MAchinE Learning for Scalable meTeoROlogy and climate

MAELSTROM

Dissemination Workshop I

Jan Mirus, Peter Dueben

www.maelstrom-eurohpc.eu

MAELSTROM

D4.7 Dissemination Workshop I

Author(s):	Peter Dueben (ECMWF) Jan Mirus
Dissemination Level:	Public
Date:	31/03/2022
Version:	1.0
Contractual Delivery Date:	31/03/2022
Work Package/ Task:	WP4/ T4.7
Document Owner:	ECMWF
Contributors:	All Partners
Status:	Final

MAELSTROM

Machine Learning for Scalable Meteorology and Climate

Research and Innovation Action (RIA)
H2020-JTI-EuroHPC-2019-1: Towards Extreme Scale Technologies and Applications

Project Coordinator:	Dr Peter Dueben (ECMWF)
Project Start Date:	01/04/2021
Project Duration:	36 months

Published by the MAELSTROM Consortium

Contact:
ECMWF, Shinfield Park, Reading, RG2 9AX, United Kingdom
Peter.Dueben@ecmwf.int

The MAELSTROM project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 955513. The JU receives support from the European Union's Horizon 2020 research and innovation programme and United Kingdom, Germany, Italy, Luxembourg, Switzerland, Norway

EuroHPC
Joint Undertaking

Contents

1 EXECUTIVE SUMMARY 5

2 INTRODUCTION 6

2.1 ABOUT MAELSTROM 6

2.2 SCOPE OF THIS DELIVERABLE 6

2.2.1 OBJECTIVES OF THIS DELIVERABLE 6

2.2.2 WORK PERFORMED IN THIS DELIVERABLE 7

2.2.3 DEVIATIONS AND COUNTER MEASURES 7

3 PLANNING, PREPARATION, AND REALIZATION OF THE WORKSHOP 8

3.1 SCHEDULING 8

3.2 INFRASTRUCTURE 8

3.3 TIMETABLE 9

3.4 PARTICIPATION 10

4 RESONANCE AND RESULTS 11

4.1 OVERALL RESONANCE AND QUESTIONS 11

4.2 POLL RESULTS AND CONCLUSIONS 11

5 CONCLUSION 15

6 ANNEX: MAELSTROM PRESENTATION DECKS 17

6.1 INTRODUCTION: PETER DUEBEN, ECMWF 17

6.2 WP1: GONG BING, JSC 19

6.3 WP2: FABIAN EMMERICH, 4CAST 21

6.4 WP3: DANIELE GREGORI, E4 22

6.5 WP3: ANDREAS HERTEN, JSC 24

1 Executive Summary

On 28th March 2022, the first of two dissemination workshops was held with the objective not only to share the progress of project MAELSTROM with the high-performance computing, machine learning and weather & climate communities, but also to collect input on the needs of those communities. While restrictions due to the COVID pandemic are being lifted, it was still necessary to choose an online format for this workshop. Registrations soared up to over 200; over 60 people gathered in the virtual room.

The first part of the workshop was dedicated to Project MAELSTROM; each of the work packages introduced its motivation and goals and gave an overview on the achievements to date:

- WP1: Machine learning applications & datasets
- WP2: Machine learning workflow tools
- WP3: Hardware Systems

The second part of the workshop provided a stage for three of our EuroHPC partner projects:

- TimeX: Time parallelization for eXascale computing
- DEEP-SEA: Programming Environment for European Exascale Systems
- RED-SEA: Network Solution for Exascale Architectures

The third part of the workshop featured external talks, giving insight into technologies with possible synergies or use cases within MAELSTROM's domains:

- MeteoSwiss: Time-consistent downscaling of atmospheric fields with GANs
- Pangeo: An open-source ecosystem for data-intensive science
- NVIDIA: Deep Learning for Earth Sciences in the HPC Context

Each part of the workshop led into an interactive phase, where a discussion was moderated by MAELSTROM; after the first part, feedback was collected in a series of polls.

The workshop was concluded MAELSTROM-internally, discussing conclusions and next steps.

2 Introduction

2.1 About MAELSTROM

To develop Europe's computer architecture of the future, MAELSTROM will co-design bespoke compute system designs for optimal application performance and energy efficiency, a software framework to optimise usability and training efficiency for machine learning at scale, and large-scale machine learning applications for the domain of weather and climate science.

The MAELSTROM compute system designs will benchmark the applications across a range of computing systems regarding energy consumption, time-to-solution, numerical precision and solution accuracy. Customised compute systems will be designed that are optimised for application needs to strengthen Europe's high-performance computing portfolio and to pull recent hardware developments, driven by general machine learning applications, toward needs of weather and climate applications.

The MAELSTROM software framework will enable scientists to apply and compare machine learning tools and libraries efficiently across a wide range of computer systems. A user interface will link application developers with compute system designers, and automated benchmarking and error detection of machine learning solutions will be performed during the development phase. Tools will be published as open source.

The MAELSTROM machine learning applications will cover all important components of the workflow of weather and climate predictions including the processing of observations, the assimilation of observations to generate initial and reference conditions, model simulations, as well as post-processing of model data and the development of forecast products. For each application, benchmark datasets with up to 10 terabytes of data will be published online for training and machine learning tool-developments at the scale of the fastest supercomputers in the world. MAELSTROM machine learning solutions will serve as blueprint for a wide range of machine learning applications on supercomputers in the future.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverable

D4.7, the first dissemination workshop, aims to

- generate awareness in the scientific and technological communities of weather & climate science, machine learning and high-performance computing for the project Maelstrom, its mission, objectives, and products
- share MAELSTROM's progress and achievements after nearly one year of project runtime with the said audiences
- collect valuable input and feed-back, showing us where our products and outcomes resonate with the community, correspond to unmet needs, and are likely to spawn use cases.

2.2.2 Work performed in this deliverable

The work to create the exploitation plan included collection of feedback from the partners in form of questionnaires and the identification of the relevant aspects pertaining to exploitation.

2.2.3 Deviations and counter measures

No deviations occurred; hence no counter measures were requisite.

3 Planning, preparation, and realization of the workshop

3.1 Scheduling

The date for the workshop, Monday, 20th March 2022, was selected to join forces with the 4-day machine learning workshop held by ECMWF on the following four workdays of this week[1].

3.2 Infrastructure

Because of COVID restrictions, respectively because of the uncertainty that had existed beforehand about the regulations, it was advisable to hold this workshop online in a virtual room. Registration was required, to allow us to gage the size of the audience and adjust our protocol accordingly.

The registration page and mechanism, just like the technical support for the webinar infrastructure, was shared with the said ECMWF workshop; an efficient solution for MAELSTROM.
The workshop was announced on Twitter and on the MAELSTROM project website.

The workshop was held with Zoom Webinar, a tool specialized for larger audiences. Attendees were able to ask questions in a dedicated Q&A area. Another tool, Slido, was used to create live polls. The workshop has been recorded with the consent of all participants.

[1] https://events.ecmwf.int/event/294/

3.3 Timetable

All times are GMT.

Part 1	**Learn about MAELSTROM**
8:30 → 08:50	**Introduction** Peter Dueben (ECMWF)
08:50 → 09:10	**WP1 Summary on Applications and Data Sets** Bing Gong (JSC)
09:10 → 09:30	**WP2 Summary on Software Tools** Greta Denisenko (4-cast)
09:30 → 09:50	**WP3 Summary on Hardware Benchmarks** Andreas Herten (JSC) and Daniele Gregori (E4)
09:50 → 10:30	**Possibility for general discussion**
10:30 → 11:00	**Coffee break**
Part 2	**EuroHPC Partner Project Talks**
11:00 → 11:20	**TimeX** Giovanni Samaey (KU Leuven) and Martin Schreiber (TU of Munich)
11:20 → 11:40	**Deep-Sea** Estela Suarez (JSC)
11:40 → 12:00	**Red-Sea** Nikos Xrysos (FORTH)
12:00 → 12:30	**Discussions between MAELSTROM and the speakers**
12:30 → 13:30	**Lunch break**
Part 3	**External Science Talks**
13:30 → 14:00	**Time-Consistent Downscaling of atmospheric fields with generative adversarial networks** Jussi Leinonen (MeteoSwiss)
14:00 → 14:30	**Pangeo: An open-source ecosystem for data-intensive science** Ryan Abernathey (Columbia University)
14:30 → 15:00	**Deep learning for earth sciences in the HPC context** Thorsten Kurth (NVIDIA)
15:00 → 15:30	**Discussions between MAELSTROM and the speakers**
15:30 → 16:00	**Coffee break**
16:00 → 17:30	**Possibility for further discussions between the working groups**

3.4 Participation

208 registrations were received. Eventually, up to 62 people participated at a single moment in time, i.e. speakers included. 34 participants contributed to the polls. The workshop started and ended as scheduled.

4 Resonance and results

4.1 Overall resonance and questions

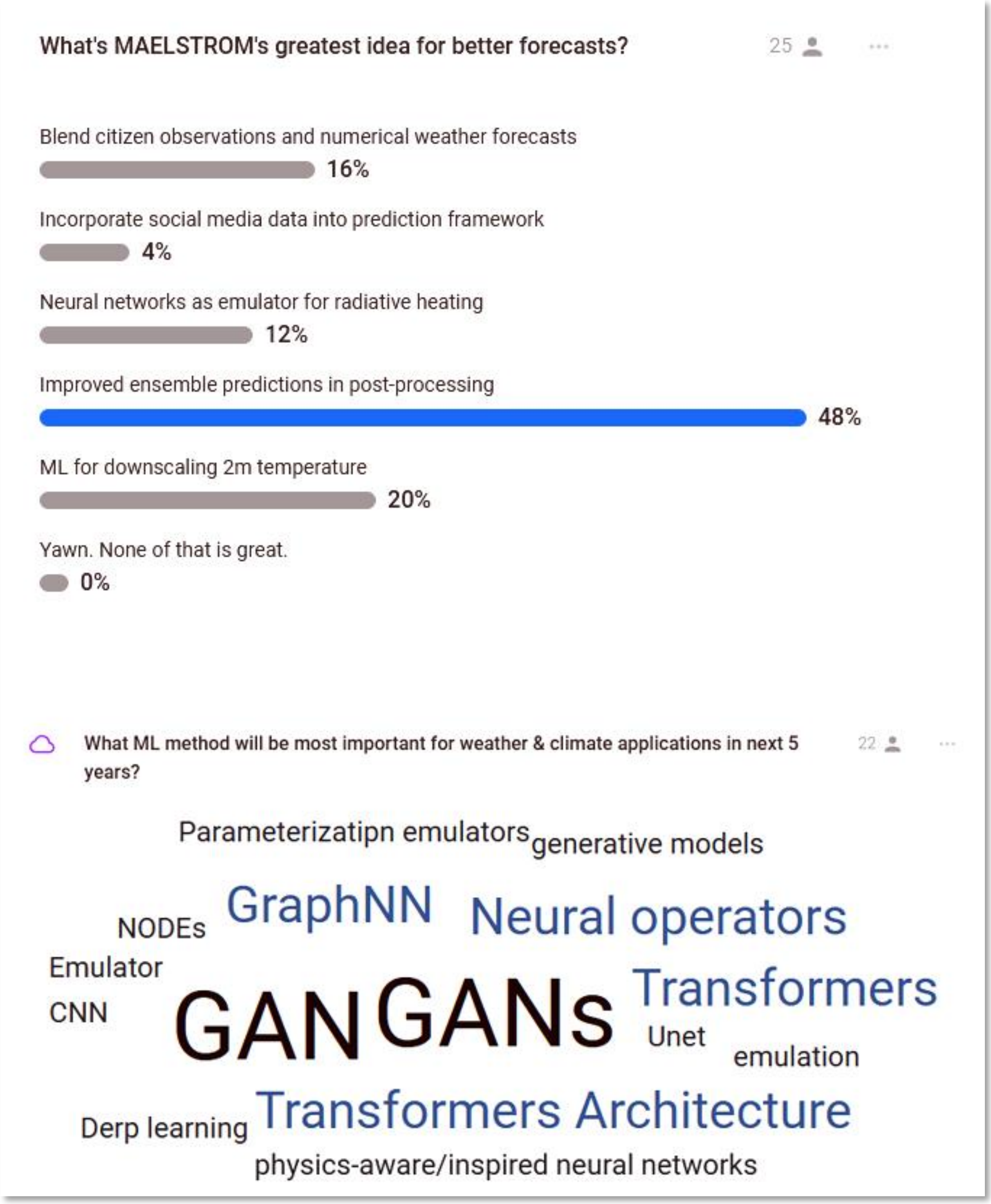
About 30 technical or scientific questions on the presented contents were collected through the Q&A tool, and either answered in verbally by the presenters, or in writing via the Q&A tool. The number, level and degree of detail of those inquiries demonstrated, that

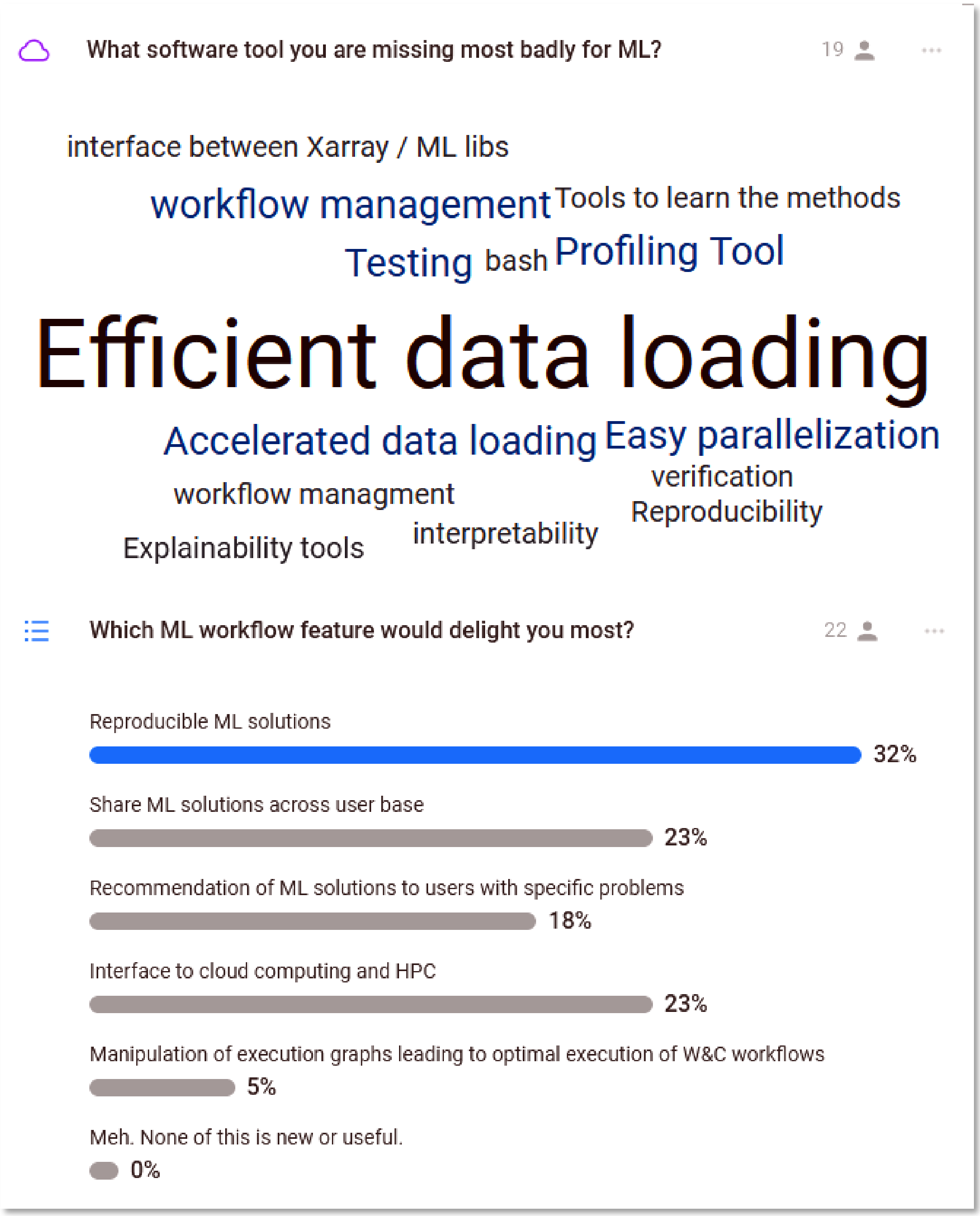
- the presented contents, such as the objectives, approaches, achievements, and products of project MAELSTROM were well understood
- project MAELSTROM pursues objectives, that will have good chances to be adopted by the technology and science community, because they respond to relevant needs.

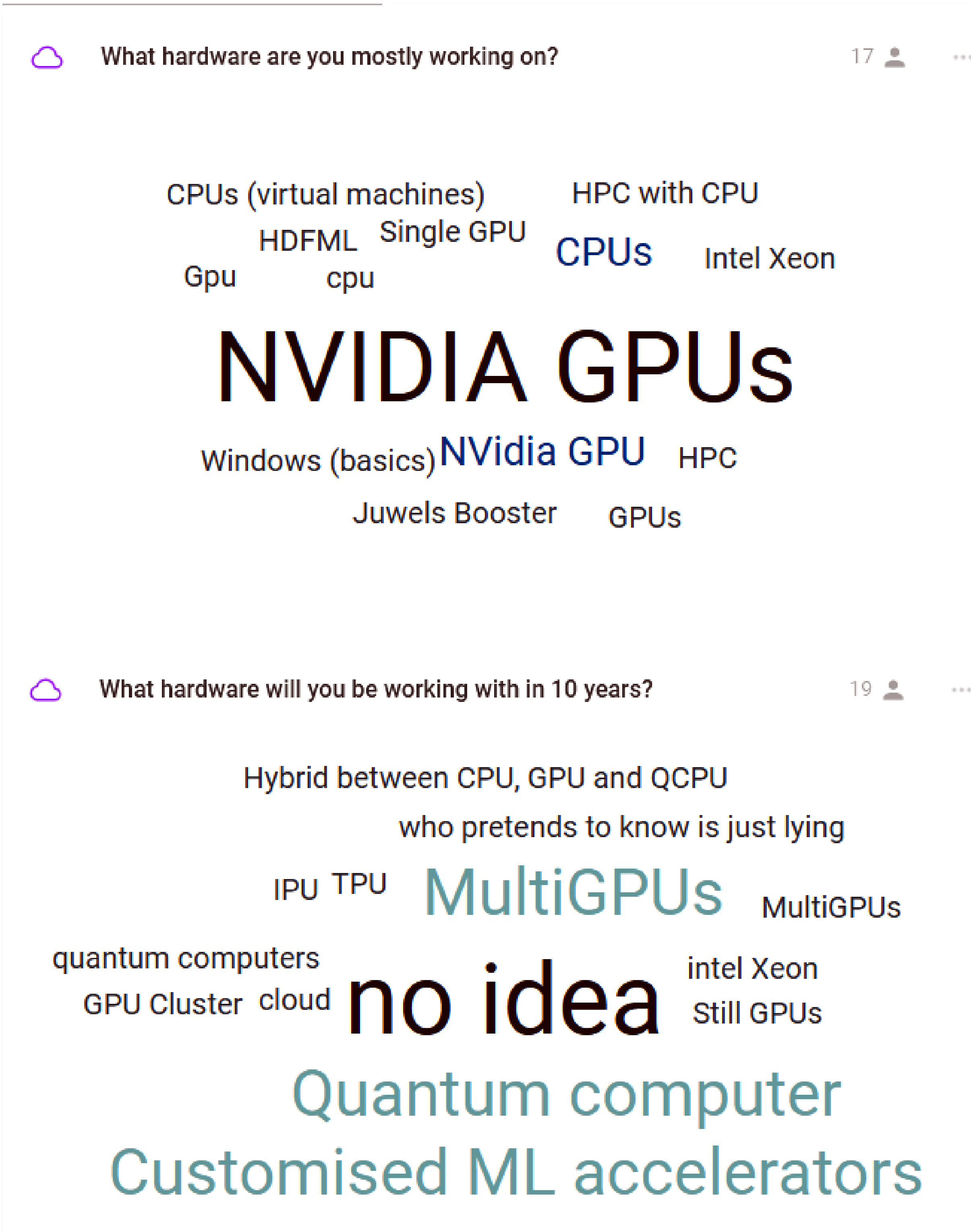
4.2 Poll results and conclusions

MAELSTROM had prepared 6 polls before the workshop, which were intended to jump-start a discussion, and enable the moderator (Peter Dueben) to structure this discussion. The questions were:

1. What's MAELSTROM's greatest idea for better forecasts? (Choice of 5 products of WP1)
2. What ML method will be most important for weather & climate prediction in the next 5 years? (open-ended)
3. What software tool are you missing most badly for ML? (open-ended)
4. Which ML workflow feature would delight you most? (Choice of 5 core features of WP2's workflow tools)
5. What hardware are you mostly working on?
6. What hardware will you be working on in 10 years?


What's MAELSTROM's greatest idea for better forecasts?
25
Blend citizen observations and numerical weather forecasts
16%
Incorporate social media data into prediction framework
4%
Neural networks as emulator for radiative heating
12%
Improved ensemble predictions in post-processing
48%
ML for downscaling 2m temperature
20%
Yawn. None of that is great.
0%
What ML method will be most important for weather & climate applications in next 5 years?
22
Parameterizatipn emulators
generative models
NODEs
GraphNN
Neural operators
Emulator
CNN
GAN
GANs
Transformers
Unet
emulation
Derp learning
Transformers Architecture
physics-aware/inspired neural networks


What software tool you are missing most badly for ML?
19
interface between Xarray / ML libs
workflow management
Tools to learn the methods
Testing
bash
Profiling Tool
Efficient data loading
Accelerated data loading
Easy parallelization
verification
workflow managment
Reproducibility
interpretability
Explainability tools
Which ML workflow feature would delight you most?
22
Reproducible ML solutions
32%
Share ML solutions across user base
23%
Recommendation of ML solutions to users with specific problems
18%
Interface to cloud computing and HPC
23%
Manipulation of execution graphs leading to optimal execution of W&C workflows
5%
Meh. None of this is new or useful.
0%



The results were discussed among and between MAELSTROM work package teams following the public part of the workshop.

5 Conclusion

With over 60 participants and a rich level of questions and discussions, MAELSTROM's first dissemination workshop can be considered successful. The format of the panel – MAELSTROM, partner projects, and relevant talks from the wider science community – has proven to be suitable and can serve as a model for the coming second dissemination workshop.

Learning from this workshop will help us develop the interactive parts – polling, discussion – further for the coming hackathons and dissemination workshop II.

A recording of the workshop, MAELSTROM presentation decks and poll results will be made available on the MAELSTROM project website. Recording for some of the talks as well as the slides are already available on Machine Learning Workshop webpage[2].

[2] https://events.ecmwf.int/event/294/timetable/

Document History

Version	Author(s)	Date	Changes
0.1	Jan Mirus	30/03/2022	Initial version
1.0	Peter Dueben (ECMWF)	30/03/2022	Final version

Internal Review History

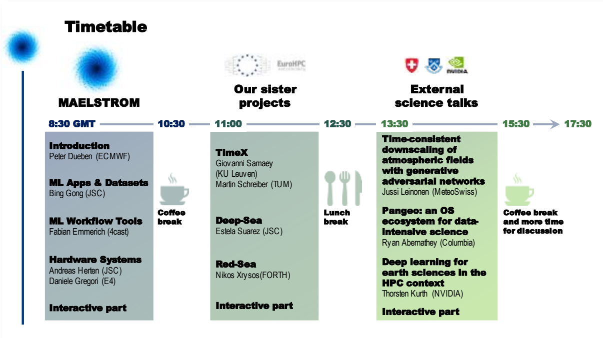
Internal Reviewers	Date	Comments
Peter Dueben (ECMWF)	30/03/2022	Approved with comments

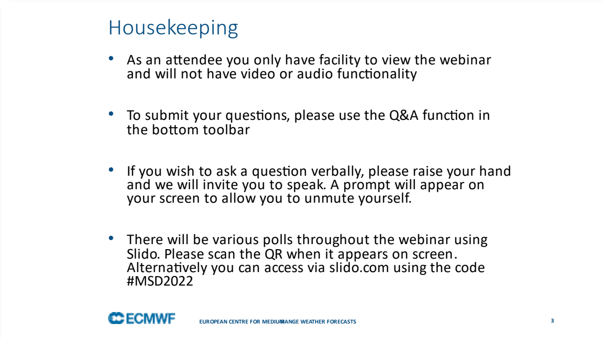
Estimated Effort Contribution per Partner

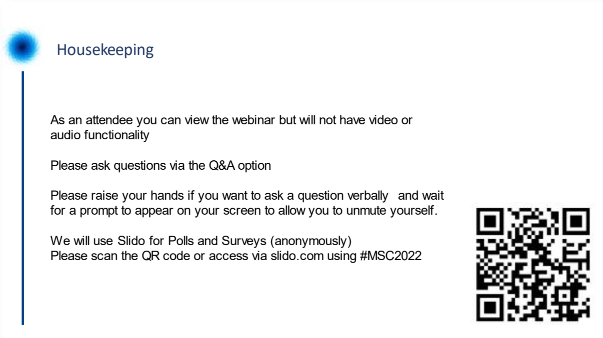
Partner	Effort
Jan Mirus	0.25
ECMWF	0.25
Total	**0.5**

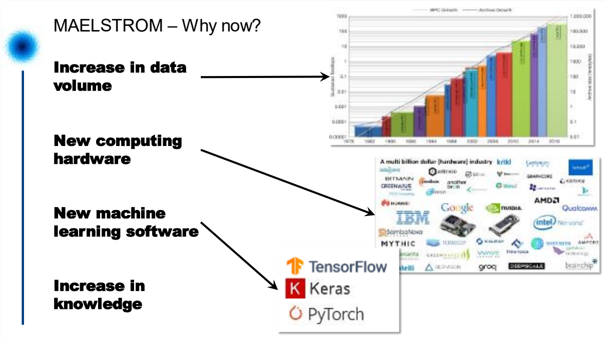
6 Annex: MAELSTROM presentation decks

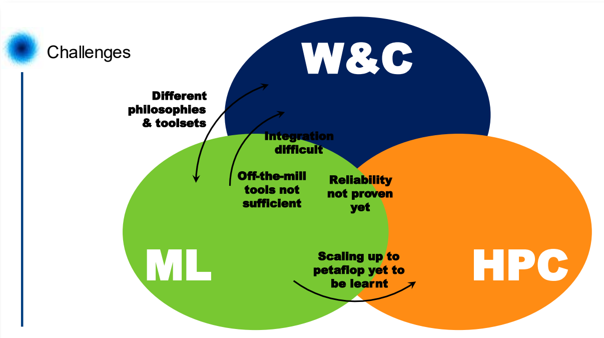
6.1 Introduction: Peter Dueben, ECMWF

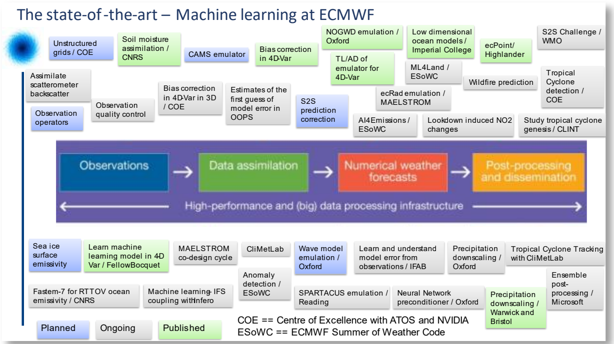




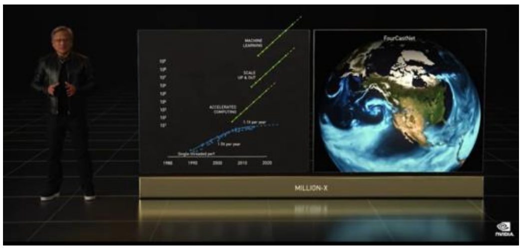








The perspective – Full ML models for weather and climate?



NIVIDA's Earth-2 is coming with FourCastNet – see Thorsten's talk

Climate?

A myriad of options…

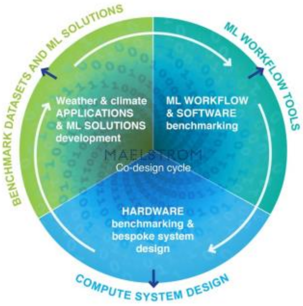
A myriad of options for machine learning approaches

Dense Neural Networks, LSTMs, ConvGru, Attention Layers, Transformer networks, # of hidden layers, different normalisation of inputs, batch normalisation, tanh, relu, gelu, softplus, elu, selu, leaky relu, softmax, sigmoid function, generative adversarial networks, recurrent neural networks, encoding/decoding networks, random forests, boosting methods, clustering techniques, singular vector decomposition, causal discovery, ablation studies, root mean square error, variational auto encoder, gradient descent, stochastic gradient decent, adagrad, adadelta, RMSprop, Adam, # of epochs, # of batches, learning rate, overfitting, dropout, Bayesian networks, Gaussian processes, half precision, sparse networks…. Argh…

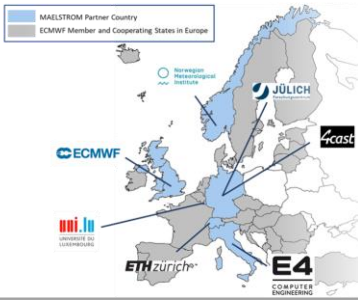
+ a myriad of options for machine learning hardware

= confused scientists

The MAELSTROM approach



We are MAELSTROM



MAELSTROM – How to get involved?

- Visit our ***project webpage***: www.maelstrom-eurohpc.eu
- Use our ***benchmark datasets***
- Use our ***machine learning blueprints*** for other applications
- Use of ***machine learning workflow tools***
- Use of ***compute system designs***
- ***Talks, deliverables and publications***
- Two ***dissemination workshops*** are planned for April 2022 and Spring 2024 at ECMWF
- Two ***hackathons*** foreseen (2022 in Juelich and 2023 at ECMWF)

Our partners have a special interest in collaborations regarding:

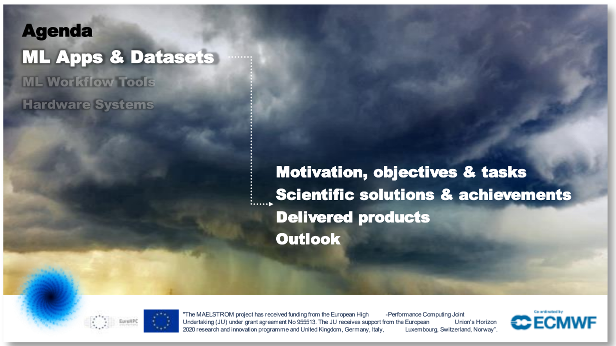
Machine learning, data movement and IO, approximate computing and mixed precision, hardware accelerators and co-design, hardware and software benchmarks and many other topics.

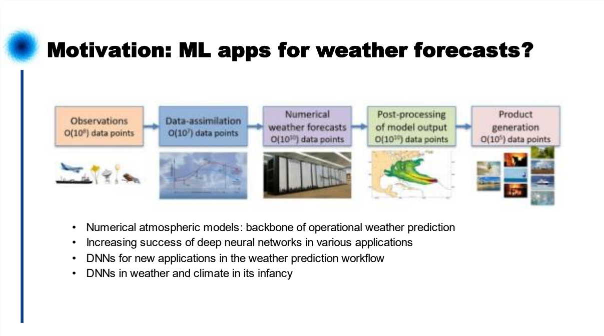
Questions?

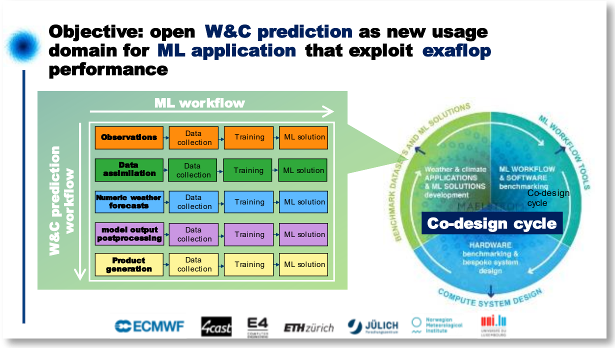
Please also join us for the ECMWF Machine Learning Workshop Tu -Fr
https://events.ecmwf.int/event/294/

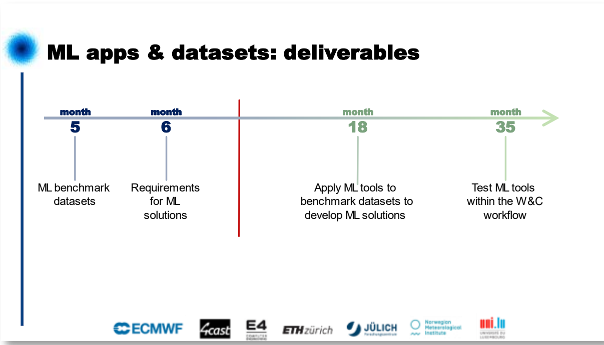
This presentation reflects the views only of the author, and the European High -Performance Computing Joint Undertaking or Commis sion cannot be held responsible for any use which may be made of the information contained therein.

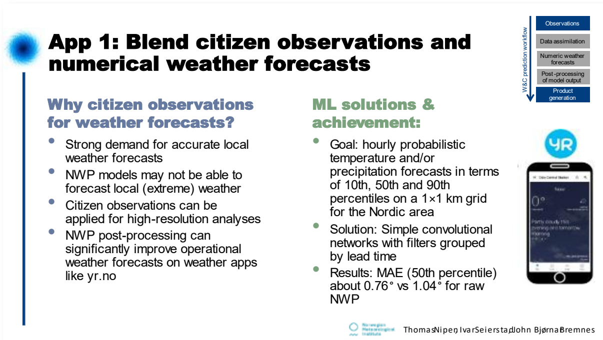
6.2 WP1: Gong Bing, JSC

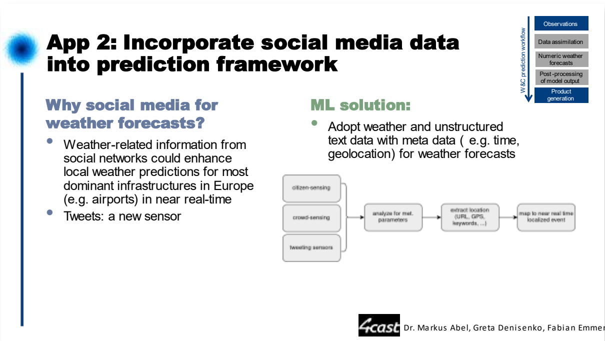


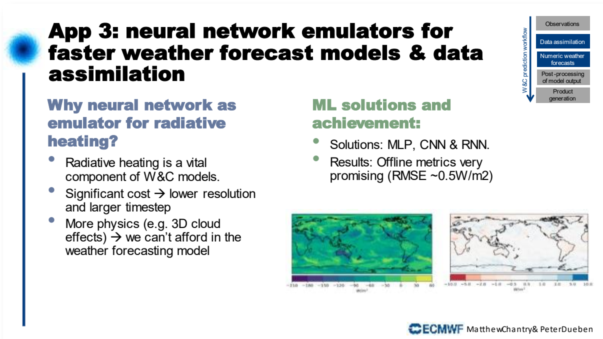


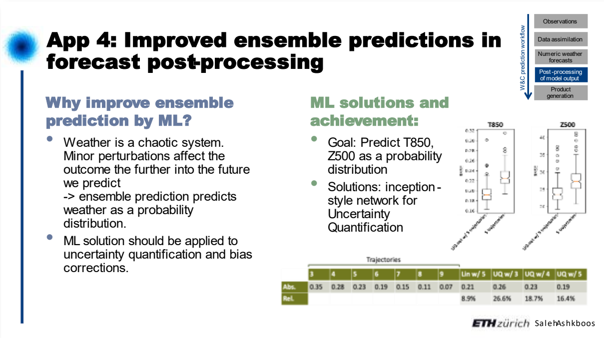


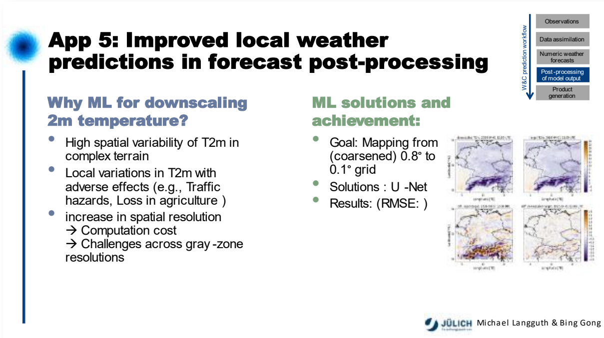


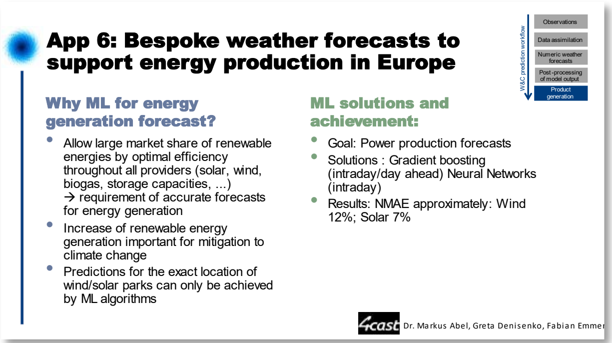


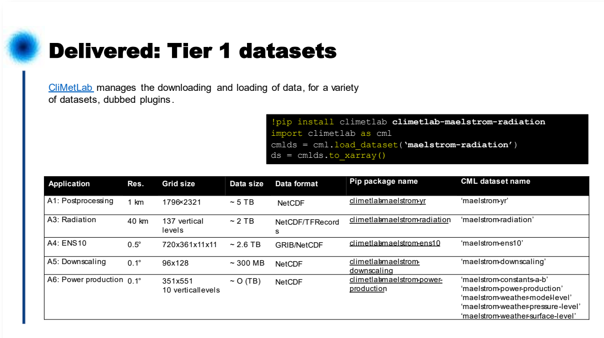


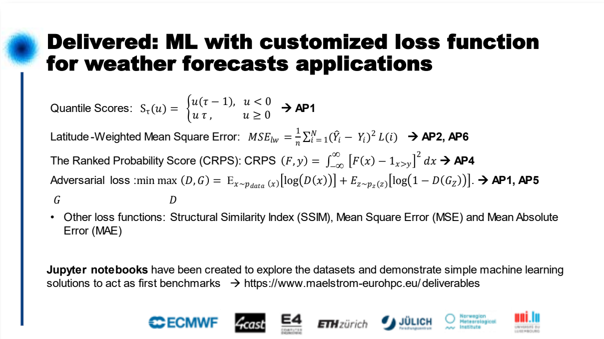


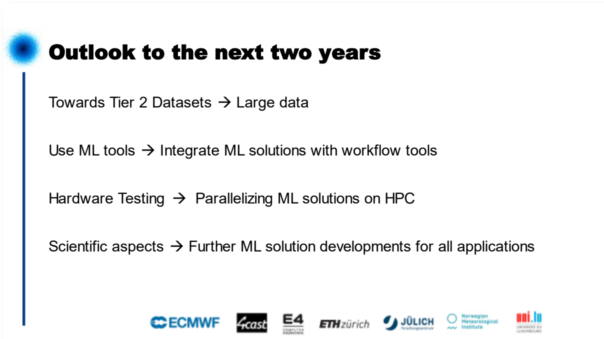

App 4: Improved ensemble predictions in forecast post-processing
Observations
Data assimilation
Numeric weather forecasts
Post-processing of model output
Product generation
W&C prediction workflow
Why improve ensemble prediction by ML?
Weather is a chaotic system. Minor perturbations affect the outcome the further into the future we predict -> ensemble prediction predicts weather as a probability distribution.
ML solution should be applied to uncertainty quantification and bias corrections.
ML solutions and achievement:
Goal: Predict T850, Z500 as a probability distribution
Solutions: inception-style network for Uncertainty Quantification
T850
Z500
Trajectories
3 4 5 6 7 8 9 Lin w/ 5 UQ w/ 3 UQ w/ 4 UQ w/ 5
Abs. 0.35 0.28 0.23 0.19 0.15 0.11 0.07 0.21 0.26 0.23 0.19
Rel. 8.9% 26.6% 18.7% 16.4%
ETH zürich Saleh Ashkboos


App 5: Improved local weather predictions in forecast post-processing
Observations
Data assimilation
Numeric weather forecasts
Post-processing of model output
Product generation
W&C prediction workflow
Why ML for downscaling 2m temperature?
High spatial variability of T2m in complex terrain
Local variations in T2m with adverse effects (e.g., Traffic hazards, Loss in agriculture)
increase in spatial resolution
→ Computation cost
→ Challenges across gray-zone resolutions
ML solutions and achievement:
Goal: Mapping from (coarsened) 0.8° to 0.1° grid
Solutions : U -Net
Results: (RMSE:)
JÜLICH Michael Langguth & Bing Gong


App 6: Bespoke weather forecasts to support energy production in Europe
Observations
Data assimilation
Numeric weather forecasts
Post-processing of model output
Product generation
W&C prediction workflow
Why ML for energy generation forecast?
Allow large market share of renewable energies by optimal efficiency throughout all providers (solar, wind, biogas, storage capacities, …)
→ requirement of accurate forecasts for energy generation
Increase of renewable energy generation important for mitigation to climate change
Predictions for the exact location of wind/solar parks can only be achieved by ML algorithms
ML solutions and achievement:
Goal: Power production forecasts
Solutions : Gradient boosting (intraday/day ahead) Neural Networks (intraday)
Results: NMAE approximately: Wind 12%; Solar 7%
4cast Dr. Markus Abel, Greta Denisenko, Fabian Emmer

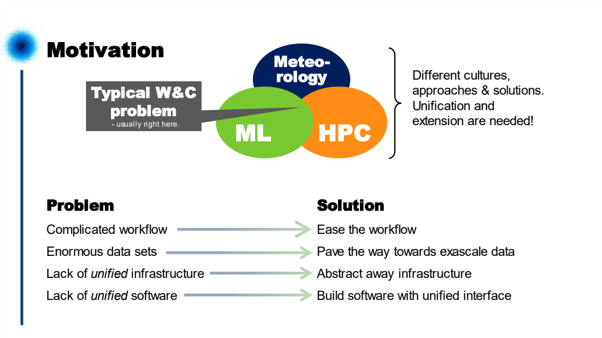

Delivered: Tier 1 datasets
CliMetLab manages the downloading and loading of data, for a variety of datasets, dubbed plugins.
!pip install climetlab climetlab-maelstrom-radiation
import climetlab as cml
cmlds = cml.load_dataset('maelstrom-radiation')
ds = cmlds.to_xarray()
Application | Res. | Grid size | Data size | Data format | Pip package name | CML dataset name
A1: Postprocessing | 1 km | 1796×2321 | ~ 5 TB | NetCDF | climetlab-maelstrom-yr | 'maelstrom-yr'
A3: Radiation | 40 km | 137 vertical levels | ~ 2 TB | NetCDF/TFRecords | climetlab-maelstrom-radiation | 'maelstrom-radiation'
A4: ENS10 | 0.5° | 720x361x11x11 | ~ 2.6 TB | GRIB/NetCDF | climetlab-maelstrom-ens10 | 'maelstrom-ens10'
A5: Downscaling | 0.1° | 96x128 | ~ 300 MB | NetCDF | climetlab-maelstrom-downscaling | 'maelstrom-downscaling'
A6: Power production | 0.1° | 351x551 10 vertical levels | ~ O (TB) | NetCDF | climetlab-maelstrom-power-production | 'maelstrom-constants-a-b' 'maelstrom-power-production' 'maelstrom-weather-model-level' 'maelstrom-weather-pressure-level' 'maelstrom-weather-surface-level'

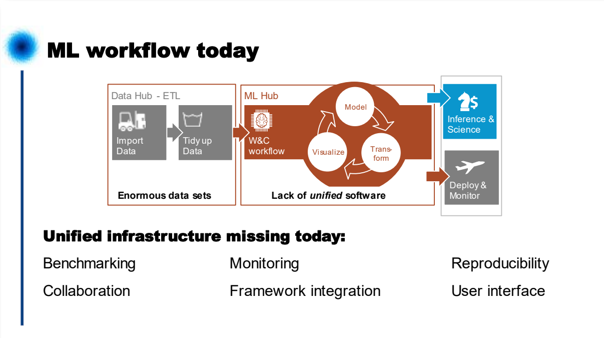

Delivered: ML with customized loss function for weather forecasts applications
Quantile Scores: $S_\tau(u) = \begin{cases} u(\tau-1), & u<0 \\ u\,\tau, & u \geq 0 \end{cases}$ → AP1
Latitude-Weighted Mean Square Error: $MSE_{lw} = \frac{1}{n}\sum_{i=1}^{N}(\hat{Y}_i - Y_i)^2 L(i)$ → AP2, AP6
The Ranked Probability Score (CRPS): CRPS $(F,y) = \int_{-\infty}^{\infty}\left[F(x) - 1_{x>y}\right]^2 dx$ → AP4
Adversarial loss : $\min_G \max_D (D,G) = \mathrm{E}_{x\sim p_{data}(x)}[\log(D(x))] + E_{z\sim p_z(z)}[\log(1-D(G_Z))]$. → AP1, AP5
Other loss functions: Structural Similarity Index (SSIM), Mean Square Error (MSE) and Mean Absolute Error (MAE)
Jupyter notebooks have been created to explore the datasets and demonstrate simple machine learning solutions to act as first benchmarks → https://www.maelstrom-eurohpc.eu/deliverables
ECMWF 4cast E4 ETH zürich JÜLICH Norwegian Meteorological Institute uni.lu

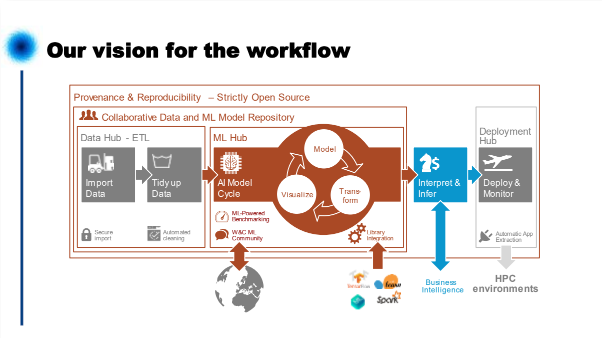

Outlook to the next two years
Towards Tier 2 Datasets → Large data
Use ML tools → Integrate ML solutions with workflow tools
Hardware Testing → Parallelizing ML solutions on HPC
Scientific aspects → Further ML solution developments for all applications
ECMWF 4cast E4 ETH zürich JÜLICH Norwegian Meteorological Institute uni.lu

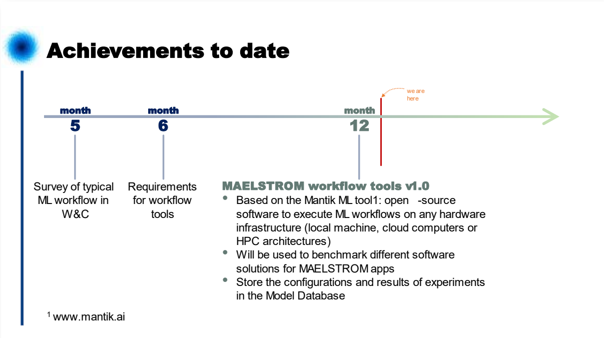

www.maelstrom-eurohpc.eu
MAELSTROM
Thank you
ECMWF 4cast E4 ETH zürich JÜLICH Norwegian Meteorological Institute uni.lu
"The MAELSTROM project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 955513. The JU receives support from the European Union's Horizon 2020 research and innovation programme and United Kingdom, Germany, Italy, Luxembourg, Switzerland, Norway".
Co-ordinated by ECMWF

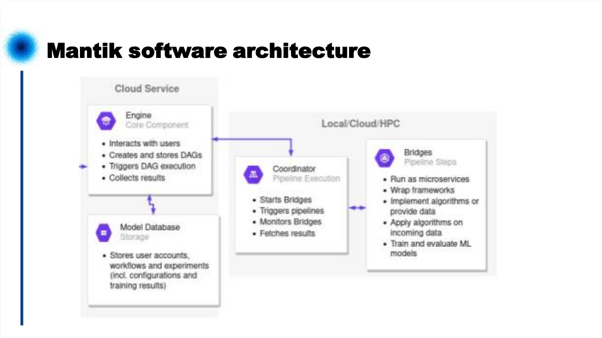
6.3 WP2: Fabian Emmerich, 4cast

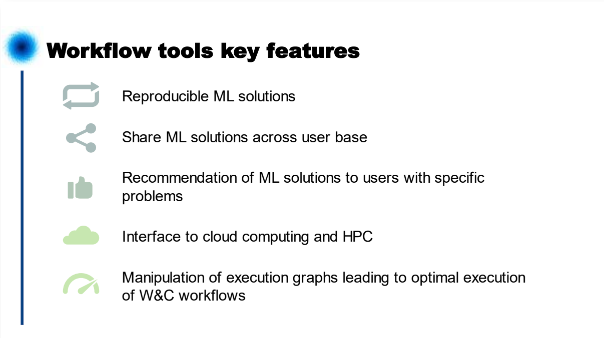


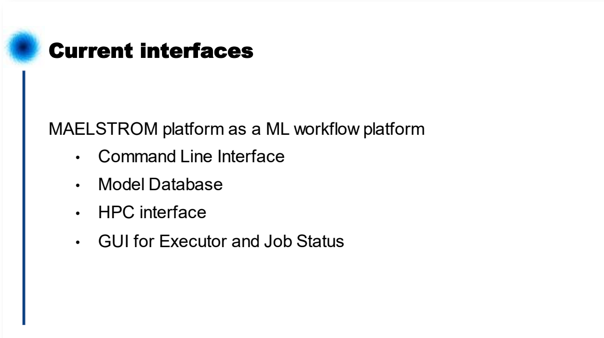


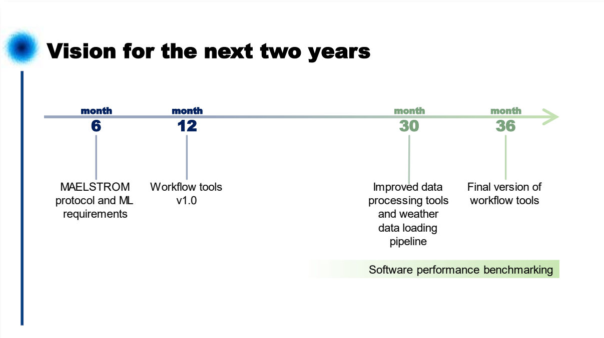




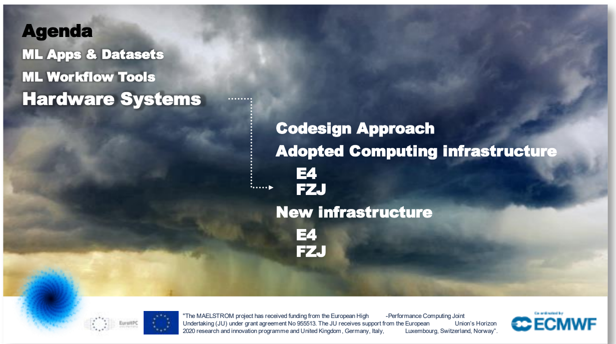


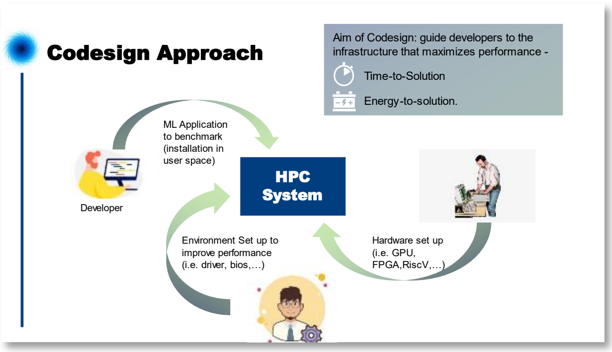


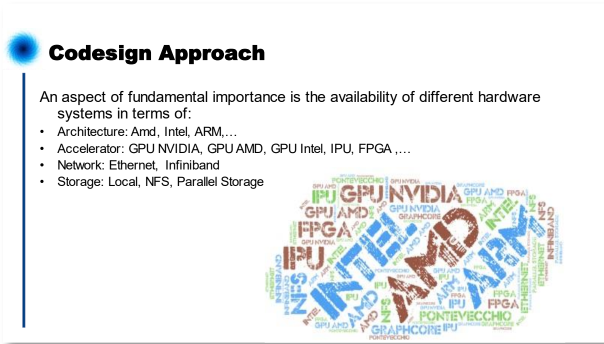


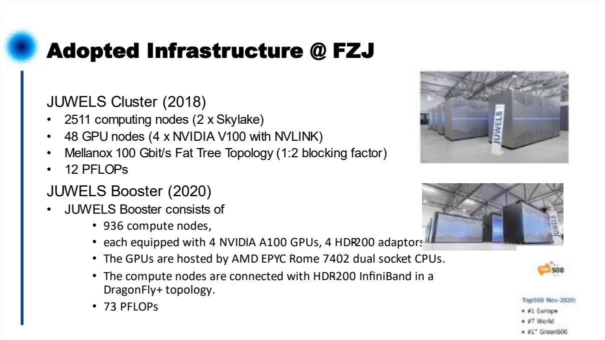


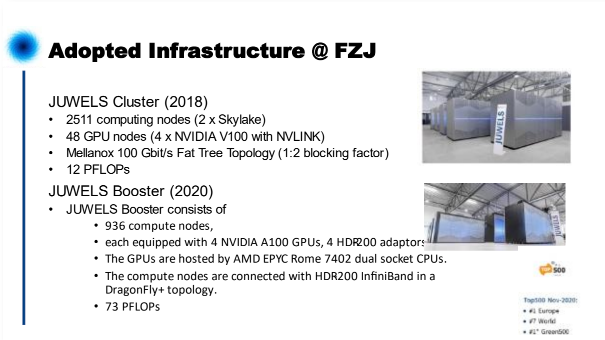
6.4 WP3: Daniele Gregori, E4

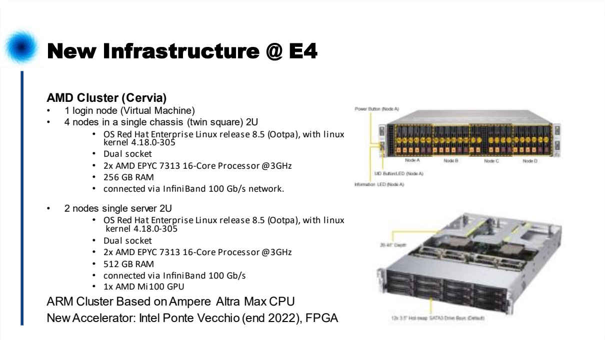


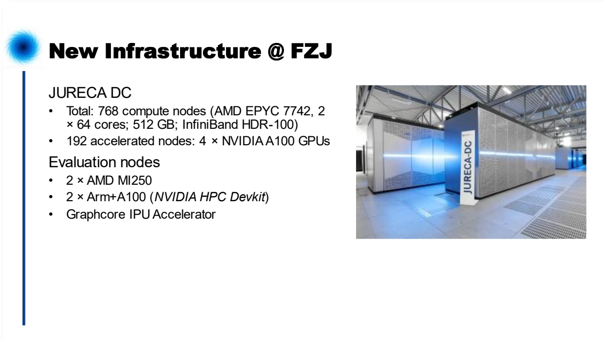





Adopted Infrastructure @ FZJ
JUWELS Cluster (2018)
2511 computing nodes (2 x Skylake)
48 GPU nodes (4 x NVIDIA V100 with NVLINK)
Mellanox 100 Gbit/s Fat Tree Topology (1:2 blocking factor)
12 PFLOPs
JUWELS Booster (2020)
JUWELS Booster consists of
936 compute nodes,
each equipped with 4 NVIDIA A100 GPUs, 4 HDR200 adaptors
The GPUs are hosted by AMD EPYC Rome 7402 dual socket CPUs.
The compute nodes are connected with HDR200 InfiniBand in a DragonFly+ topology.
73 PFLOPs


Adopted Infrastructure @ FZJ
JUWELS Cluster (2018)
2511 computing nodes (2 x Skylake)
48 GPU nodes (4 x NVIDIA V100 with NVLINK)
Mellanox 100 Gbit/s Fat Tree Topology (1:2 blocking factor)
12 PFLOPs
JUWELS Booster (2020)
JUWELS Booster consists of
936 compute nodes,
each equipped with 4 NVIDIA A100 GPUs, 4 HDR200 adaptors
The GPUs are hosted by AMD EPYC Rome 7402 dual socket CPUs.
The compute nodes are connected with HDR200 InfiniBand in a DragonFly+ topology.
73 PFLOPs


New Infrastructure @ E4
AMD Cluster (Cervia)
1 login node (Virtual Machine)
4 nodes in a single chassis (twin square) 2U
OS Red Hat Enterprise Linux release 8.5 (Ootpa), with linux kernel 4.18.0-305
Dual socket
2x AMD EPYC 7313 16-Core Processor @3GHz
256 GB RAM
connected via InfiniBand 100 Gb/s network.
2 nodes single server 2U
OS Red Hat Enterprise Linux release 8.5 (Ootpa), with linux kernel 4.18.0-305
Dual socket
2x AMD EPYC 7313 16-Core Processor @3GHz
512 GB RAM
connected via InfiniBand 100 Gb/s
1x AMD Mi100 GPU
ARM Cluster Based on Ampere Altra Max CPU
New Accelerator: Intel Ponte Vecchio (end 2022), FPGA


New Infrastructure @ FZJ
JURECA DC
Total: 768 compute nodes (AMD EPYC 7742, 2 × 64 cores; 512 GB; InfiniBand HDR-100)
192 accelerated nodes: 4 × NVIDIA A100 GPUs
Evaluation nodes
2 × AMD MI250
2 × Arm+A100 (NVIDIA HPC Devkit)
Graphcore IPU Accelerator
JURECA-DC


Questions?
ECMWF
4cast
E4
ETH zürich
JÜLICH
Norwegian Meteorological Institute
This presentation reflects the views only of the author, and the European High -Performance Computing Joint Undertaking or Commis sion cannot be held responsible for any use which may be made of the information contained therein.

6.5 WP3: Andreas Herten, JSC