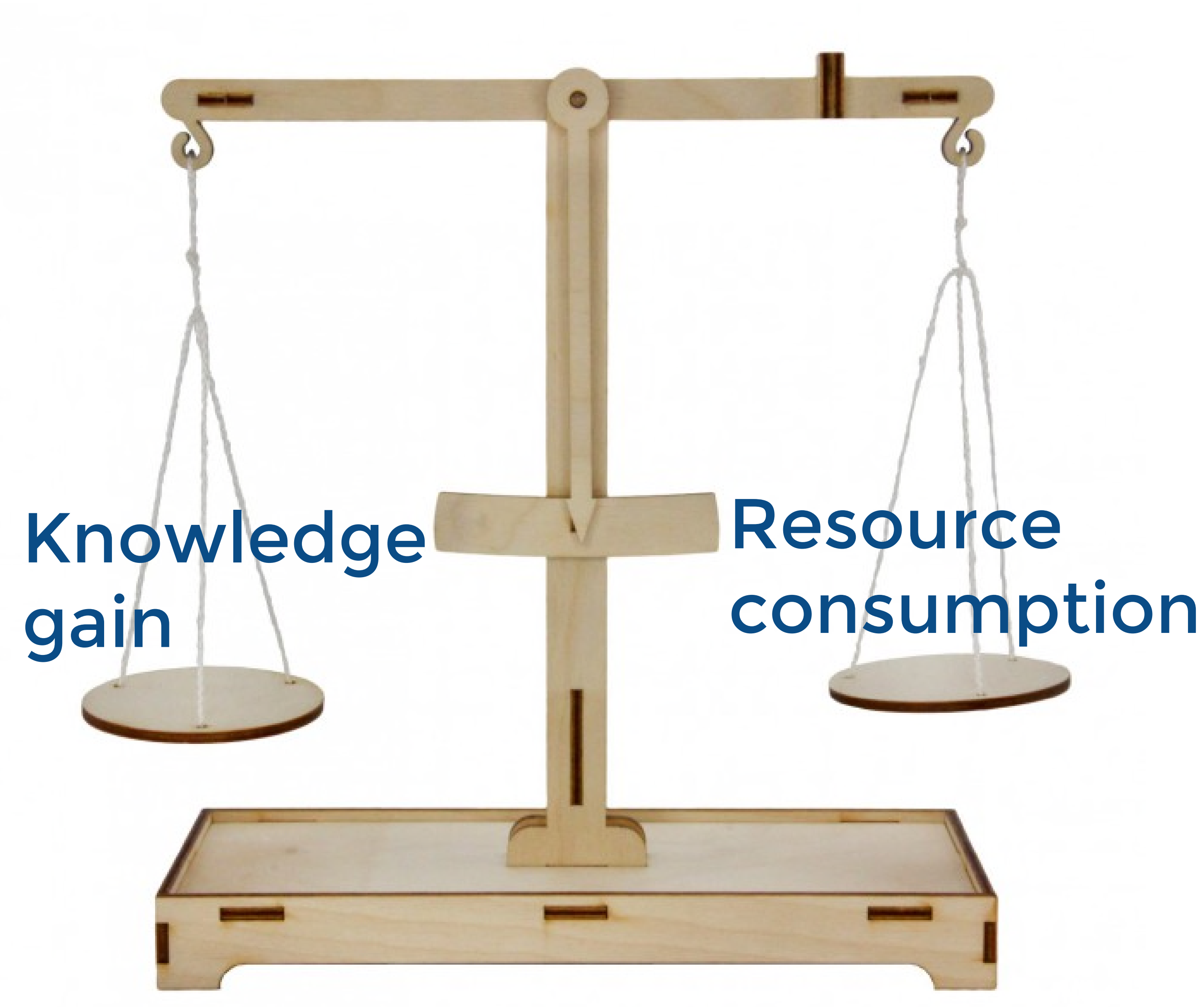
Sustainability

Martin Erdmann, RWTH Aachen

28-Aug-2024

https://indico.desy.de/event/37480



ERUM DATA HUB

White paper Resource-aware Research on Universe and Matter: Call-to-Action in Digital Transformation

https://arxiv.org/abs/2311.01169

1 The challenge
2 Smart data transformation
3 Software engineering and data analysis
4 Algorithms and artificial intelligence
5 Computing and infrastructures:
- Renewable energy
- Hardware lifetime
- Information flow and middleware

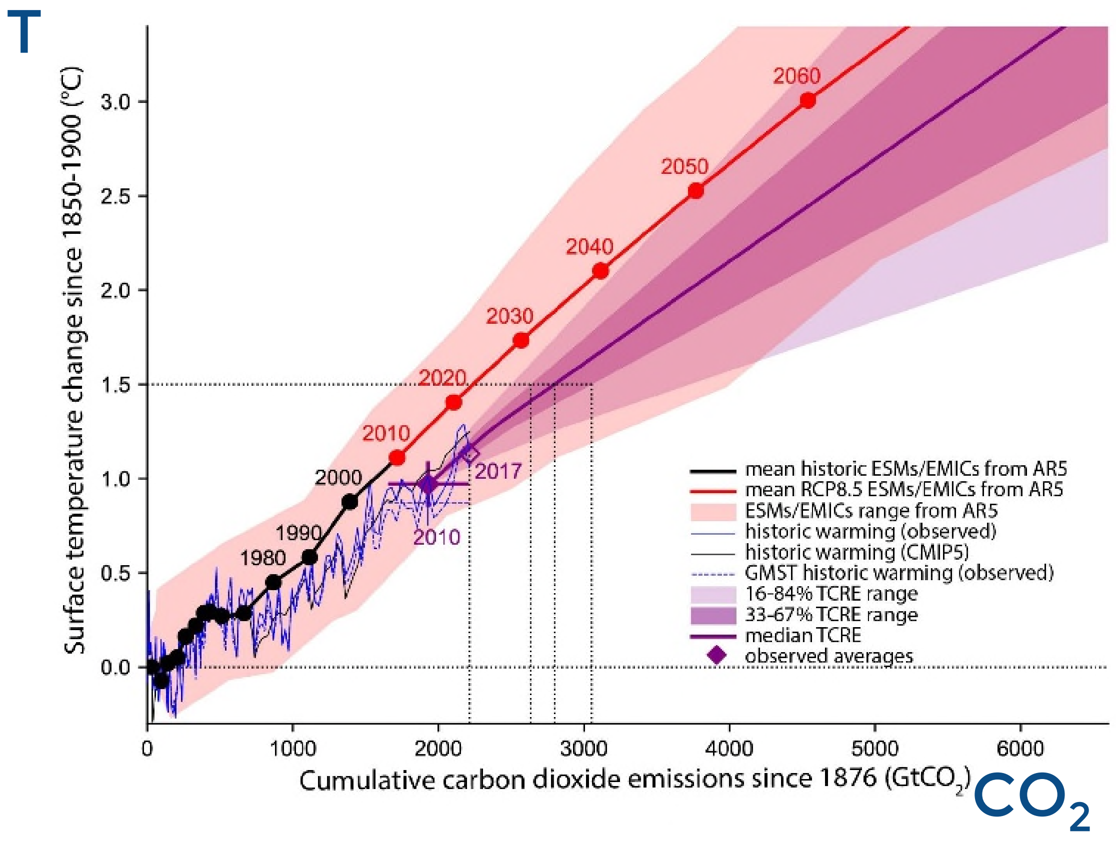
6 Developing a **culture for sustainable science**
7 Funding and institutional support

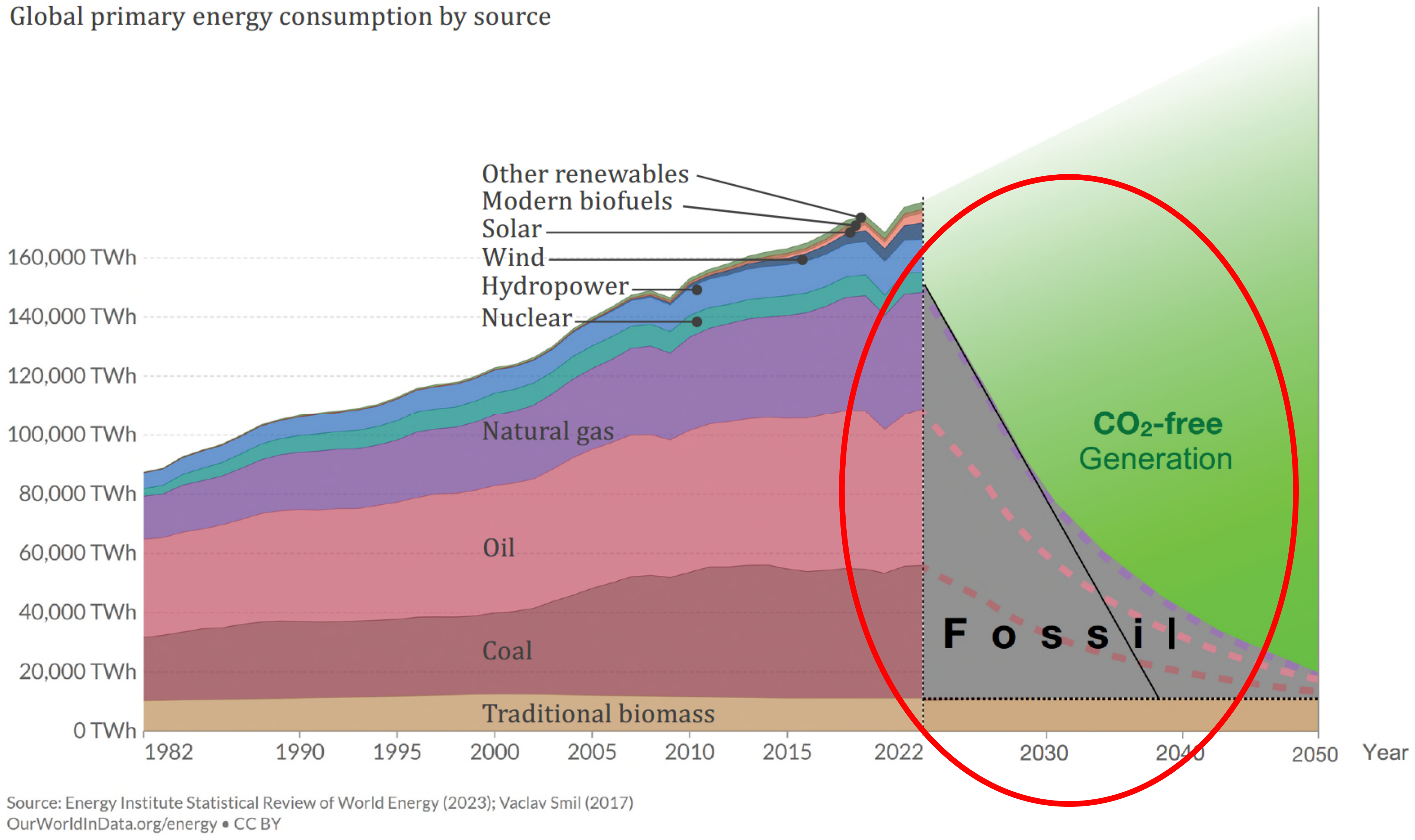
Ben Bruers[1], Marilyn Cruces[2], Markus Demleitner[3], Guenter Duckeck[4], Michael Düren[5], Niclas Eich[6], Torsten Enßlin[7], Johannes Erdmann[6], Martin Erdmann[6*], Peter Fackeldey[6], Christian Felder[8], Benjamin Fischer[6], Stefan Fröse[9], Stefan Funk[10], Martin Gasthuber[1], Andrew Grimshaw[11], Daniela Hadasch[9,12], Moritz Hannemann[8], Alexander Kappes[2], Raphael Kleinemühl[13], Oleksiy M. Kozlov[14], Thomas Kuhr[4], Michael Lupberger[15], Simon Neuhaus[13], Pardis Niknejadi[1], Judith Reindl[16], Daniel Schindler[17], Astrid Schneidewind[8], Frank Schreiber[18], Markus Schumacher[19], Kilian Schwarz[1], Achim Streit[20], R. Florian von Cube[20], Rodney Walker[4], Cyrus Walther[9], Sebastian Wozniewski[17], Kai Zhou[21]

Item	Call-to-action
	Immediately or on **short time scale** with little effort these measures can be implemented:
S1	Raise awareness of the climate challenge at all levels.
S2	Disseminate knowledge of measures to address the challenge.
S3	Monitor and report energy consumption at job level.
S4	Consider carbon footprint for all investments and project plans.
S5	Enhance awareness of the trade-off between research benefit and climate impact.
	On a **medium time scale** of a few years the following measures can be realized:
M1	Make data FAIR to promote reuse.
M2	Reduce and compress data having the anticipated scientific value of the retained information and the resource requirements in mind.
M3	Optimize the choice of storing intermediate results against re-calculating them.
M4	Optimize job orchestration and scheduling in workflows.
M5	Use workflow management to make processing FAIR.
M6	Make software FAIR and reliable by following good software development practices and ensuring sustainable support.
M7	Design software for optimized energy consumption and provide tools to measure it.
M8	Continue research on potential of AI or other new technologies for efficient use of resources, but balance gain of research action against resource consumption of these developments.
M9	Monitor and report energy consumption at site and project level, provide information of the individual use per scientist/project/publication.
M10	Extend monitoring of resources beyond CO_2e (water, material etc.).
M11	Train scientists in good practices.
M12	Regularly review and update the CO_2e reduction plan.
M13	Strive to become a role model at all levels and help to establish sustainability in everyday life.
	A **longer term** coordinated planning is required for the following measures:
L1	Adjust computing in space and time to the availability of renewable energy, e.g. computing centers close to off-shore wind parks with a job scheduling using only or mainly the surplus of renewable energy available at a given time.
L2	Develop software and middleware that can respond dynamically to the availability of renewable energy.
L3	Optimize power usage effectiveness.
L4	Re-use of produced heat.
L5	Adjust hardware lifetime considering emissions due to procurement and operation.
L6	Include the resources needed for continuous IT support into project planning.

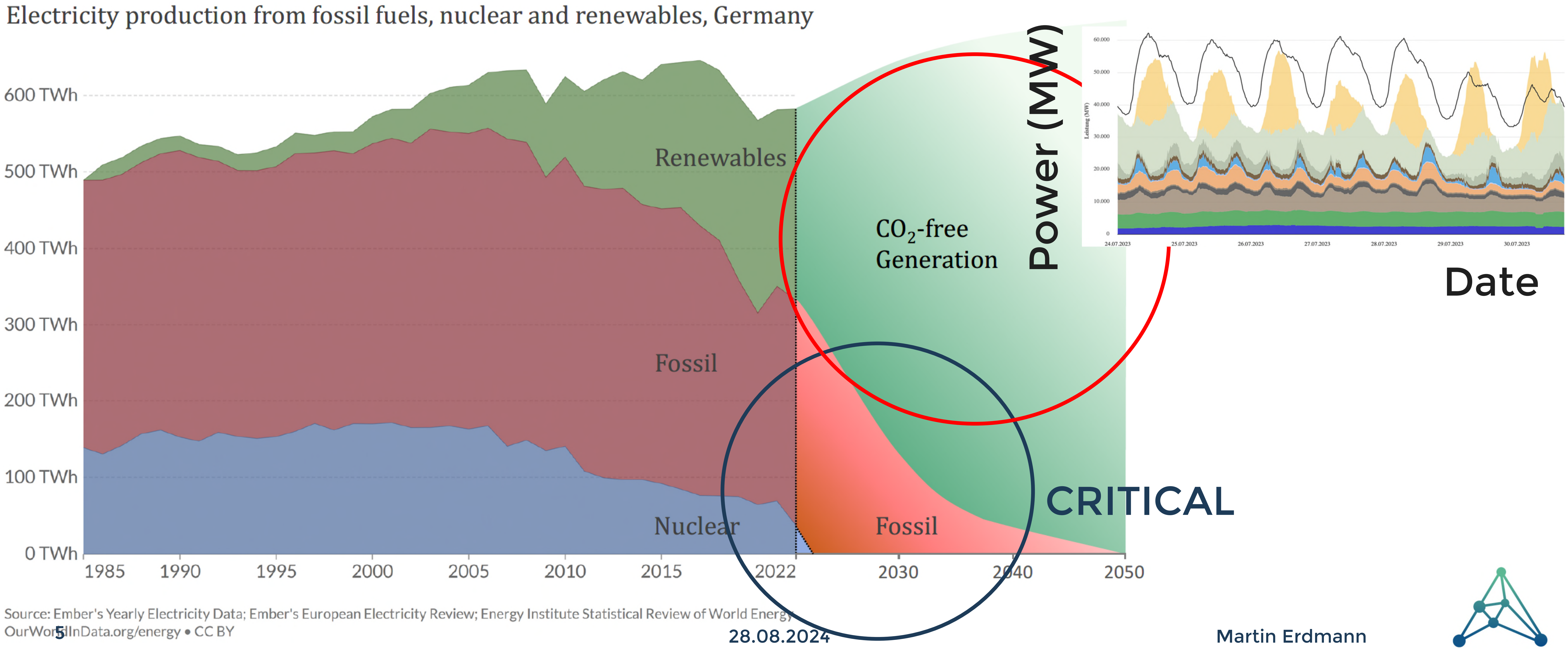
Portfolio of measures ordered in time scales

CO2 & worldwide energy generation





Electricity generation in Germany



LANCIUM, Texas USA

Lancium Footprint: 5,000 acres with ~7.5 GW Potential Capacity

LANCIUM




Name	1 Fort Stockton	2 Abilene	3 Childress	4 – 8 Various
Description / Status	Lancium Clean Campus 1 Under Construction 111 Acres	Lancium Clean Campus 2 Under Construction 1,000+ Acres	Lancium Clean Campus 3 In Planning 1,400 Acres	Lancium Clean Campus 4-8 In Planning 1,861+ Acres
Target Date	Q2-2022 (Phase 1) Q1-2023 (Phase 2)	1Q-2023 (Phase 1) Q2-2024 (Phase 2)	Q1-2025	2024+
Installed Capacity	25 MW (Phase 1) 300 MW (Phase 2)	200 MW (Phase 1) +1,000 MW (Phase 2)	1,000 MW	5,000+ MW
Key Stats	Operational	All utility work done	Interconnection approved.	Under Lancium control
Interconnection Status	TNMP 25MW Load Request Contracted & Paid TNMP 300MW Load Request Contracted & Paid	AEP 200MW Load Request Contracted & Paid AEP 1,000MW Load Request Approved	AEP 1,000MW approved.	NA

move computing site for scientific computing to energy provider

Aachen: VISPA – small-size computing cluster under our control

- **Interactive analyses in the browser**

- Small scale cluster with:
  - 14 machines with **30 GPUs**
  - 2.2 TB RAM and around 500 GB VRAM
  - Software installation for Data Analysis and Deep Learning

- Users (>2000 registered):
  - ~10 Academic researchers: CMS, Pierre-Auger (cosmic-ray)
  - Students in experimental physics lectures (~100 users)

HTCondor

- Test bed for enhancing sustainability in physics research

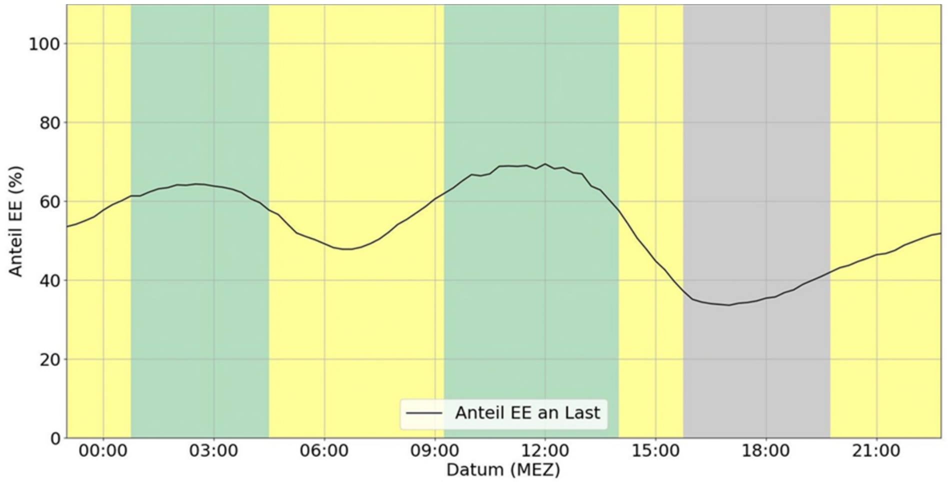
Steps towards Sustainability (team of 3 students, supported by our experts)

③ Sustainable Operation
- e.g. automated scheduling

② User Behaviour
- e.g. raise resource-awareness

① Monitoring
- e.g. monitor power consumption for informed decisions

Traffic light prediction renewables



Current efforts towards: users indicate urgency of their jobs at submission

- renewables
- urgent

PETRA BOT

Programmed Energy Tracker for more Resource Awareness

User-specific power consumption

- Calculation of power consumption per user

$$E_{user} = \sum_{machines} \frac{R_{user,machine}}{\sum_{user} R_{i,machine}} \cdot E_{machine}$$

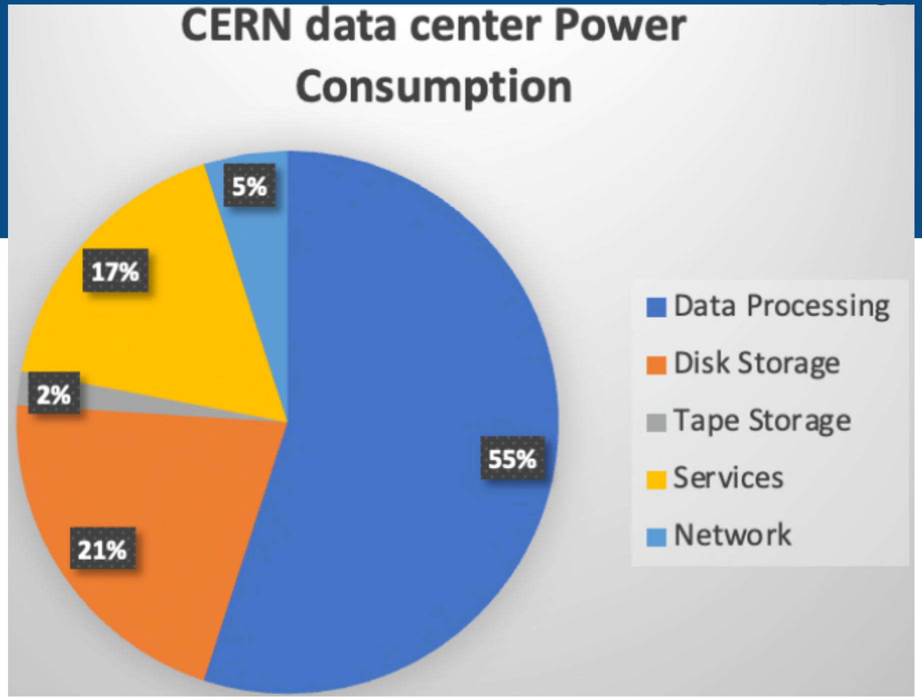
- The resource usage is given by

$$R_{user,machine} = \sum_{jobs} U \cdot Runtime$$

- U is the CPU or GPU (memory) usage
- CPU and GPU power consumption of a user is summed up in the end

➢ Method is accurate for the heterogenous user group of our cluster

Hardware



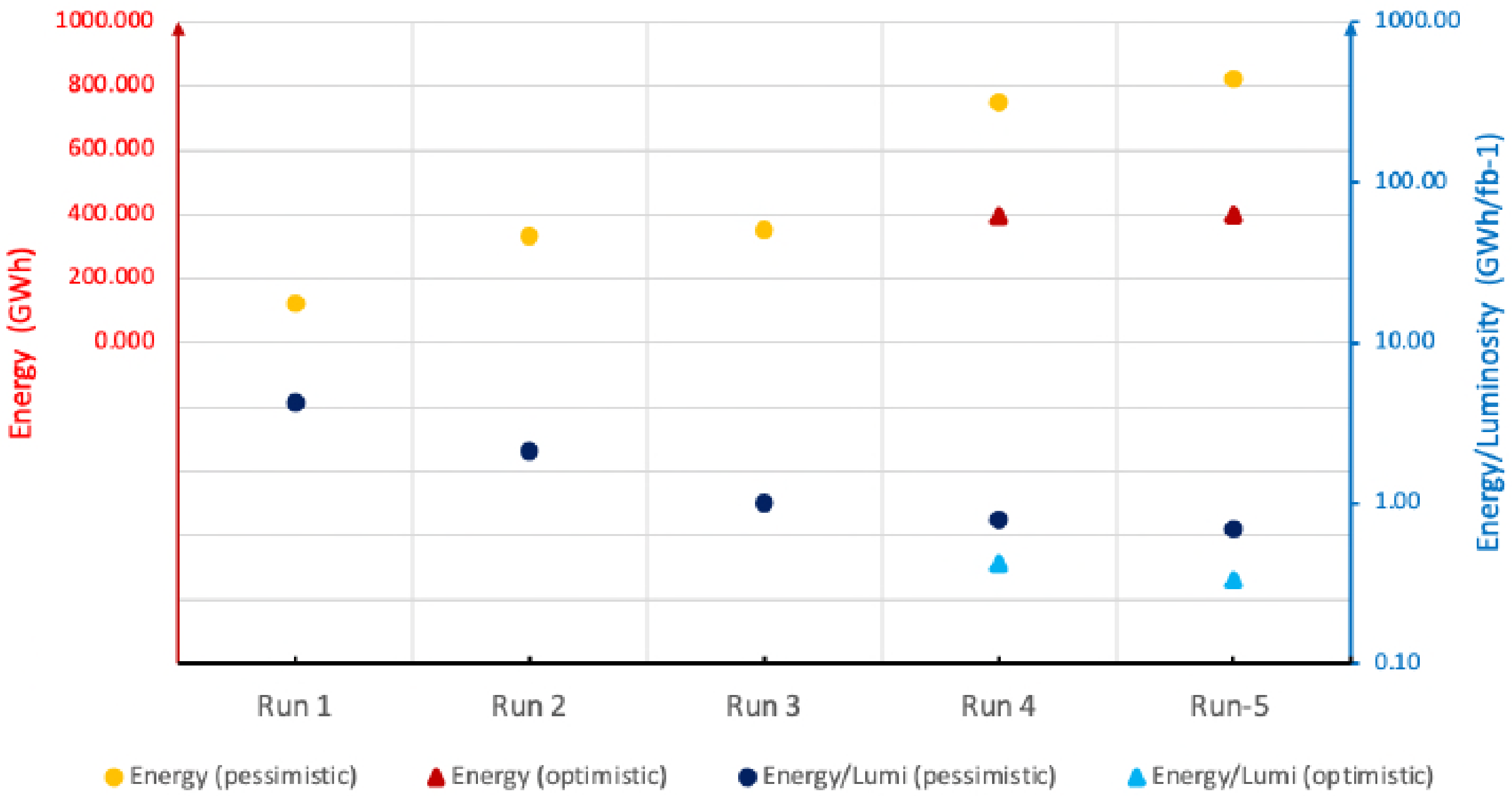
PUE (Power Usage Efficiency) factor 1.35 – 1.65

dual AMD 7302 processors, 4 TB SSD, 256 GB memory, 10 Gbit NIC

- a performance value of 1040 HS06,
- an idle power value of 120 W
- and a full load value of up to 420 W

we assume a 5 years lifecycle

GWh/fb^{-1} for WLCG



https://wiki.erumdatahub.de/2023_sustainability/wlcgenergyneeds.pdf

Conclusions

- Moral/ethical motivation

- I strongly believe: resource awareness will boost scientific progress: new methods, new technologies, more direct paths towards obtaining physics message