CATALYST
Leveraging HPC to Drive Innovation in AI

Dennis Hoppe (HLRS)

HLRS

How HLRS addresses …

CHALLENGES OF AI

European Strategy for AI — Three Pillars [1]

- Boosting the EU's technological and industrial capacity and **AI uptake across the economy**
  - Supporting AI research excellence centers across Europe
  - Bringing AI to all small businesses and potential users
- Preparing for **socioeconomic changes**
  - Focus on jobs that are likely to be transformed or to disappear; leverage chances of new job creations
- Ensuring an appropriate **ethical and legal framework**
  - Citizens and businesses alike need to be able to trust the technology they interact with

[1] European Commission: Communication Artificial Intelligence for Europe, 2018.

Why does AI need HPC? Why does HPC need AI?

- **AI solutions require immense compute-resources**
  - CPU, network, storage, accelerators, …
- **Simulations** such as climate models **hit the wall**
  - computing physical processes right down to the last detail is very compute-intensive
- **Information overload** will continue to increase
  - 5G, IoT, autonomous driving and flying, …
- HLRS addresses these challenges through different channels
  - **Economy**, **Society**, **Research**

Challenges of AI: Economy / Society / Research

- **Economy** (with focus on SMEs)
  - missing AI expertise
  - no in-house AI hardware
  - security concerns (GDPR)
- **Society**
  - AI is seen as a blackbox model
  - low acceptance rates of AI solutions
  - security concerns (privacy intrusion)
- **Research**
  - Support of hybrid HPC/AI workflows on HPC systems
  - Multitude of complementary requirements (e.g. software)
  - AI experts are no HPC experts
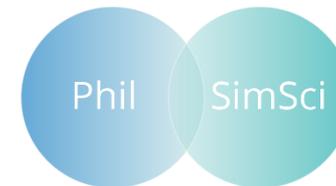
Combining HPC and HPDA for Academia and Industry

CATALYST

CATALYST. Overview [2016–2021]

- Our customers tend to run more and more **data-intensive applications** resulting in **vast amounts of output data**
  - A single turbulence & acoustics simulation of an axial fan with just four rotations results in 80 TB of data
  - Domain experts are no longer able to analyze data manually
- Close cooperation between **HLRS** and **Cray** ($\rightarrow$ HPE)

- Evaluate requirements that arise when combining AI and HPC
  - Hardware + software environment
    - Cray Urika-GX (DA/ML), CS-Storm (DL), HPE Apollo (HPC)
    - Open-source software stack
  - Perform **case studies** with both academia and industry

Case Study (1/5): Big Data Processing ("3D City over Night") – nFrames

The illustration shows a textured 3D mesh of San Francisco. The data was provided by courtesy of Geomni. Copyright nFrames.

Case Study (2/5): AI meets Art ("Who is drawing?") – Lunar Ring

Image courtesy: Johannes Stelzer, Lunar Ring.

Case Study (3/5): AI Ethics (Ethical Framework) – AIEI

- Framework is proposed by the **AI Ethics Impact Group** (AIEI)
  - With contributions by Andreas Kaminksi, Philosophy at HLRS
- **Traffic-light based rating system** to give an AI system an ethics label
  - Each rating is based on a set of criteria making the rating system itself transparent.
- Various **dimensions**, e.g.,
  - Transparency
  - Privacy
  - Reliability
  - …

Transparency
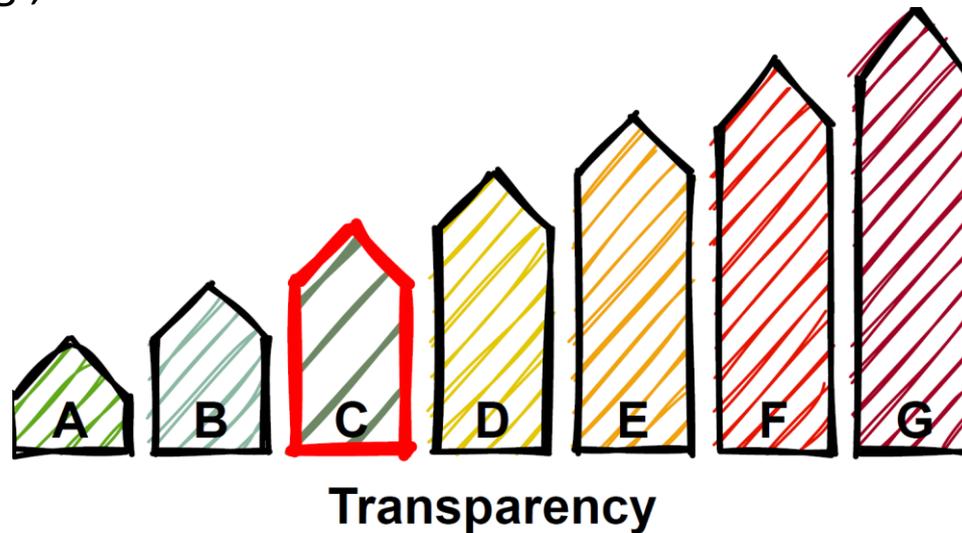
Case Study (4/5): Data-Analytics-as-a-Service – LandesCloud

- We need to **support SMEs**!
  - while 25% of industrial companies leverage AI, only 15% of small-and-medium enterprises use AI technology [2]
- Specifically SMEs face many hurdles
  - many times no AI expertise in-house
  - often no powerful hardware in-house
  - concerns about **data security** and transfer (cf. GDPR)
- LandesCloud provides DaaS on top of HLRS infrastructure
  - KNIME, TensorFlow, PyTorch, Spark
  - **CrowdiQ platform** for virtual hackathons, AI team-working across company boundaries

landescloud

Support during the starting phase: Research & Development

↓

Interactive data analytics and training of AI models

↓

Going into production; automatic and continuous process optimization
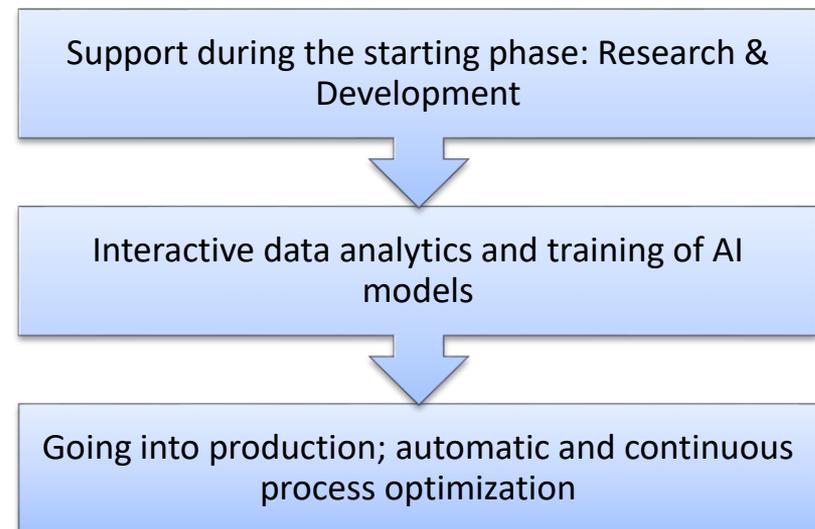
[2] Study by BMWi about potentials of AI for German companies, 2018.

Case Study (5/5): Hybrid HPC/AI Workflows

a) Synthetic data generation
b) Define parameters for the simulation
c) Initial solution is given by an AI model
d) Optimise the parameter space of simulations
e) Iterations are alternating between AI and simulation
f) Surrogate models (functions; equations)
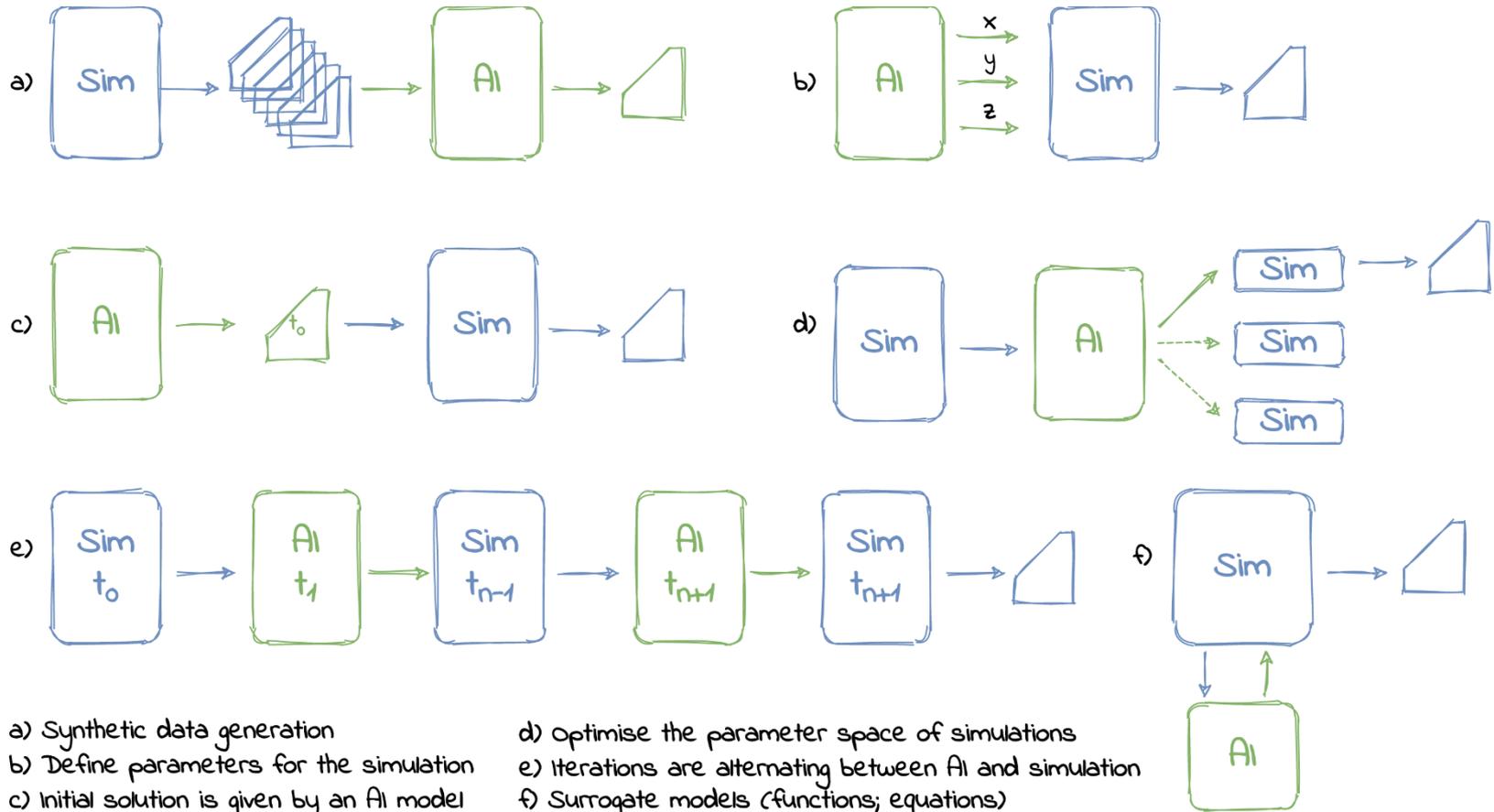
CATALYST. Combining HPC and HPDA for Academia and Industry

CONCLUSIONS

Take Away Messages

- AI is a **many-faceted domain**
- HPC community needs to **see the bigger picture**
- **Complementary views**
  - **HPC needs AI**
    - AI for **parameter sweeps**: reduce #jobs; save costs
    - AI for **simulations**: e.g., surrogate models
    - …
  - **AI needs HPC**
    - Computational and data needs exceed commodity hardware
    - (near) real-time responses
- As a HPC center we have to address all of the above!

Further Information

- CATALYST (https://www.hlrs.de/bigdata)
- Photogrammetry (www.nframes.com)
- Art Meets AI
  - Website: http://www.lunar-ring.ai/
  - YouTube Channel: https://bit.ly/3kzaLXl
- Ethical AI Framework (https://www.ai-ethics-impact.org/en)
- Hybrid HPC/AI Workflows (https://bit.ly/3lCUFNR)
- LandesCloud (https://www.landes.cloud/)
- Simulierte Welten (https://simulierte-welten.de/)
- Philosophy at HLRS (https://philo.hlrs.de/)

Thank you !

This work was supported by the research project CATALYST funded by the Ministry of Science, Research and the Arts of Baden-Württemberg, Germany (2016–2021).

Dennis Hoppe
High-Performance Computing Center Stuttgart (HLRS)
Nobelstraße 19
70569 Stuttgart, Germany
Email: hoppe@hlrs.de
Web: https://hlrs.de/bigdata