Towards Climate Awareness in NLP Research

Daniel Hershcovich
Department of Computer Science
University of Copenhagen
dh@di.ku.dk

Nicolas Webersinke **Mathias Kraus**
FAU Erlangen-Nuremberg
{nicolas.webersinke,
mathias.kraus}@fau.de

Julia Anna Bingler
ETH Zurich
binglerj@ethz.ch

Markus Leippold
University of Zurich
markus.leippold@bf.uzh.ch

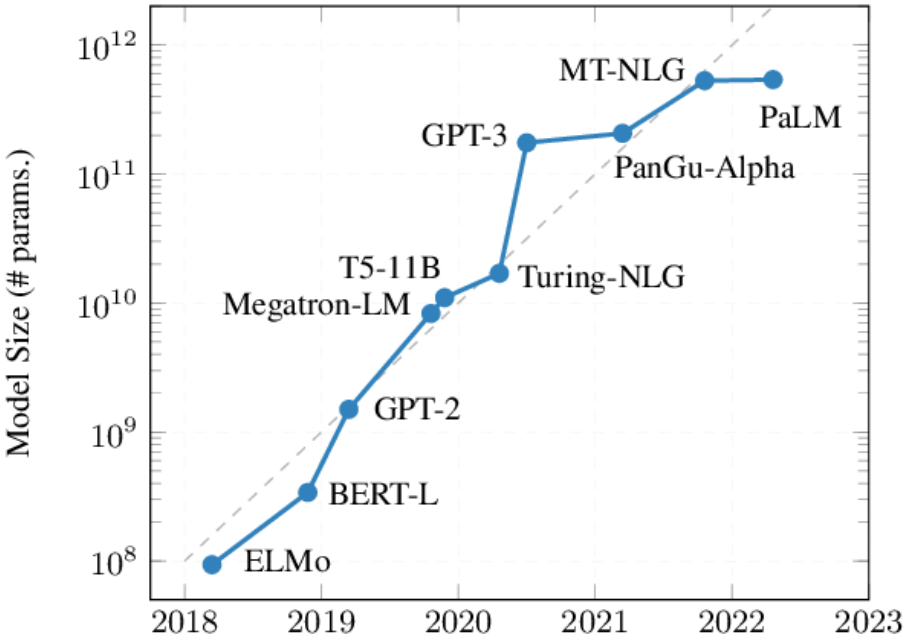
EMNLP 2022

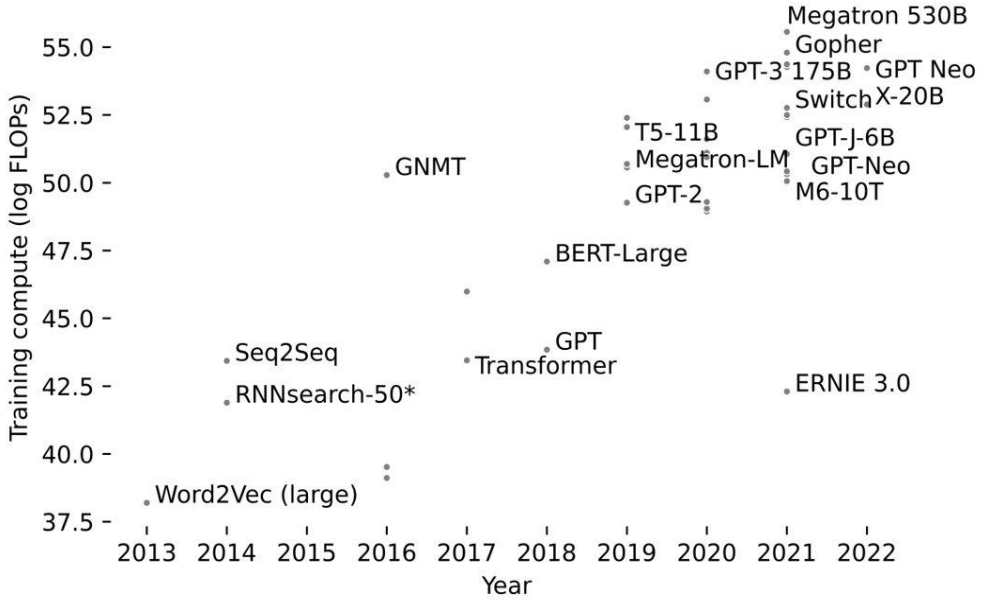
UNIVERSITY OF COPENHAGEN

https://arxiv.org/abs/2205.05071

Climate impact of NLP

Models grow exponentially and take longer to train

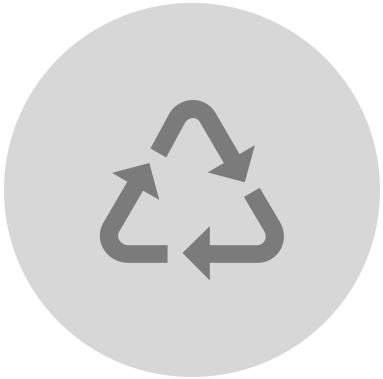




Efficient Methods for Natural Language Processing: A Survey (Treviso et al., 2022)

Climate performance model card

1. Model publicly available?

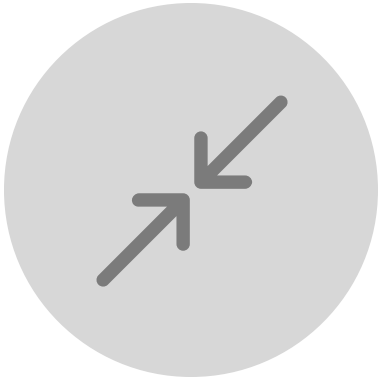


REUSE

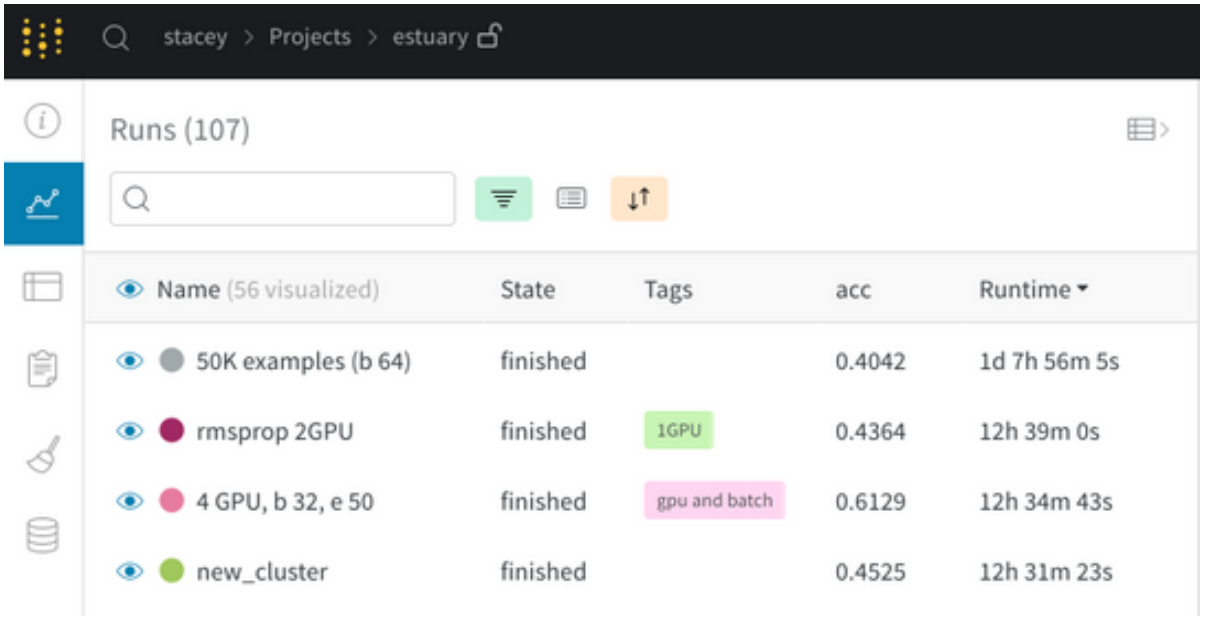
https://huggingface.co/models

Climate performance model card

1. Model publicly available?
2. Time to train final model
3. Time for all experiments



REDUCE



https://wandb.ai/

Energy consumption scales with size

Model	Consumption [MWh]
1.3B	2.1
2.7B	4.8
6.7B	11.8
13B	22.9

A Holistic Assessment of the Carbon Footprint of Noor, a Very Large Arabic Language Model (Lakim et al., BigScience 2022)

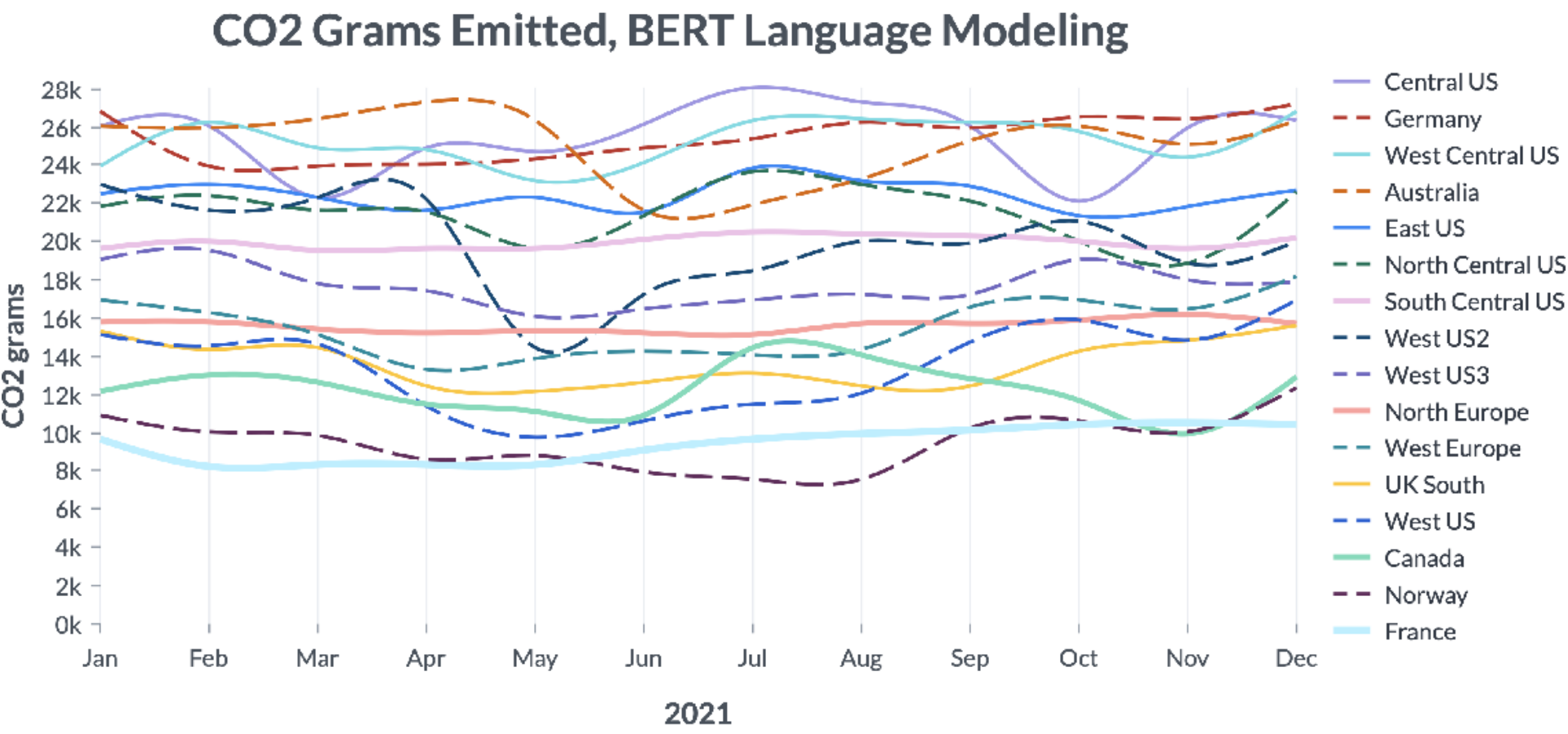
Climate performance model card

1. Model publicly available?
2. Time to train final model
3. Time for all experiments
4. Power of GPU and CPU

	A100 80GB PCIe	A100 80GB SXM
GPU Memory	80GB HBM2e	80GB HBM2e
GPU Memory Bandwidth	1,935 GB/s	2,039 GB/s
Max Thermal Design Power (TDP)	300W	400W ***

https://www.nvidia.com/en-us/data-center/a100/#specifications

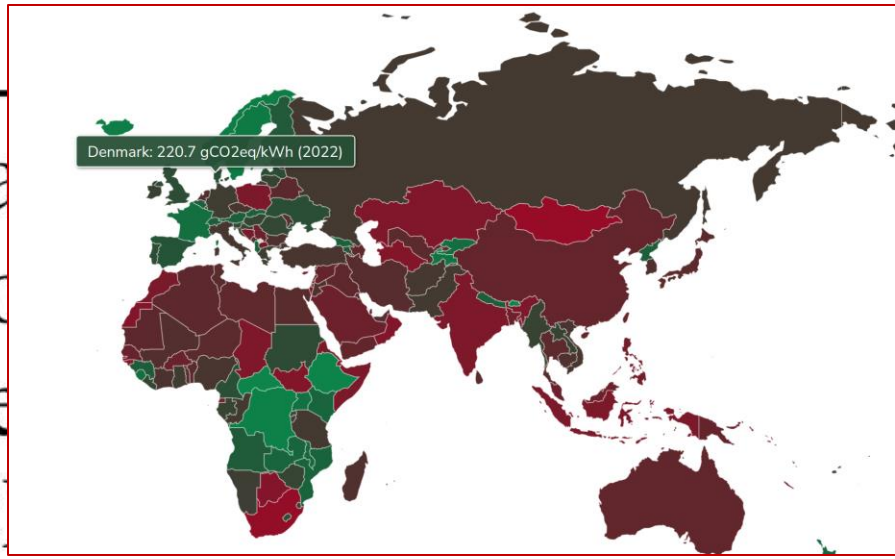
CO_2e footprint depends on time and location



Measuring the Carbon Intensity of AI in Cloud Instances (Dodge et al., FAccT 2022)

Climate performance model card

1. Model publicly availa
2. Time to train final mo
3. Time for all experime
4. Power of GPU and C
5. Location for computations
6. Energy mix at location

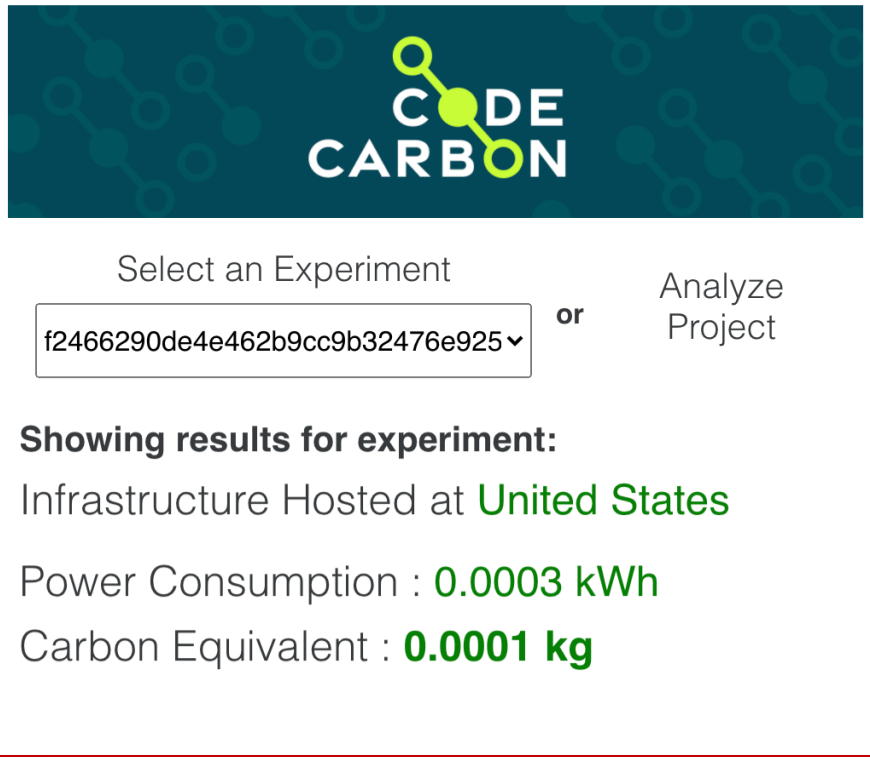


https://lowcarbonpower.org/map-gCO2eq-kWh

RELOCATE

Climate performance model card

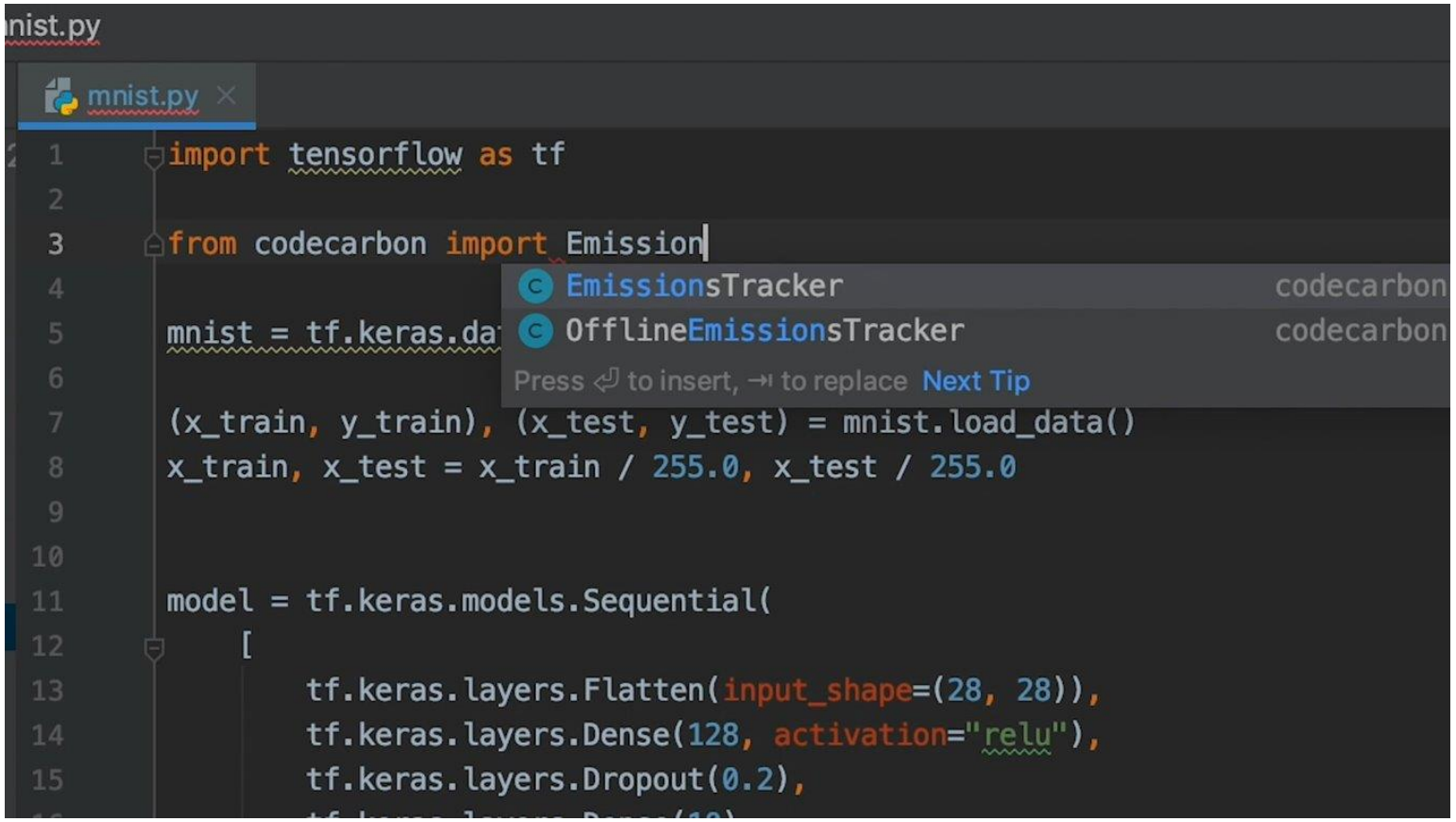
1. Model publicly availa
2. Time to train final mo
3. Time for all experime
4. Power of GPU and C
5. Location for computa
6. Energy mix at locatio
7. CO_2eq for final model
8. CO_2eq for all experiments



https://codecarbon.io/

Automated reporting

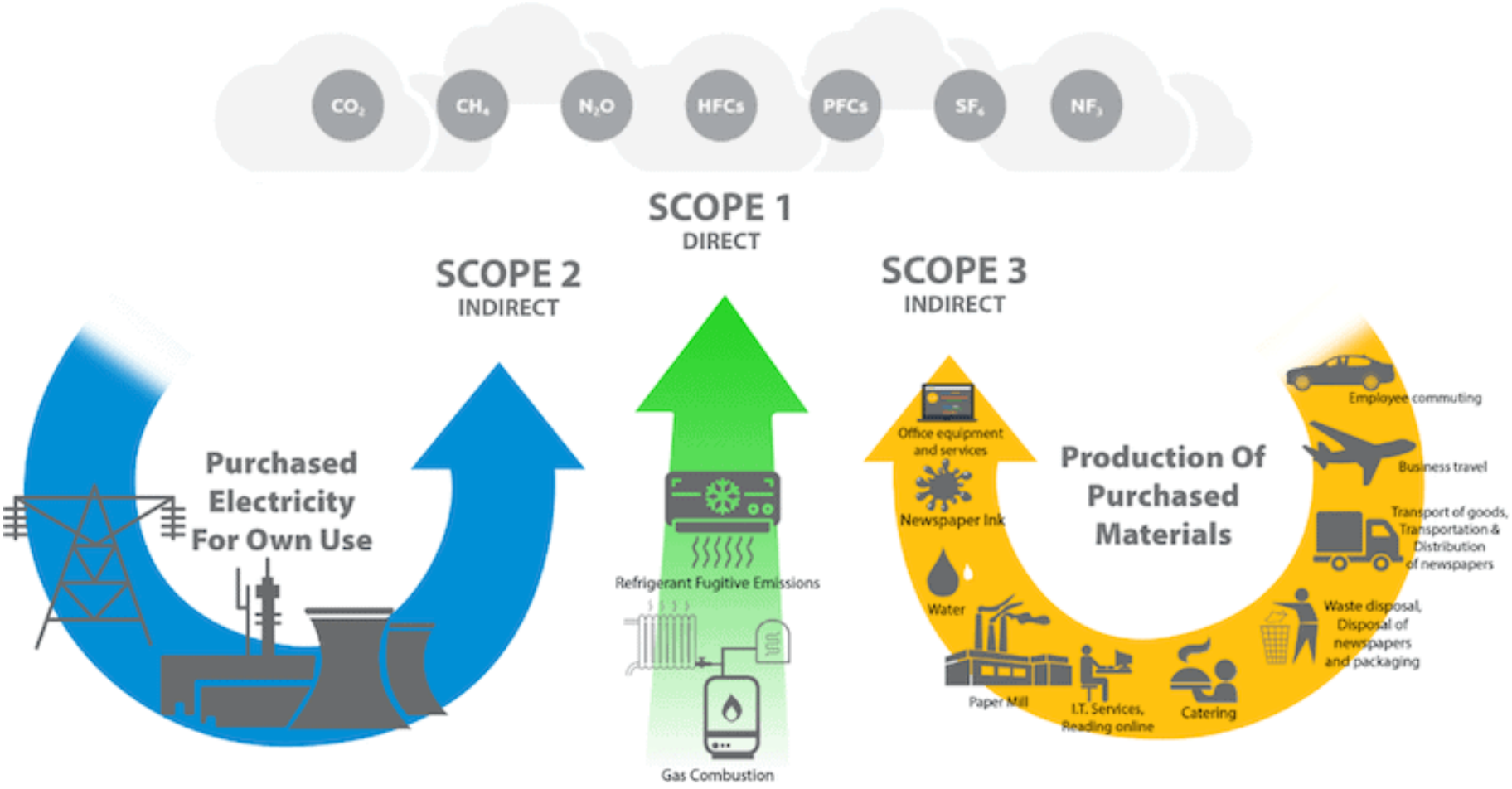
Convenient and reproducible



But measures vary significantly between tools

Evaluating the carbon footprint of NLP methods: a survey and analysis of existing tools (Bannour et al., SustaiNLP 2021)

Greenhouse Gas Protocol

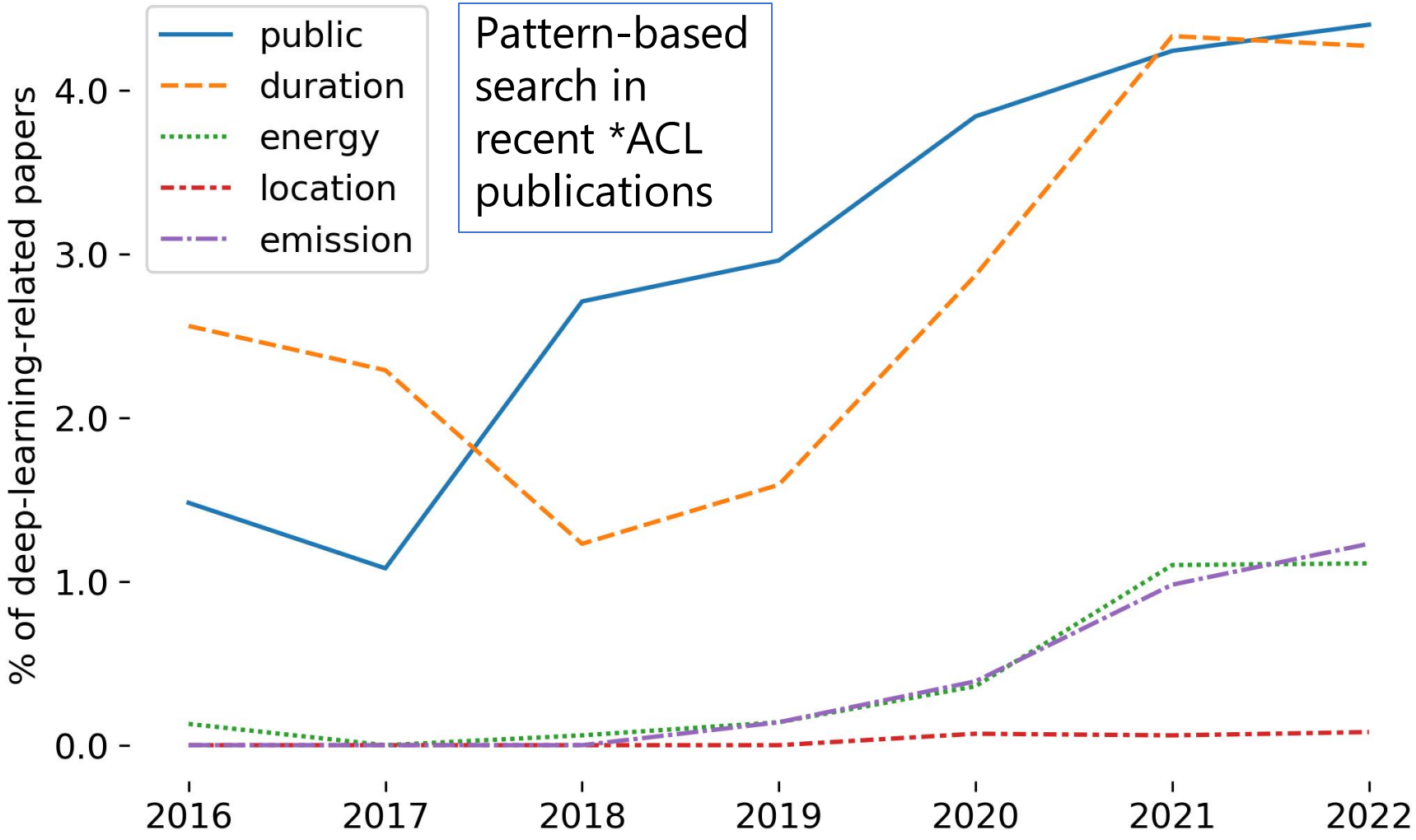


https://www.greenelement.co.uk/blog/carbon-footprint-scope-1-2-3/

Climate performance model card

1. Model publicly available?
2. Time to train final model
3. Time for all experiments
4. Power of GPU and CPU
5. Location for computations
6. Energy mix at location
7. CO_2eq for final model
8. CO_2eq for all experiments
9. Average CO_2eq for inference per sample

Climate awareness in NLP



Climate awareness is increasing but is still "niche"

Greenwashing?

"Better than some alternative" ≠ sustainable

Principles of climate performance assessment

RELEVANCE

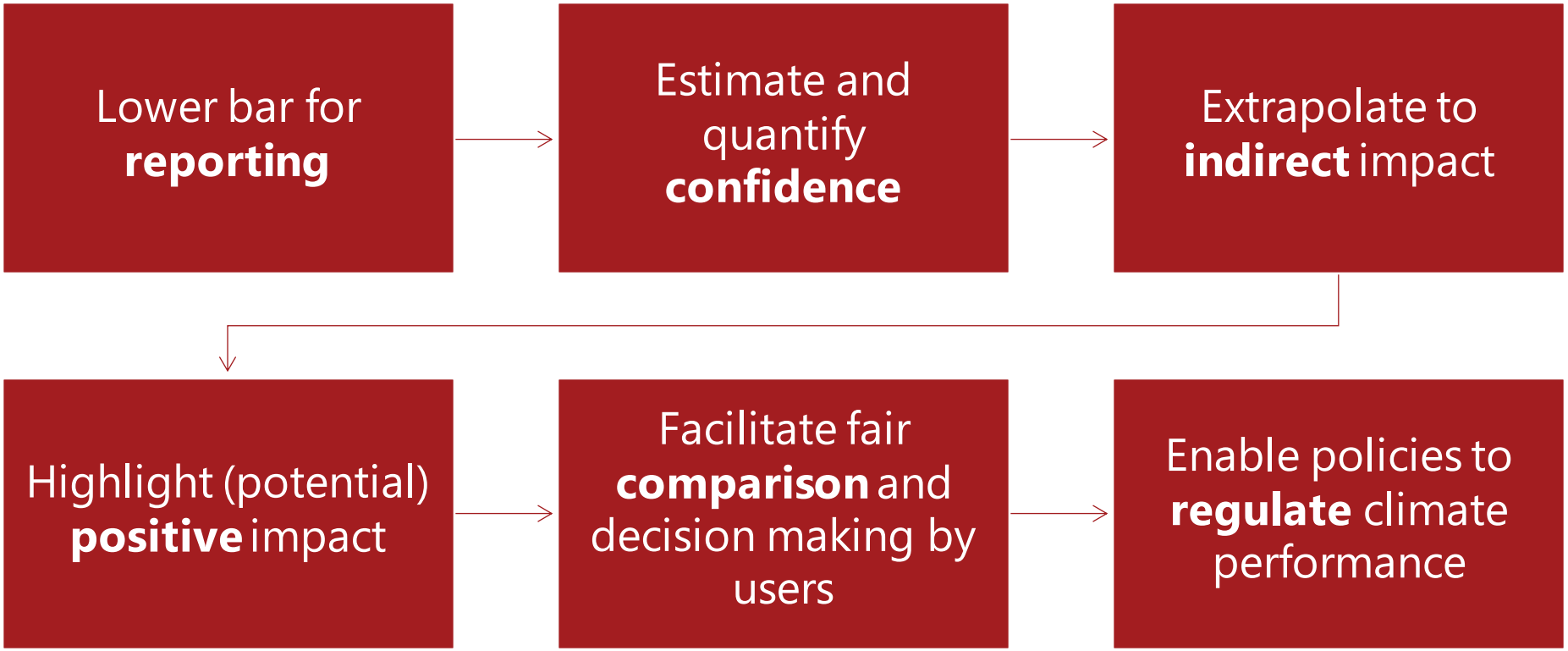
COMPLETENESS

CONSISTENCY

TRANSPARENCY

ACCURACY

Pragmatic reporting



Carbon offsetting?

Compensating for emissions is controversial

Mitigation by reduction is essential and feasible

Model card example

ClimateBert	
1. Model publicly available?	Yes
2. Time to train final model	8 hours
3. Time for all experiments	288 hours
4. Power of GPU and CPU	0.7 kW
5. Location for computations	Germany
6. Energy mix at location	470 gCO_2eq/kWh
7. CO_2eq for final model	2.63 kg
8. CO_2eq for all experiments	94.75 kg
9. Average CO_2eq for inference per sample	0.62 mg

ClimateBert: A Pretrained Language Model for Climate-Related Text (Webersinke et al., 2021)

Model card example

MCWQ `mT5-base+RIR`

Information	Unit
1. Model publicly available?	Yes
2. Time to train final model	592 hours
3. Time for all experiments	1315 hours
4. Energy consumption	2209.2 kWh
5. Location for computations	Denmark
6. Energy mix at location	191 gCO2eq/ kWh
7. CO2eq for final model	189.96 kg
8. CO2eq for all experiments	421.96 kg

Compositional Generalization in Multilingual Semantic Parsing over Wikidata (Cui et al., TACL 2022)

Some avenues for positive climate impact

Detecting greenwashing
(e.g., Bingler et al., 2022)

Analyzing public discourse
(e.g., Hansen & Hershcovich, NLP4PI 2022)

Understanding consumer behavior

Facilitating behavior change

…

How Good Is NLP? A Sober Look at NLP Tasks through the Lens of Social Impact
(Jin et al., Findings 2021)

Summary

Climate awareness should be mainstream

REDUCE

REUSE

RELOCATE

Consistent and pragmatic reporting via model cards

Code and templates: github.com/danielhers/climate-awareness-nlp

Efficient is not enough – net positive impact needed

Contact:
danielhers.github.io
dh@di.ku.dk