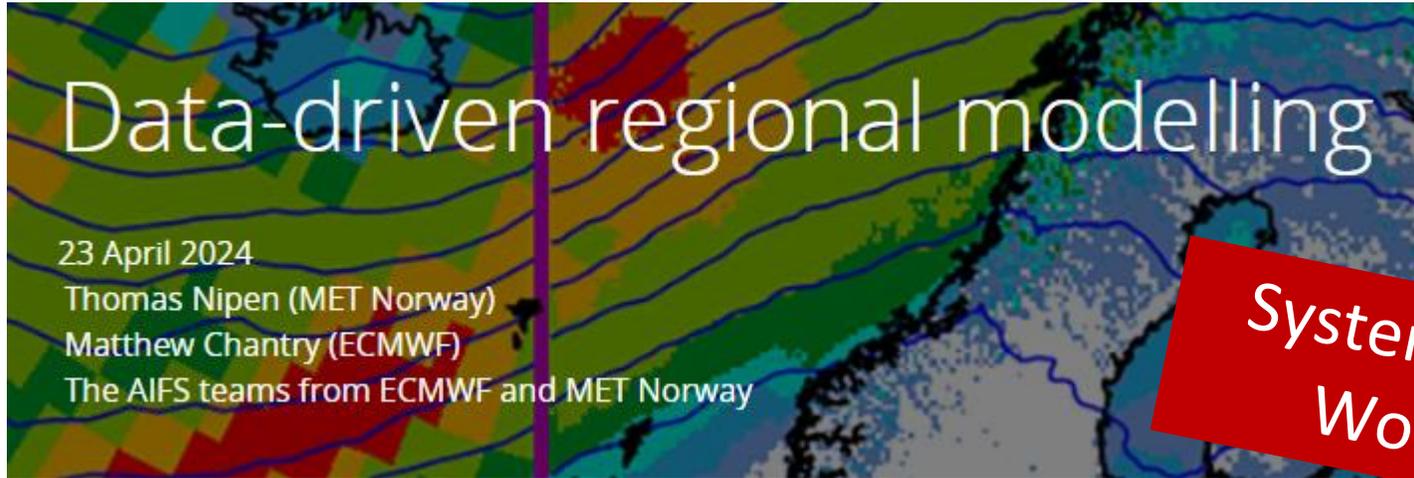
EUMETNET

E-AI Cafe

GeoSphere Austria

Activities @GeoSphere Austria using ML/AI – a brief overview

Aitor Atencia, Alexander Kann, Çağlar Küçük, Christian Möstl, Irene Schicker, Harald Schellander, Markus Dabernig, Pascal Gfäller, Petrina Papazek, Sandro Oswald, Sebastian Lehner, Thomas Muschinski
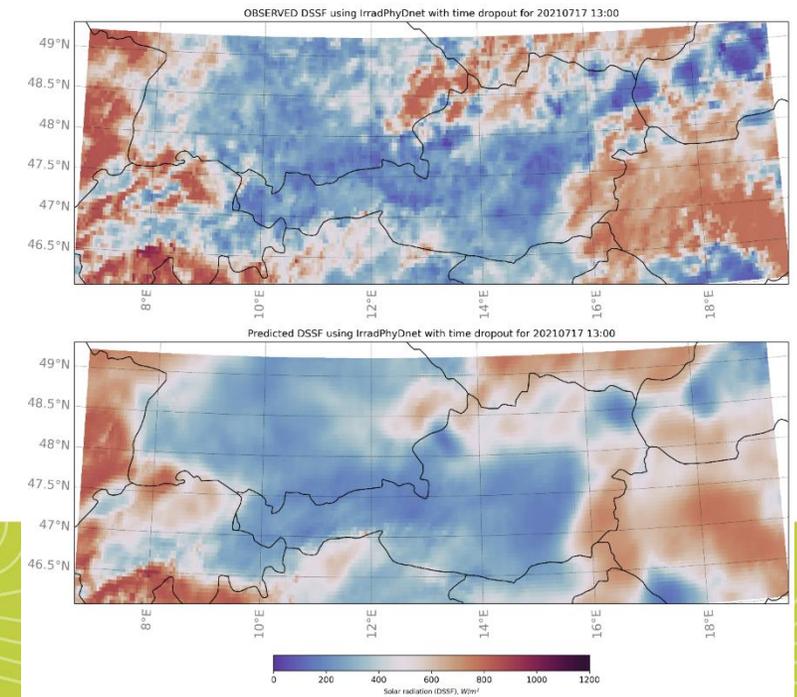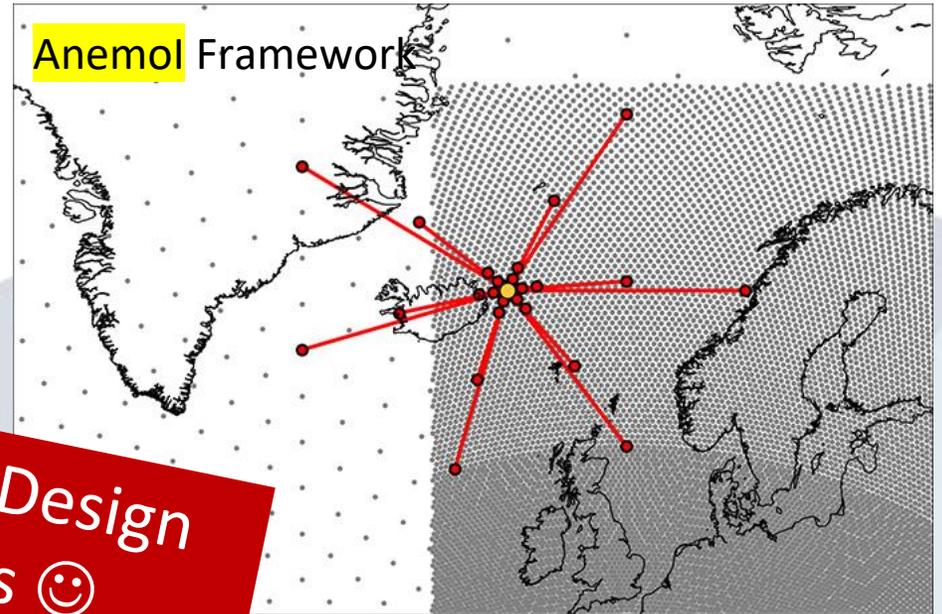
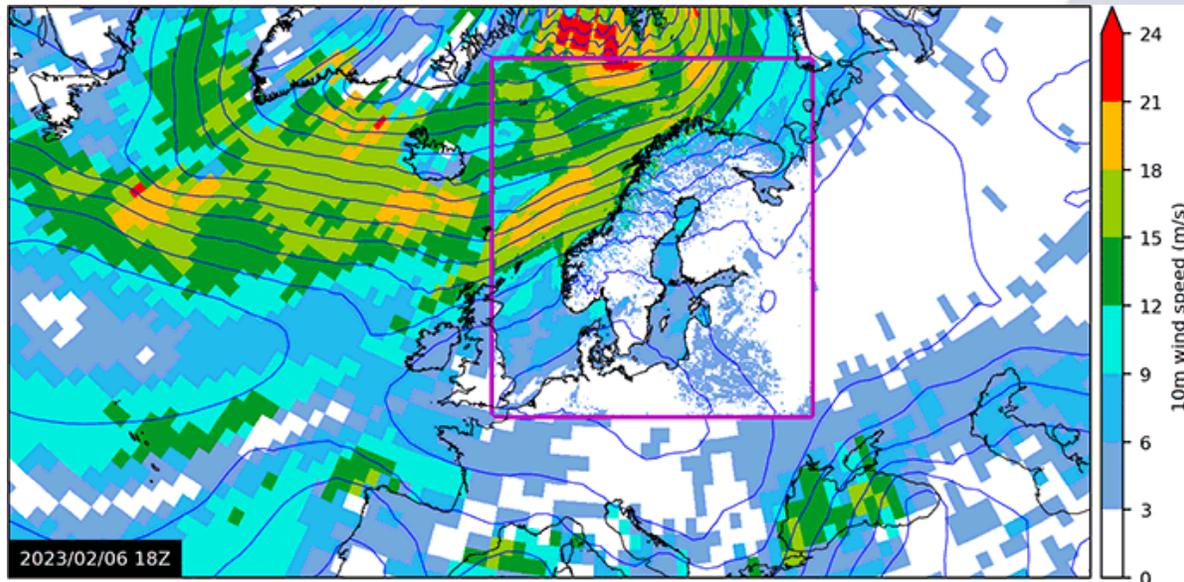May 8, 2024

Data Driven Regional Modeling

Data-driven regional modelling

23 April 2024
Thomas Nipen (MET Norway)
Matthew Chantry (ECMWF)
The AIFS teams from ECMWF and MET Norway

Anemol Framework

System Design Works ☺

2023/02/06 18Z

10m wind speed (m/s)

In a collaboration between MET Norway and ECMWF, they have adapted the tooling for the global AIFS approach, to take advantage of high-resolution data over the Nordics.

In an ECMWF pilot project on ML, led by MET Norway and MeteoSwiss and with contributions from 13 meteorological organizations across Europe, we will further develop these tools and explore together what the role of data-driven models is for weather prediction.

EUMETNET

Data Driven Global Forecasts

AIFS Forecasts

Pangu Weather

FourCastNet

GraphCast

Foundation Models

Neural LAM

AICON Forecast

System Design Works ☺

ICON Reanalysis

AICON Forecast 21h

02/2022 FourCastNet (NVIDIA)

11/2022 Pangu-Weather (Huawei)

12/2022 GraphCast (Google)

10/2024 AIFS (EZMW)

4/2024 AICON Prototype working (DWD)

MetOffice: **Moonshot Project AI4NWP** multi-model, multi-mode training

AtmoRep Foundation Model

MetCoOp: **Neural-LAM**

https://github.com/198808xc/Pangu-Weather
https://www.nature.com/articles/s41586-023-06185-3

https://github.com/google-deepmind/graphcast

Analysis, Modeling, Postprocessing

(1) AI Nowcasting

(2) AI Forecasting (NWP)

(1) AI Ensembles / Uncertainty

(2) AI Data Assimilation

https://arxiv.org/abs/2406.00390

(3) AI Postprocessing

Workshop activity of ECMWF in 2024 both for collecting input for the ==analysis and forecasting ECMWF pilot== project as well as on more general AI/ML based developments.

ECMWF Pilot Project

WP0/Co-ordination:
Jørn Kristiansen/ Katrin Ehlert

(Project co-lead Norway/Switzerland)

WP1: Data-driven forecasting
Co-leads: M. Plu, L. Denby, O. Fuhrer

WP2: Ensemble Forecasting
Co-leads: J. Keller, H. Frederiksen

WP3: Data assimilation
Co-leads: M. Mile, J. Keller

WP4: Infrastructure and MLOps
Co-leads: Kasper Hintz, MeteoSwiss(tbd)

WP5: Training and support
Co-leads: ECMWF, L. Denby

Truth with Obs at eval step 008

Analysis at eval step 008

EUMETNET

Working Groups on Products and Services

Group	Description
WG 1	AI Data Readiness and Curation (e.g. Data Sets Preparation and Management, Standards, IPR, Labeling)
WG 2	LLM Implementation and LLM Services
WG 3	AI Operationalization, Maintenance and QC, MLOps
WG 4	AI Sensing and Novel Observations (e.g Phenomena and Identification, Webcams and Co.)
WG 5	AI based Information Condensation and Communication, AI based HIW Warnings, Personalization
WG 6	Nowcasting, Obs2Forecast, and Data Fusion, Downscaling, PostProcessing (together with NWC and PP Programme of EUMETNET)
WG 7	User Optimized Products and Impact (e.g. Cross Border Aviation Products)
WG 8	Ethics Questions
WG 9	AI Frameworks
ATG	Anemoi Technical Group (ATG), established, weekly meetings

EUMETNET

Kick-Off Meetings

- The Kick-off meeting of the working group WG01_Data_Curation on data management and curation is scheduled for Montag, 14. October 2024 13:00 – 15:00, facilitated by Roope Tervo, EUMETSAT, Arianna Valmassoi, DWD and Stephan Siemen, ECMWF.

- The Kick-off meeting of the working group WG02_LLM on the use of large language models is scheduled for Wednesday, 2nd October 2024, 2:00pm - 4:00pm (CEST) facilitated by Frank Guibert, MeteoFrance and Johannes Schenk, DWD

- The Kick-off meeting of the working group WG03_Operations on MLOps and ML Frameworks October 15, 2024, 3:00pm - 5:00pm (CEST) facilitated by Frank Guibert, MeteoFrance and Gabriela Aznar, MeteoSchwiss

- A first initiative meeting of a subgroup of WG06_NWC (MLCast) has been taken place at September 23, 2024. A Kick-Off Meeting is scheduled for Oct 21, 2024, 13-14:00 CEST facilitated by Irene Kruse, Geosphere Austria, Irene Schicker and Leif Denby, DMI.

Further Kick-Off Meetings are being prepared …

EUMETNET

Material / Tutorials

To get access to
EUMETNET **SharePoint**
write an email to
marek.Jacob@dwd.de

EUMETNET
E-AI Sharepoint

https://tlnt19059.sharepoint.com/sites/E-AI/SitePages/CollabHome.aspx

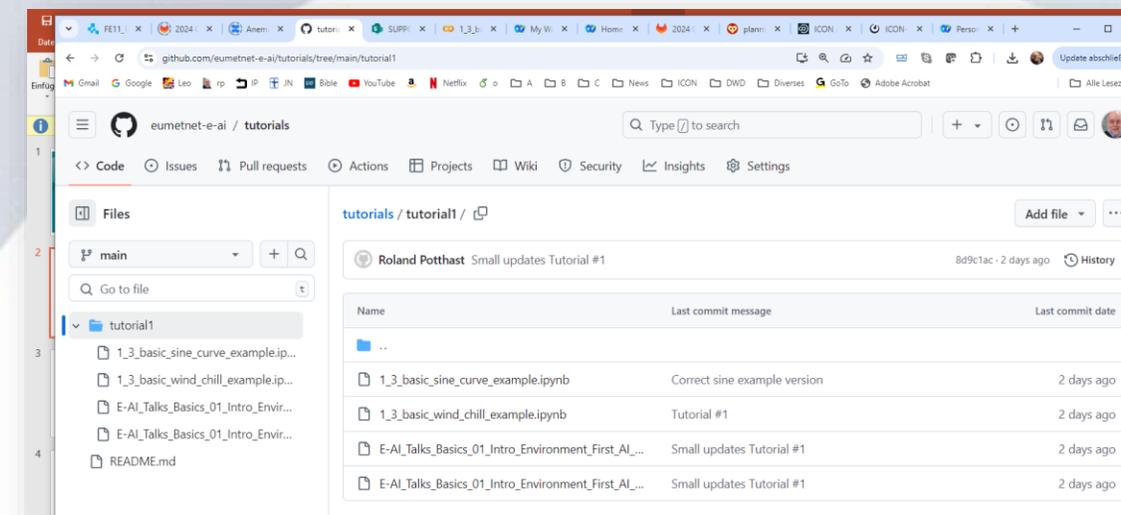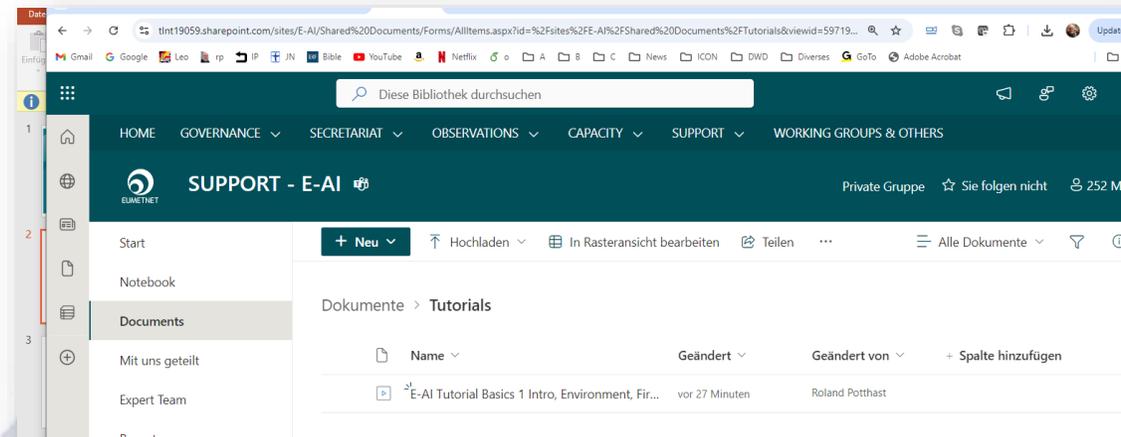**Github eumetnet-e-ai Repo**

https://github.com/eumetnet-e-ai/tutorials/

Recordings available
on Sharepoint

Codes and Slides
available on Github

EUMETNET

e) Communication and Training

Communication

Communication and Training for Development and Employment of Machine Learning Techniques

AI/ML Intro Talks

AI/ML Café

Series contributed by Members

Leadership Team to be built

- EUMETSAT Feb 2024

AI/ML Workshops

- ECMWF Pilot project June 2024

- E-AI Products & Services June 2024

AI/ML Hackathons

Under Discussion

E-AI Cafe Meeting at May 8, 13-15 CET: ML/AI Activities @GeoSphere Austria – a brief overview

AI/ML Training Course — *ECMWF Standard*

AI/ML Topical Trainings — *E-Learning Approach*

Link to EUMETCAL established

- General Intro to AI/ML
- Learning AI/ML via Python
- TensorFlow
- PyTorch
- Regression Problems
- Model Emulation
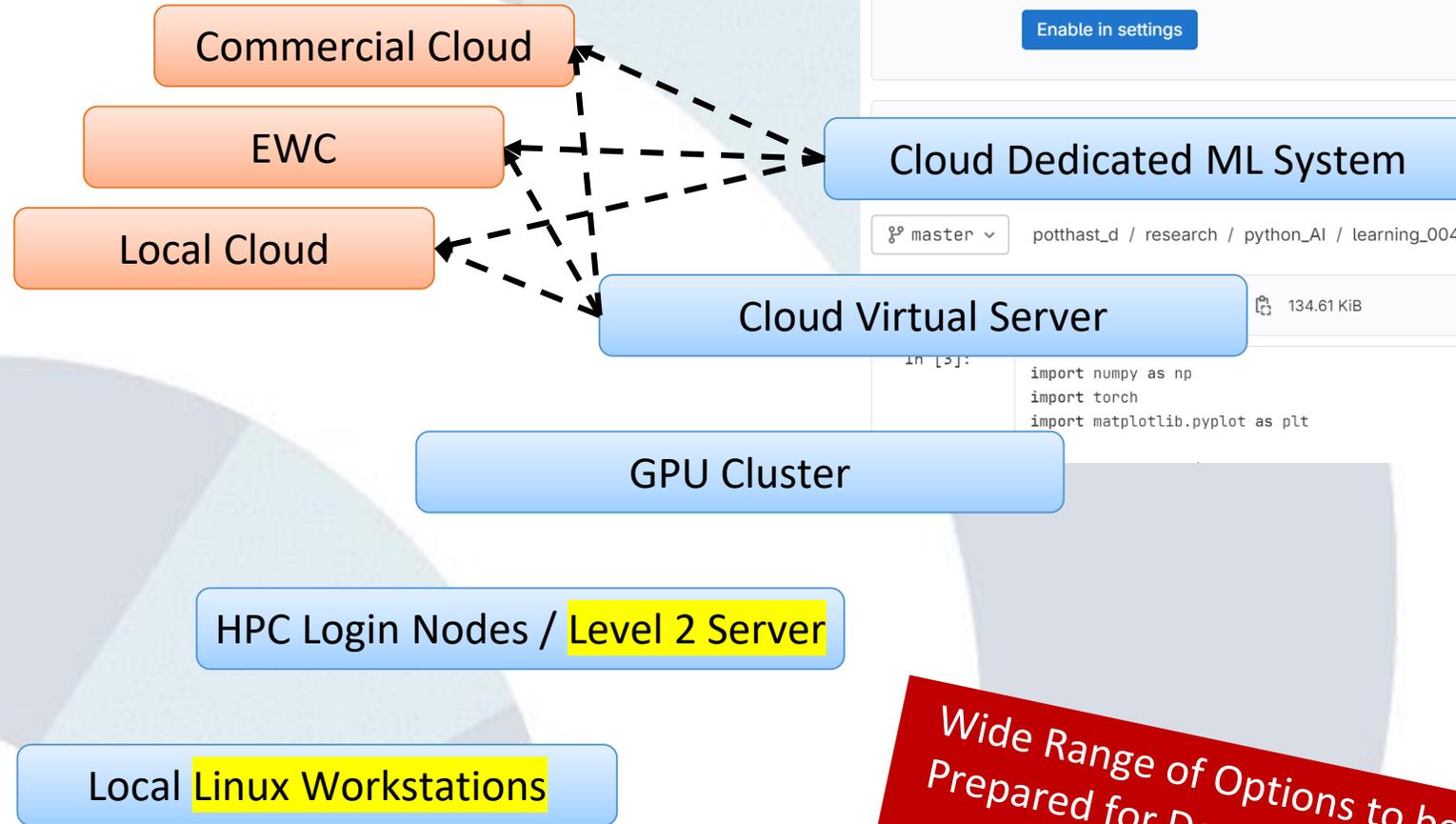- LLM Use, Options, Adaptation
- Pipelining
- Using Cloud Resources

EUMETNET

E-AI Basic Tutorials

1. Tutorial E-AI Basics 1: September 16, 2024, 11-12 CEST
2. Intro, Environment, First Example
3. 1.1 Basic Ideas of AI Techniques (20') [SH]
4. 1.2 Work Environment (20') [RP]
5. 1.3 First Example for AI - hands-on (20') [JK]

6. Tutorial E-AI Basics 2: September 18, 2024, 13-14 CEST
7. Dynamics, Downscaling, Data Assimilation Examples
8. 2.1 Dynamic Prediction by a Graph NN (20') [RP]
9. 2.2 Downscaling via Encoder-Decoder (20') [SH]
10. 2.3 AI for Data Assimilation (20') [JK]

11. Tutorial E-AI Basics 3: September 23, 2024, 11-12 CEST
12. LLM Use, Transformer Example, RAG
13. 3.1 Intro to LLM Use and APIs (20') [RP]
14. 3.2 Transformer for Language and Images (20') [JK]
15. 3.3 LLM Retrieval Augmented Generation (RAG) (20') [SH]

AI generated Images,
© Roland Potthast 2024

EUMETNET

Workflows

Workflow Scenarios for Development and Employment of Machine Learning Techniques

Towards a hierarchy of established Workflows

Wide Range of Options to be Prepared for Developers and Users

- Commercial Cloud
- EWC
- Local Cloud
- Cloud Dedicated ML System
- Cloud Virtual Server
- GPU Cluster
- HPC Login Nodes / Level 2 Server
- Local Linux Workstations
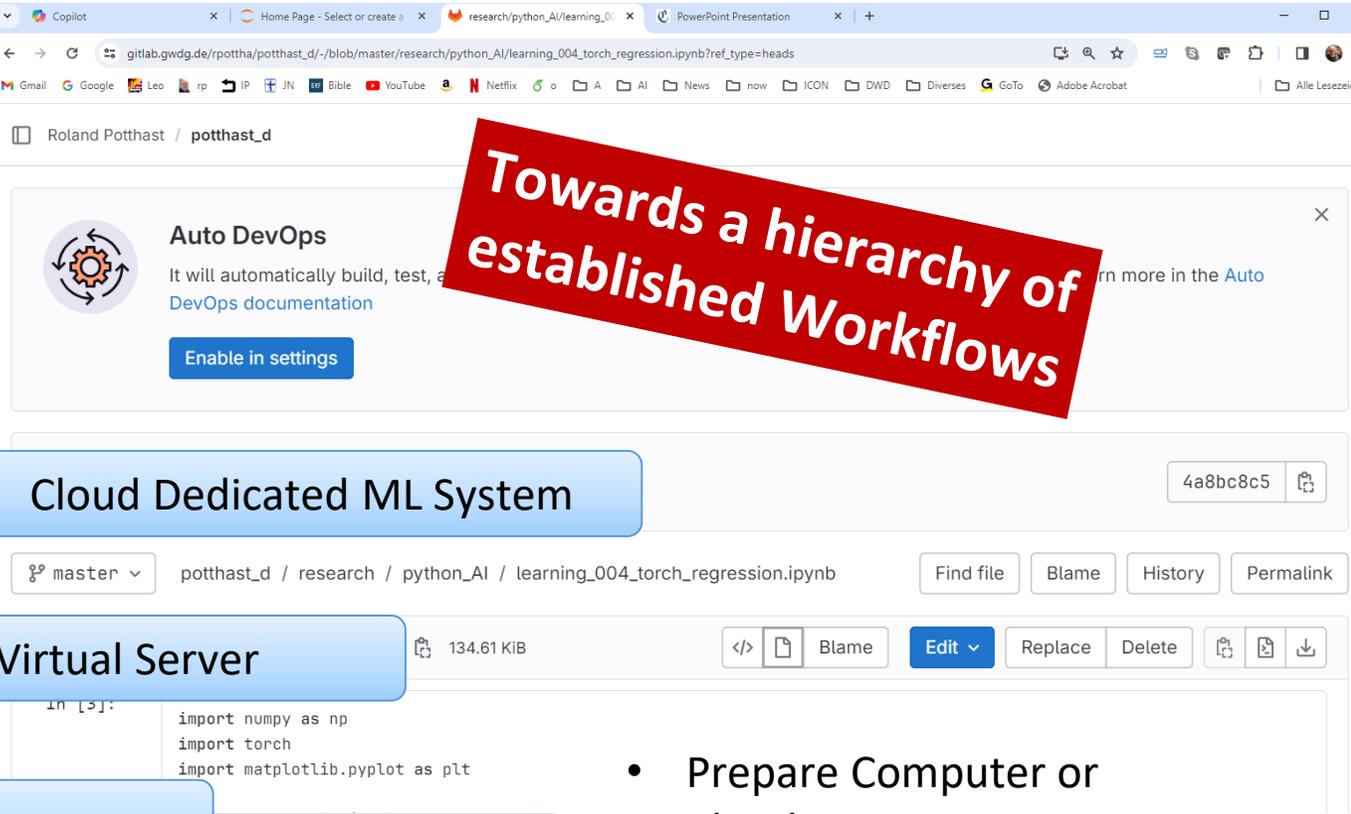
- Prepare Computer or Cloud Access
- Make Libraries Available
- Manage Data
- Enable Software Systems
- Develop Applications
- Test and Operationalize

EUMETNET

Graph Neural Networks – Basic Idea 1/2

Graph Neural Networks (GNNs) are a type of neural network designed to work on graph-structured data, where entities (nodes) are connected by relationships (edges). Unlike traditional neural networks, GNNs use graph structures to capture both the features of individual nodes and their connectivity within the graph.

By passing information between nodes through edges, GNNs allow for the modelling of complex relationships, making them effective for tasks like social network analysis, molecular modelling, and recommendation systems.

GNNs leverage techniques like message passing to aggregate and update node representations based on their neighbours.

Graph structure

grid to mesh - ICON

cell values from adjacent triangles
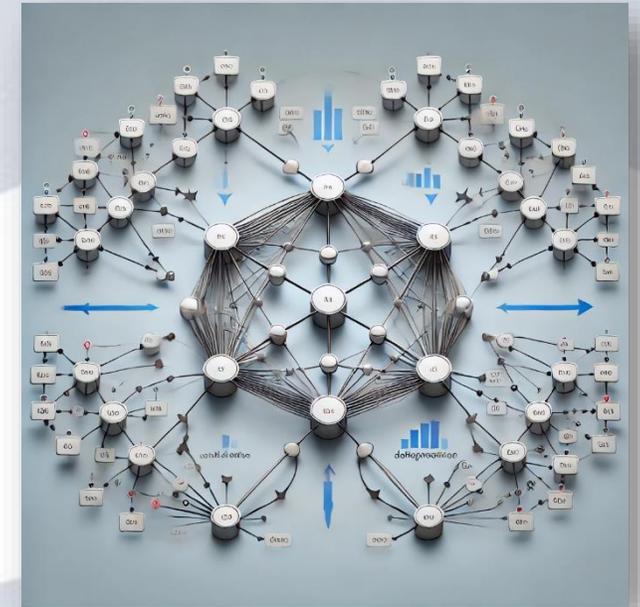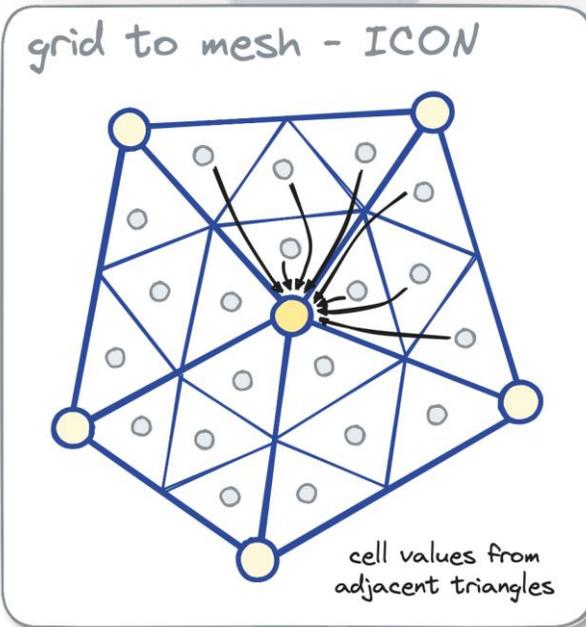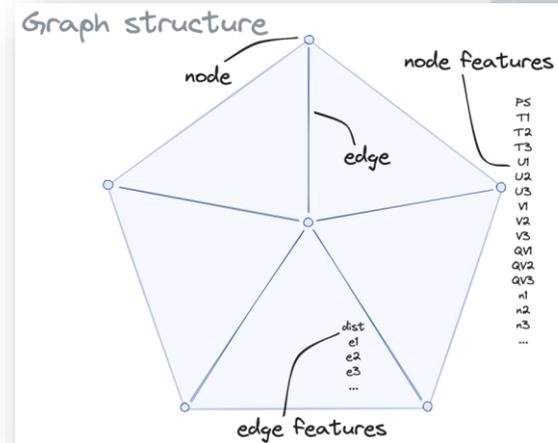
To construct a graph, start by defining its **nodes (vertices)**, which represent individual entities in your data. Then, establish **edges** between these nodes, representing the relationships or connections between the entities, forming the structure of the graph.

Reconstruct Missing Data

- We can ==remove some parts== of the patterns

- Let the autoencoder ==reconstruct== the missing parts.

- This can also be added to the ==training loop== to prepare the system.

| Original	Input	Reconstruction

19
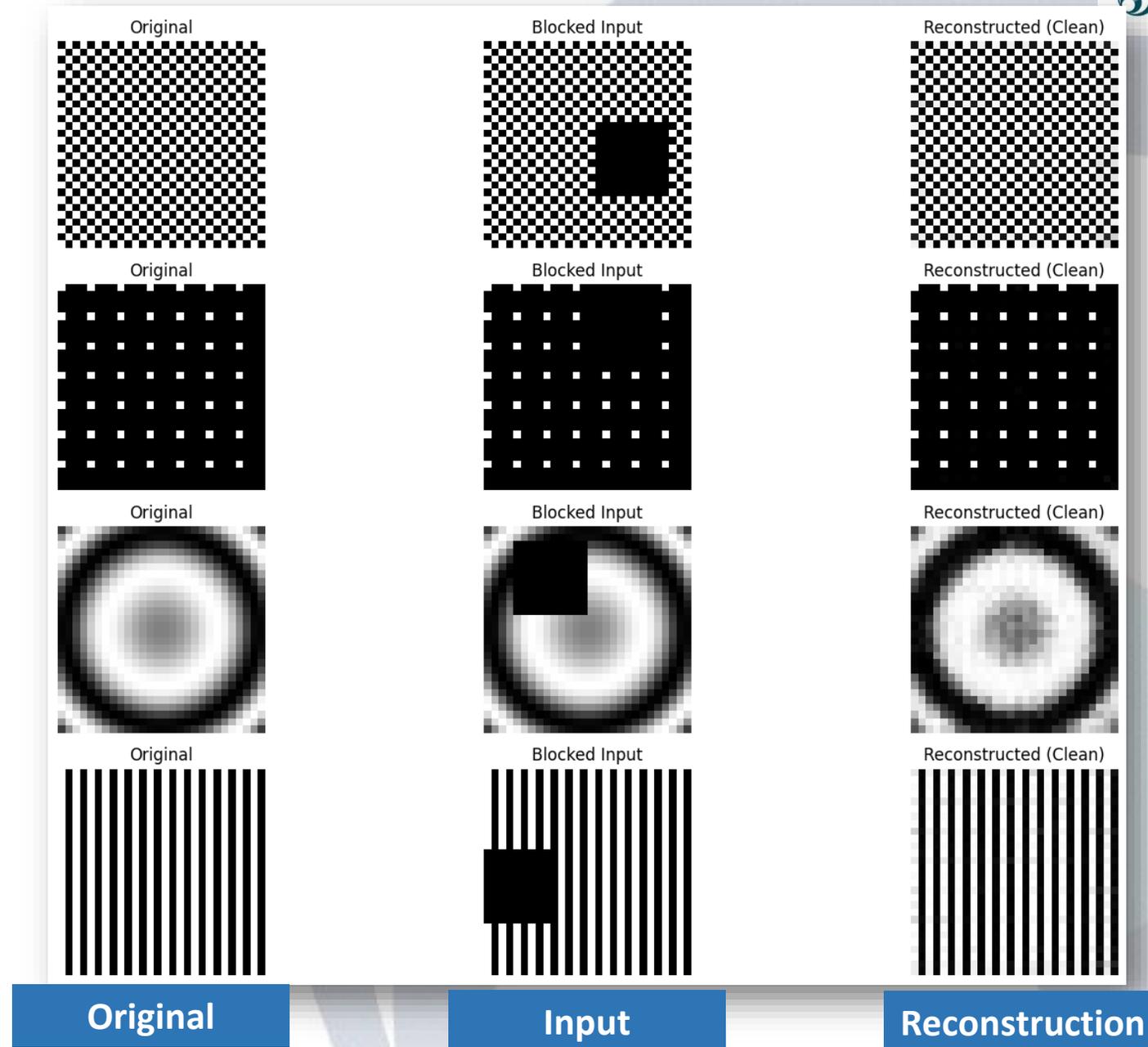
Data Assimilation Results

```python
output=model(train_data)
fig,axs=plt.subplots(2, 2, figsize=(10, 8))
for i in np.arange(4):
    ax=axs[int(i/2),i%2]
    ax.plot(z[train_data["no"].detach().numpy()[i*100],:],color="#777777",linewidth=2)
    ax.plot(train_data["fg"].detach().numpy()[i*100,:],color="#0088bb",linewidth=2)
    ax.scatter(train_data["idx"].detach().numpy()[i*100,:]*nx,train_data["obs"].detach().numpy()[i*100,:],color="#0088bb")
    ax.plot(output.detach().numpy()[i*100,:],color="#dd7700aa",linewidth=2)
fig.tight_layout()
```
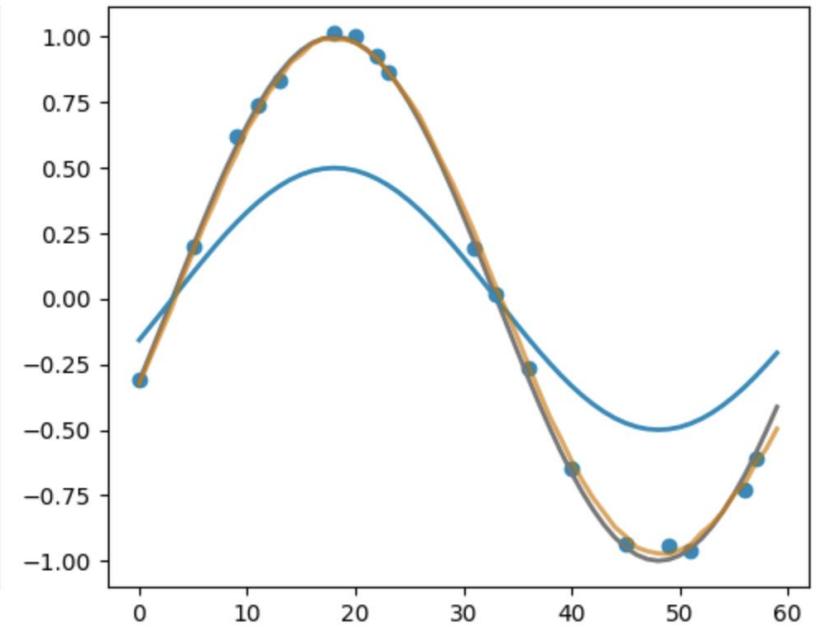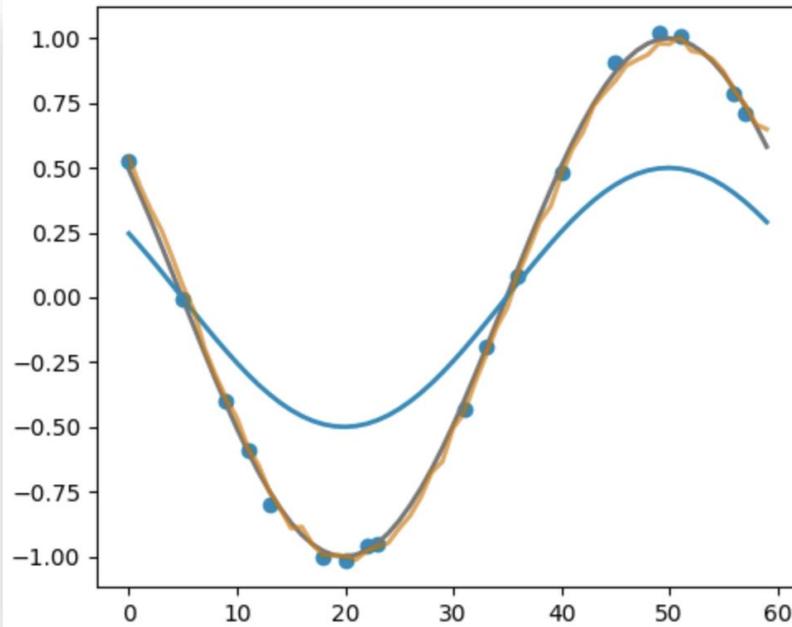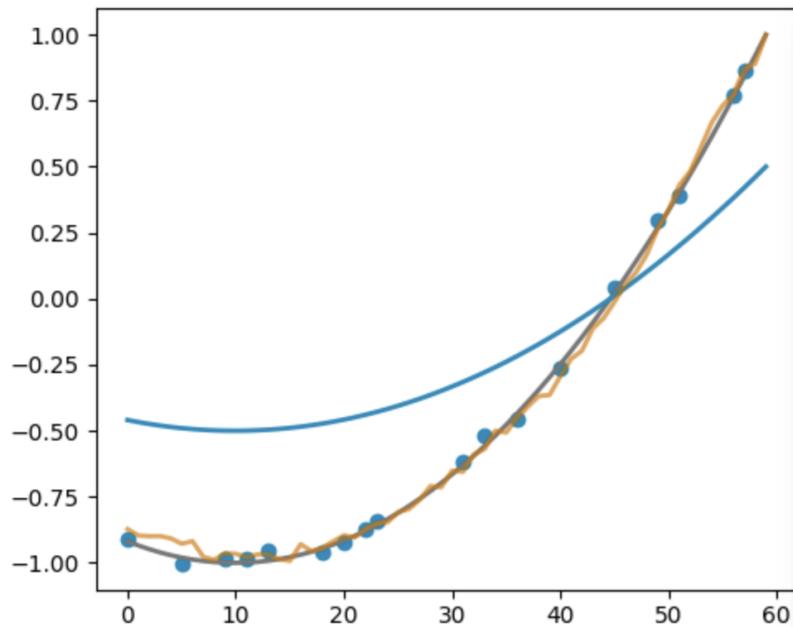
Large Language Models (LLMs)

- LLMs are models based on deep learning techniques, specifically neural networks, to process and understand patterns in the text.
- LLMs are trained on massive datasets that include books, websites, and other text sources.
- Through training, LLMs learn how words, sentences, and paragraphs relate to each other, enabling them to understand and generate coherent text.
- After training, LLMs can generate responses, summarize information, and assist with various language tasks based on user input.

Retrieval-Augmented Generation

Retrieval-Augmented Generation (RAG)
is an approach that combines information retrieval from ==external knowledge sources== with a ==language model== to generate more informed and accurate responses by ==retrieving== relevant documents or data before generating an answer.

Core Components for RAG are

- **LLM**: The pre-trained language model (e.g., BERT, GPT) that generates the responses. It has been trained on a broad, general corpus during training.

- **Corpus**: The external, domain-specific information provided to the RAG system to improve retrieval and enhance the LLM's output. The LLM uses this corpus at inference time to "look up" relevant information rather than relying solely on its pre-trained knowledge.

Transformers for images
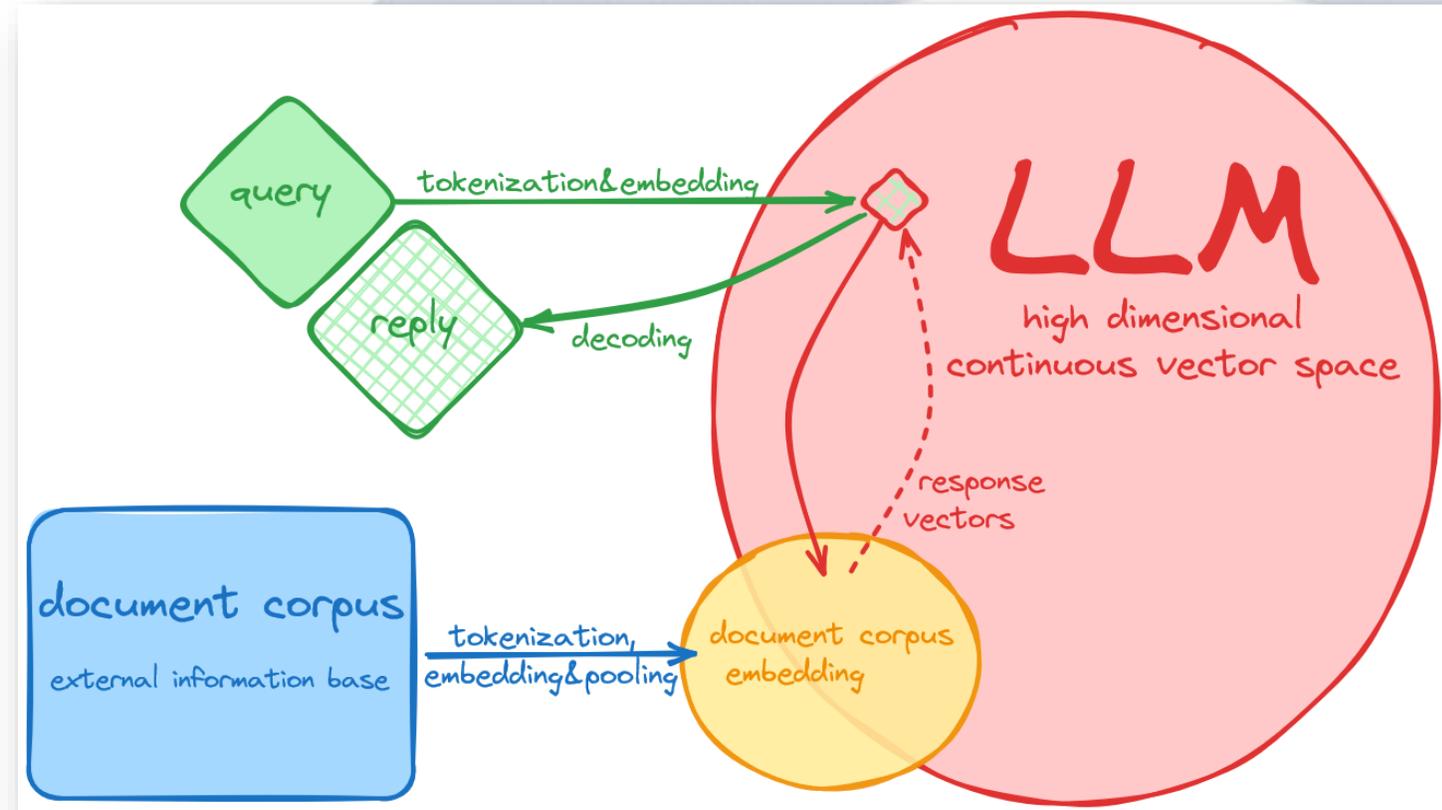
Explainable AI as Important Branch of Development

Visualization of attention as a heatmap

Dog →

Cat →

Elephant →

Zebra →
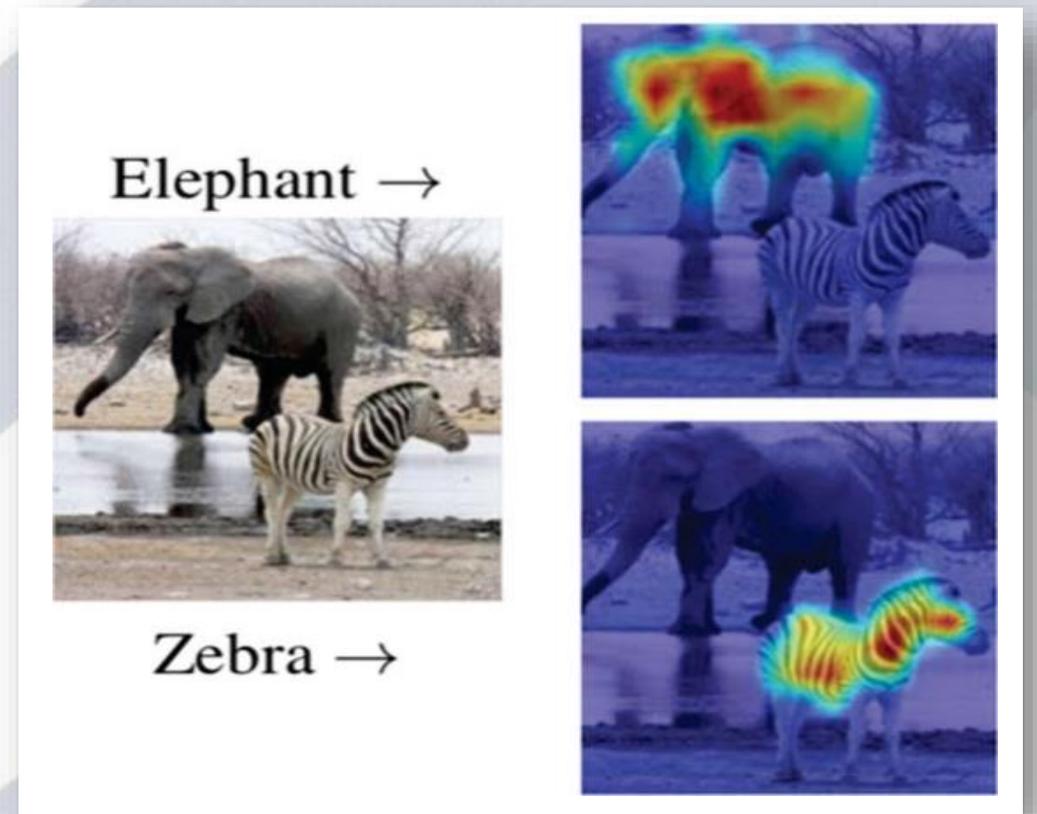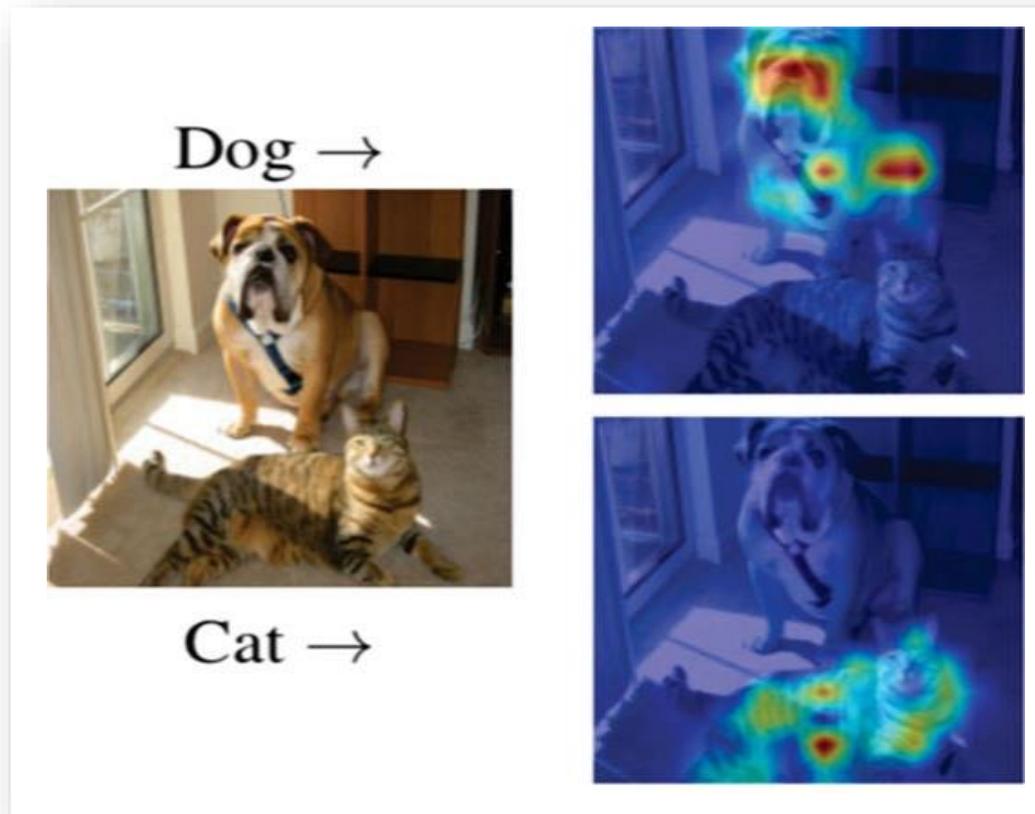
Adaption von Xu et al. (2022)

EUMETNET

A Series of Tutorials on Working with AI Tools V1.0

Python Basics	Model Fields, ECcodes & Cartopy	Putting LLM into Action	Pipelines and ML Flow
Linux → HPC GPU ← Cloud	Observations & Feedback Files	AI Frameworks Intro	Training & Fine Tuning LLM
Jupyter und Webserver	Xarray & Co.	NN Architectures	AI Model Forecasting
Container as Tools	Verification	AI for Grids	AI Data Assimilation
	xxx	AI for Time Series	AI Ensemble Prediction

Tutorials under Development

EUMETNET

Perspectives Weather, LLMs and AI/ML Techniqes

- **Huge diverse area, many techniques and architectures!**

- **Changes the way we work**

- **Changes customer relationships**

- **Changes products**

- **Changes services**

We have a lot to give here ☺

EUMETNET

CONTACT DETAILS

Prof. Dr. Roland Potthast

E-AI Programme Manager

EIG EUMETNET
European Meteorological Services' Network
www.eumetnet.eu

EIG EUMETNET Secretariat
c/o L'Institut Royal Météorologique de Belgique
Avenue Circulaire 3
B-1180 Brussels, Belgium

Registered Number 0818.801.249 - RPM Bruxelles

Tel: 0049 69 8062 2713

Email: Roland.Potthast@dwd.de

EUMETNET