UNIVERSITY OF CAMBRIDGE
Department of Public Health and Primary Care
British Heart Foundation
NIHR | Cambridge Biomedical Research Centre
HDRUK Health Data Research UK
Baker HEART & DIABETES INSTITUTE
Software Sustainability Institute
Jesus College Cambridge

ESTIMATING THE CARBON FOOTPRINT OF COMPUTING IN SCIENCE

Loïc Lannelongue, PhD
Research Associate, University of Cambridge
Software Sustainability Institute Fellow
Post-Doctoral Associate, Jesus College, Cambridge

2023 EIC Workshop

LL582@medschl.cam.ac.uk | *@Loic_Lnlg* | *www.lannelongue.eu*

WHAT ABOUT THE SCIENCE WE DO? **BIOLOGY**

GWAS of 1,000 traits in UK Biobank

(225h / 100 GB per trait)

17.3 T CO_2e

Metagenome assembly
of 100 soil samples
(1h / 130 GB)

14 kg CO_2e

RNA read alignments
10 million 100-bp to *P. falciparum*
(1h30 / 13 GB)

240 g CO_2e

WHAT ABOUT THE SCIENCE WE DO? AI + BIOLOGY

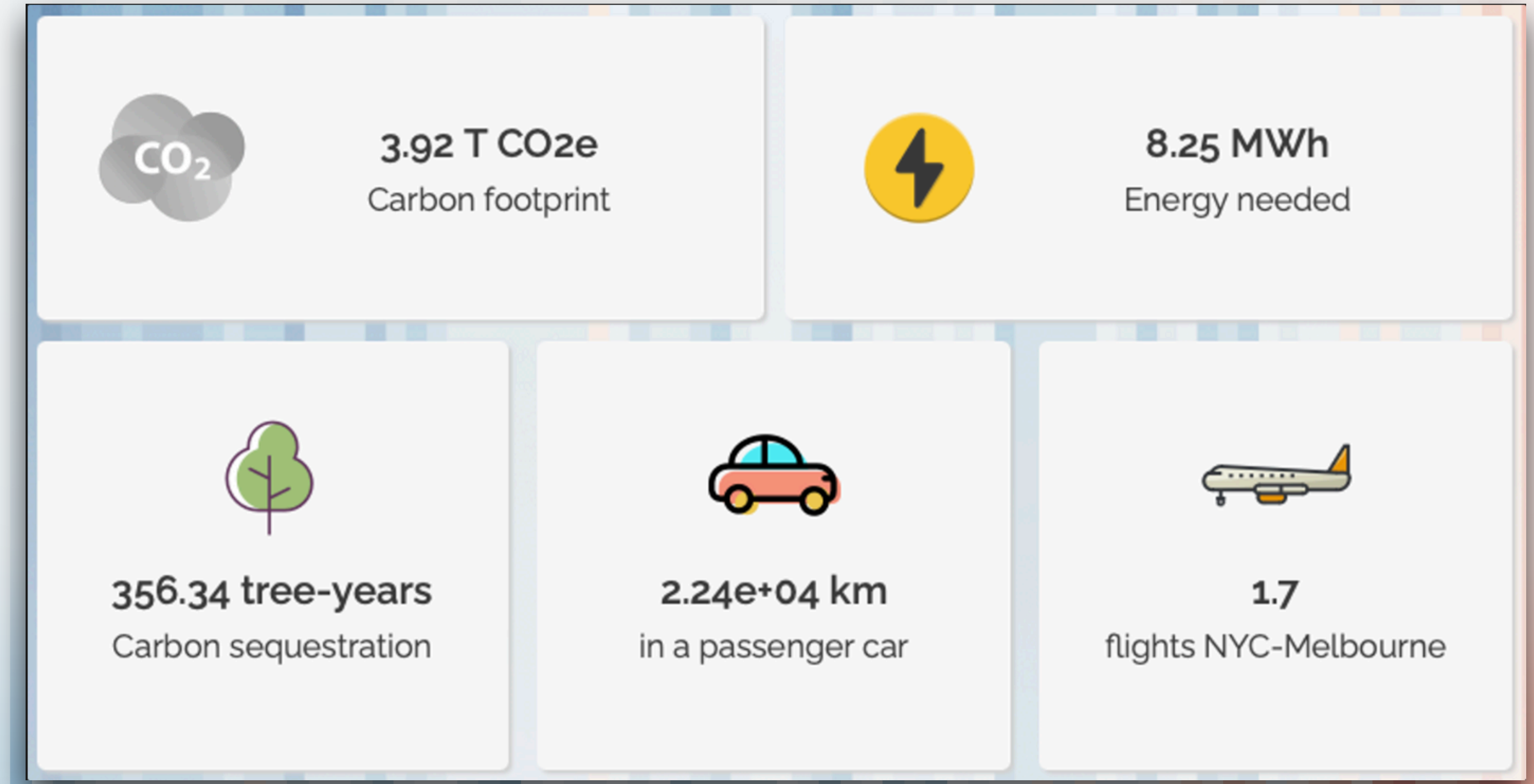
Article | Open Access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, ... Demis Hassabis + Show authors

Nature **596**, 583–589 (2021) | Cite this article

Training



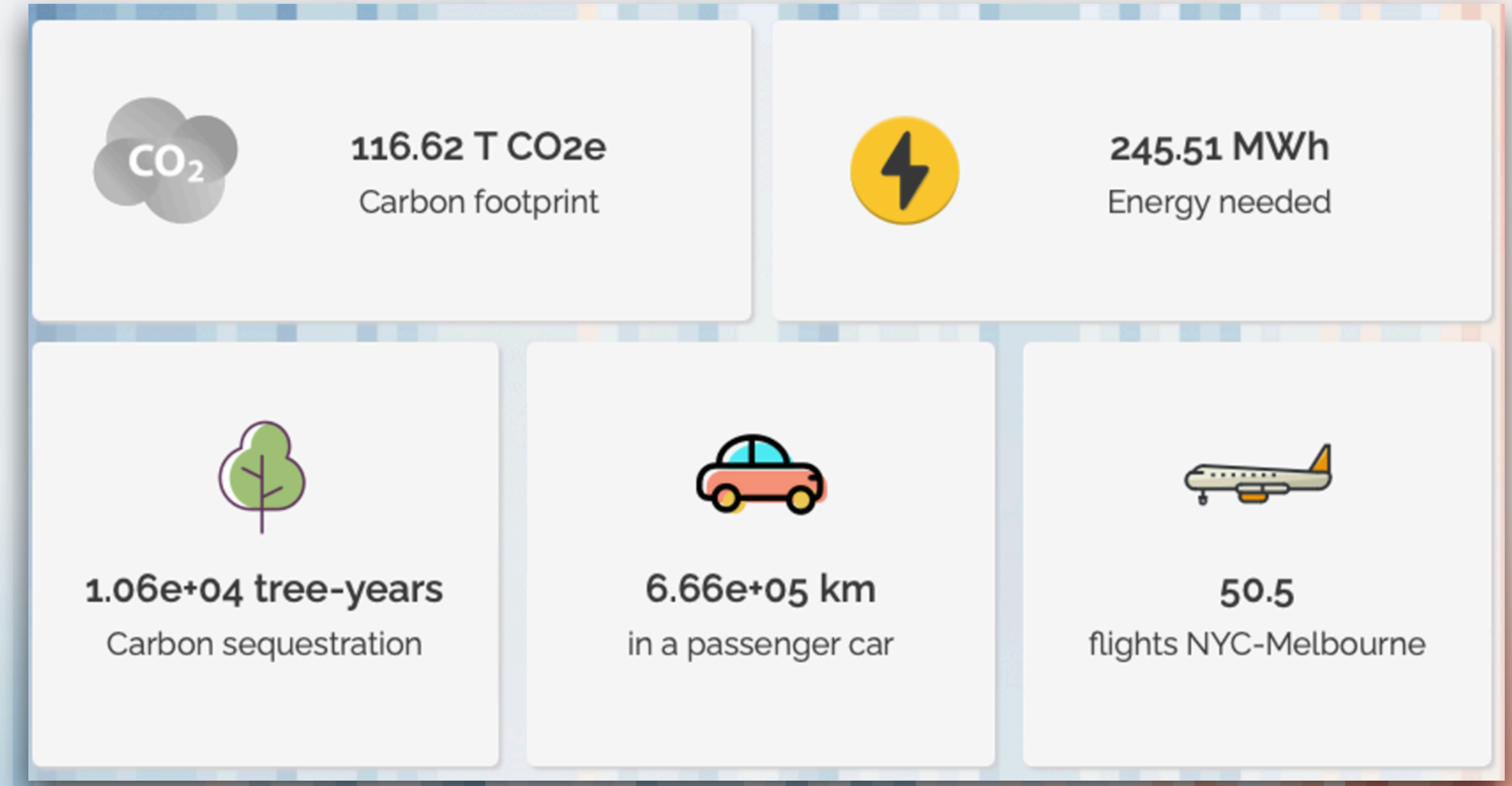
WHAT ABOUT THE SCIENCE WE DO? AI + BIOLOGY

RESEARCH ARTICLE | STRUCTURE PREDICTION

Evolutionary-scale prediction of atomic-level protein structure with a language model

ZEMING LIN, HALIL AKIN, ROSHAN RAO, BRIAN HIE, ZHONGKAI ZHU, WENTING LU, NIKITA SMETANIN, ROBERT VERKUIL, ORI KABELI, [...], AND ALEXANDER RIVES +5 authors Authors Info & Affiliations

Training the 15B model



UNFORTUNATELY…

Computing is **not** free

But from a **user** point of view, **it may look like it**

FOCUSING ON **COMPUTING**

Day-to-day computing
Emails
Writing on Words
Web surfing
Zoom

Intense computations:
long runtimes (several hours)
and/or large memory requirements (10s GB)

FOCUSING ON **COMPUTING**

Day-to-day computing

Emails

Writing on Words

Web surfing

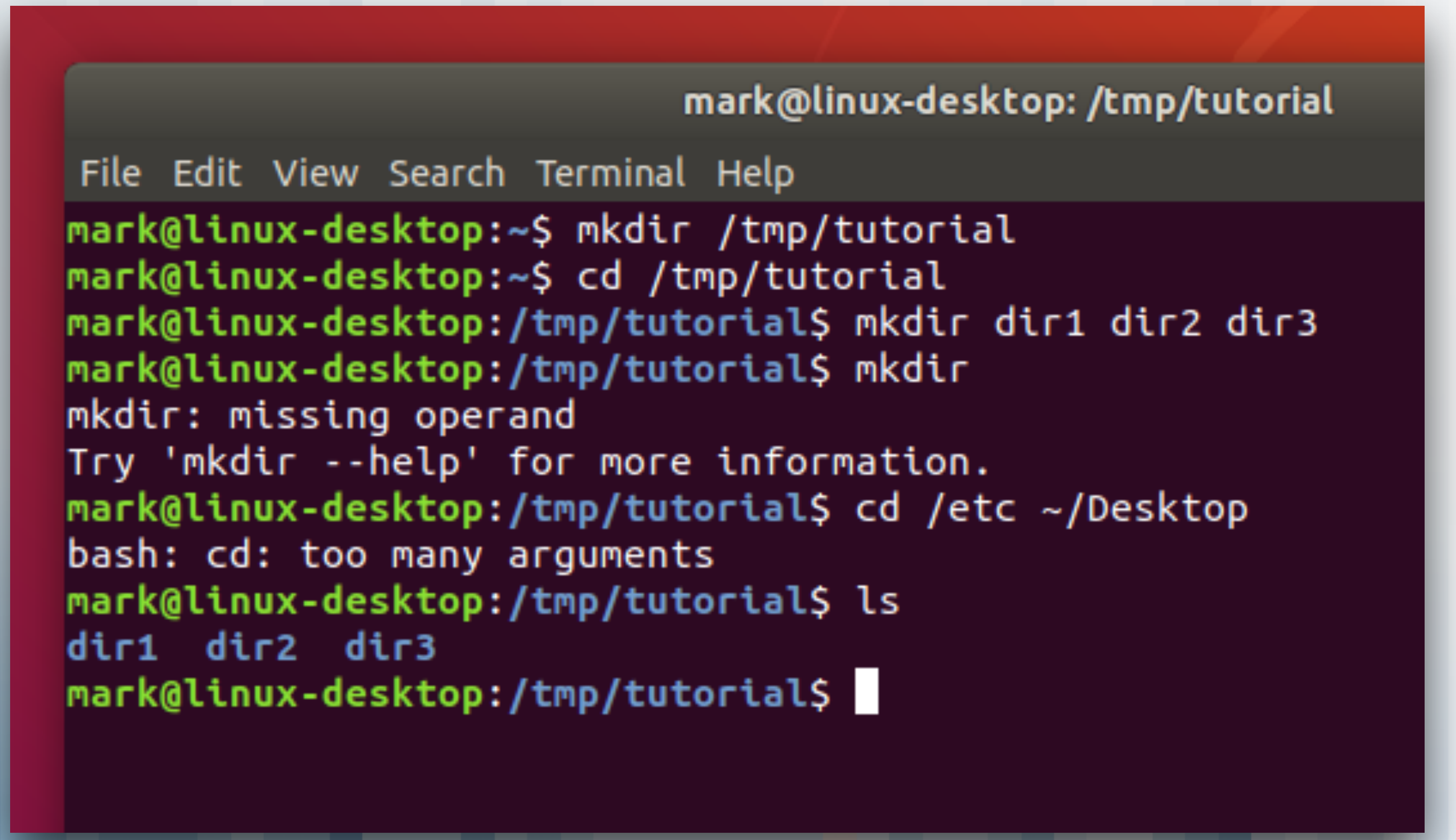
Zoom

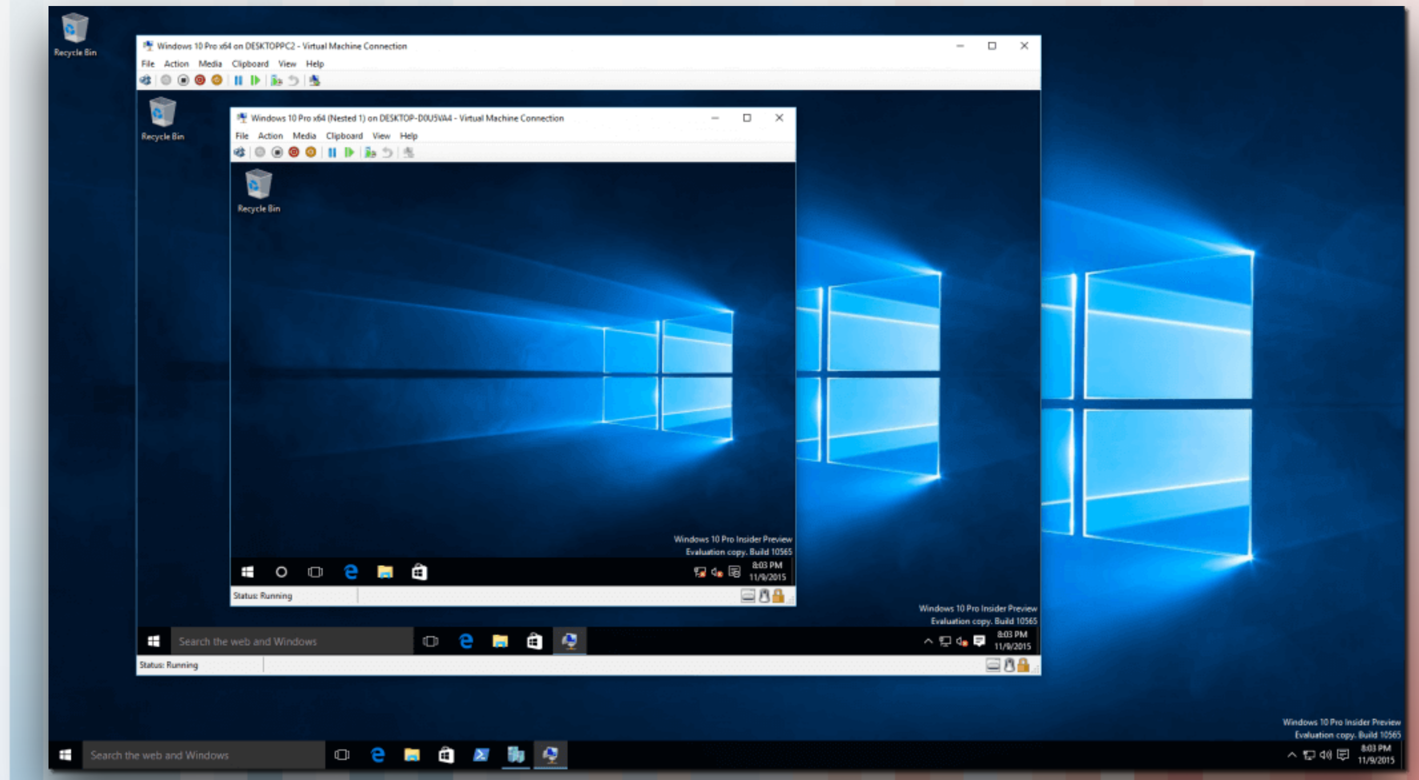
Intense computations:

long runtimes (several hours)

and/or large memory requirements (10s GB)

IT'S ALL THE SAME (ISH)





BREAKING DOWN THE ENVIRONMENTAL IMPACTS OF COMPUTING

Environmental impact
of computing

BREAKING DOWN THE ENVIRONMENTAL IMPACTS OF COMPUTING

Environmental impact of computing

Powering the computer

Long-term data storage

Life-cycle footprint of the hardware

BREAKING DOWN THE ENVIRONMENTAL IMPACTS OF COMPUTING

Environmental impact of computing

Powering the computer

Long-term data storage

Life-cycle footprint of the hardware

Keep, Repair, Reuse

70-90% of the cradle-to-grave impact is from production (consumer devices)

BREAKING DOWN THE ENVIRONMENTAL IMPACTS OF COMPUTING

Environmental impact of computing

Powering the computer

Long-term data storage

Life-cycle footprint of the hardware

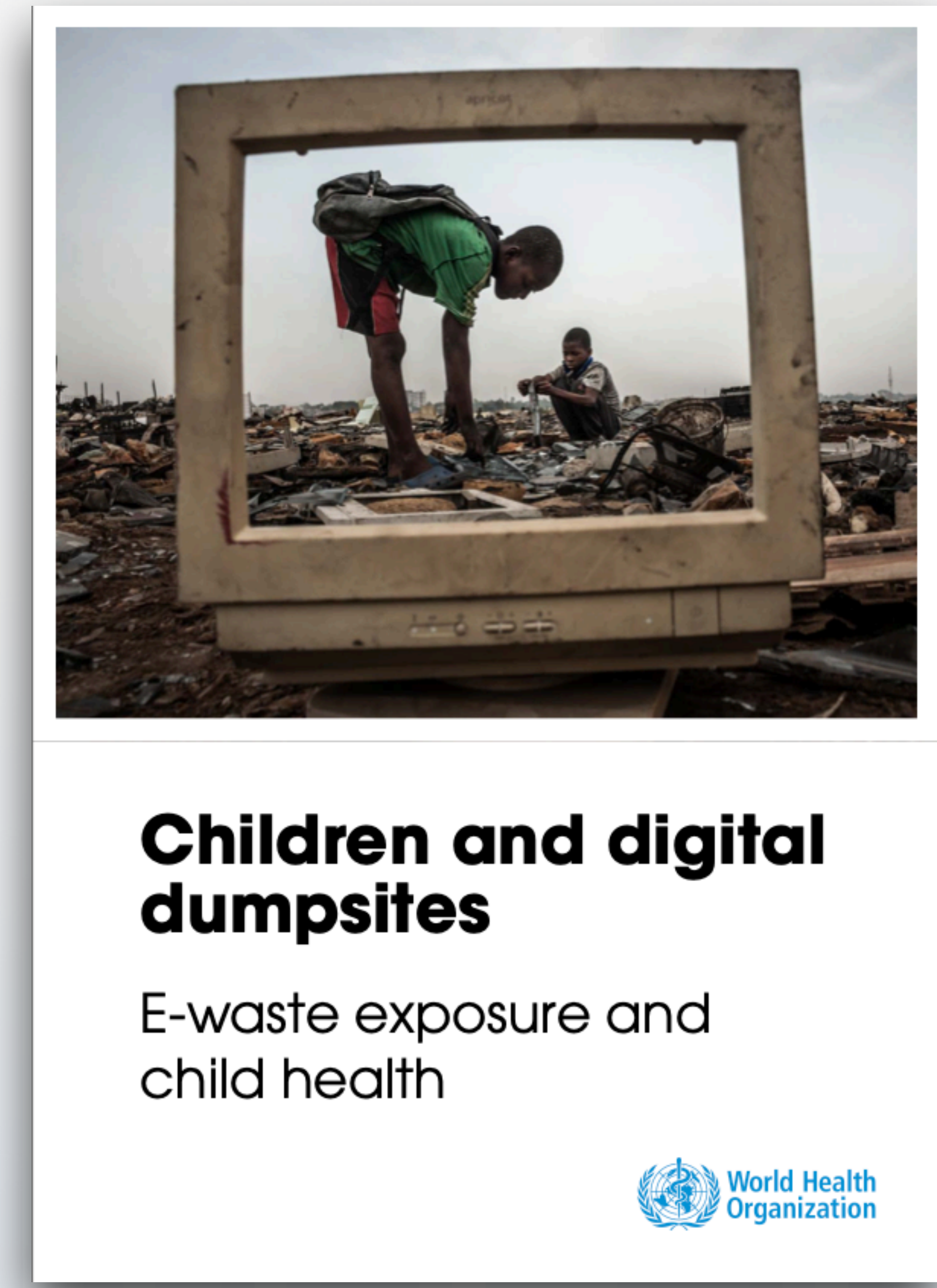
Sustainability should be accounted for in renewing policies

15-40% for servers in data centres

Keep, Repair, Reuse

70-90% of the cradle-to-grave impact is from production (consumer devices)

BREAKING DOWN THE ENVIRONMENTAL IMPACTS OF COMPUTING



Environmental impact of computing

computer

Long-term data storage

Life-cycle footprint of the hardware

Keep, Repair, Reuse

>82% of the 54m of tonnes of e-waste are handled by 12-56m informal waste workers worldwide

18m children work in industries involving waste processing

E-waste are predicted to raise by 40% by 2030

BREAKING DOWN THE ENVIRONMENTAL IMPACTS OF COMPUTING

Environmental impact of computing

Powering the computer

Energy bill of a data centre:

- *50% servers*
- *10% storage*
- *40% overheads (cooling)*

Long-term data storage

Life-cycle footprint of the hardware

BREAKING DOWN THE ENVIRONMENTAL IMPACTS OF COMPUTING

Environmental impact of computing

Powering the computer

Energy bill of a data centre:

- *50% servers*
- *10% storage*
- *40% overheads (cooling)*

Long-term data storage

~10 $kgCO_2e$/TB/year

Life-cycle footprint of the hardware

Don't store useless data

https://www.seagate.com/gb/en/global-citizenship/product-sustainability/

BREAKING DOWN THE ENVIRONMENTAL IMPACTS OF COMPUTING

Environmental impact of computing

Powering the computer

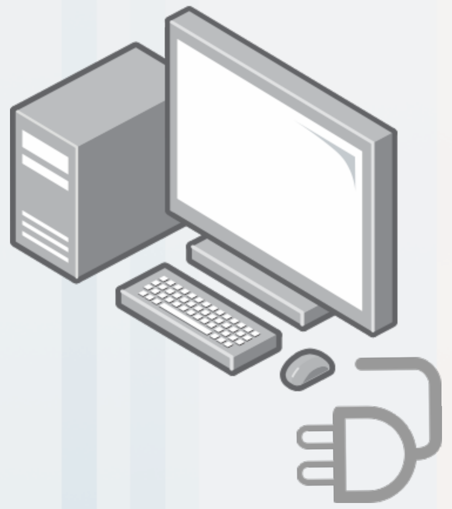
Long-term data storage

Life-cycle footprint of the hardware

THE CARBON FOOTPRINT OF COMPUTATION

Carbon footprint = energy used x carbon intensity

gCO_2e kWh gCO_2e/kWh



THE CARBON FOOTPRINT OF COMPUTATION: **ENERGY NEEDED**

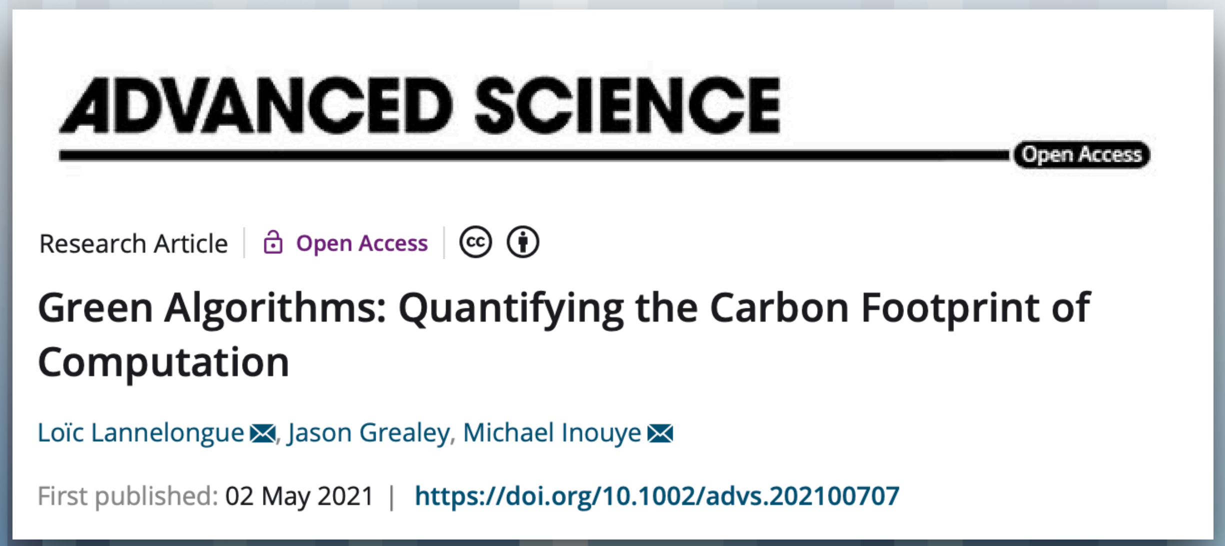
$$E = t \times (P_c + P_m) \times PUE$$

Running time (h)

Power draw of processing cores (W)

Power draw from memory (W)

Efficiency of the data centre


ADVANCED SCIENCE
Open Access

Research Article | Open Access

Green Algorithms: Quantifying the Carbon Footprint of Computation

Loïc Lannelongue, Jason Grealey, Michael Inouye

First published: 02 May 2021 | https://doi.org/10.1002/advs.202100707

THE CARBON FOOTPRINT OF COMPUTATION: **ENERGY NEEDED**

$$PUE = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$$

$$E = t \times (P_c + P_m) \times PUE$$

Running time (h)

Power draw of processing cores (W)

Power draw from memory (W)

Efficiency of the data centre

ADVANCED SCIENCE Open Access

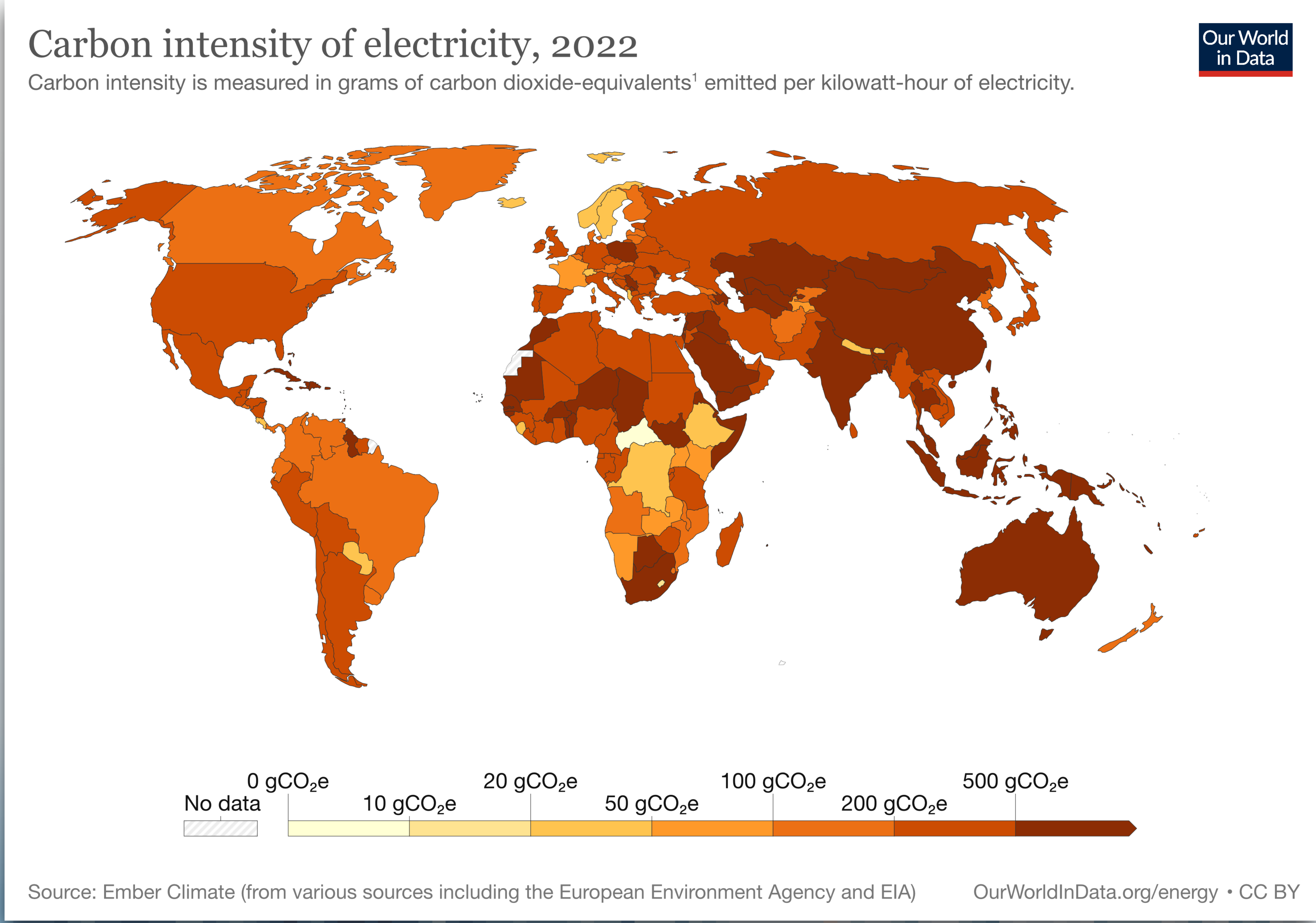
Research Article | Open Access

Green Algorithms: Quantifying the Carbon Footprint of Computation

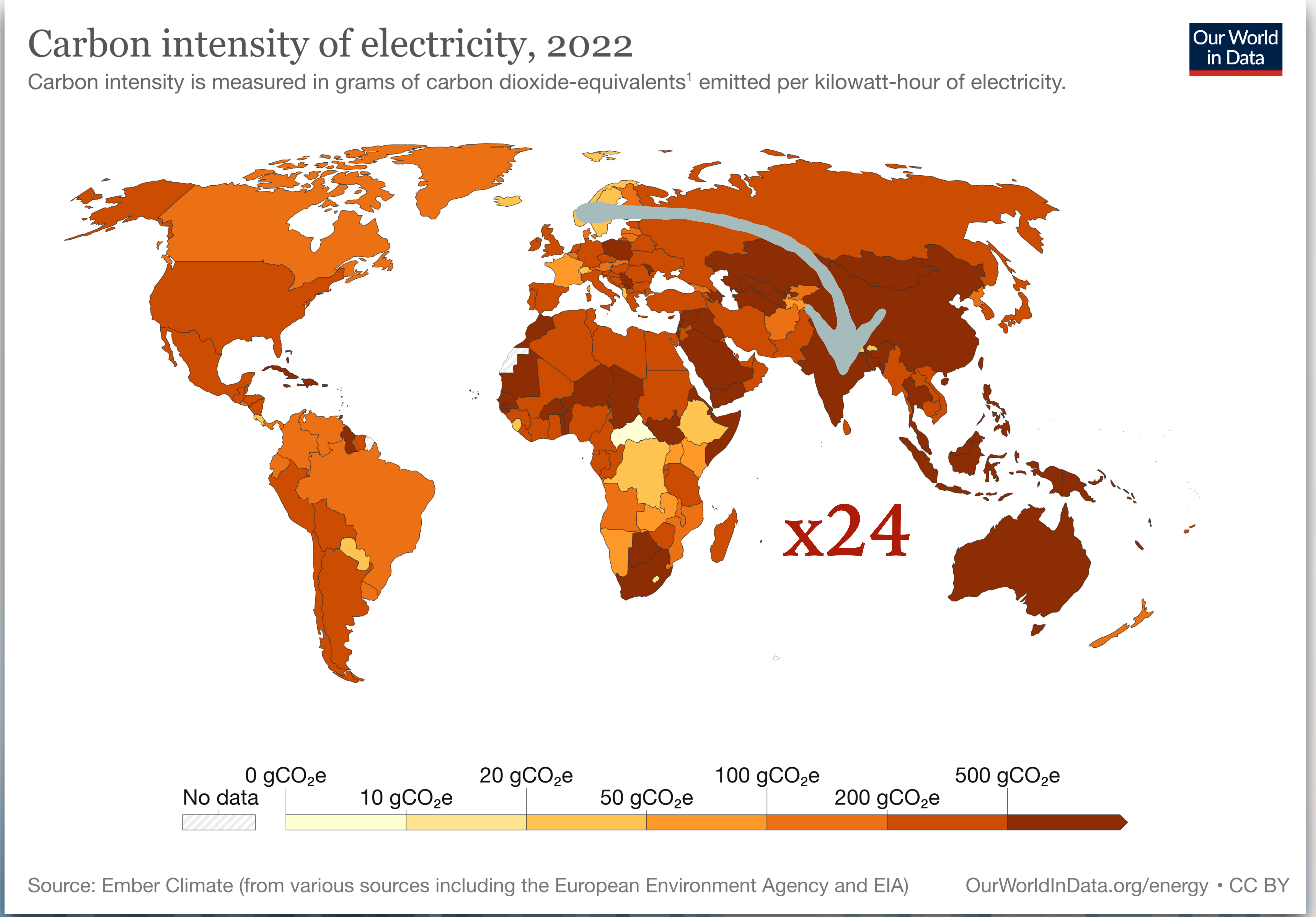
Loïc Lannelongue, Jason Grealey, Michael Inouye

First published: 02 May 2021 | https://doi.org/10.1002/advs.202100707

THE CARBON FOOTPRINT OF COMPUTATION: CARBON INTENSITY

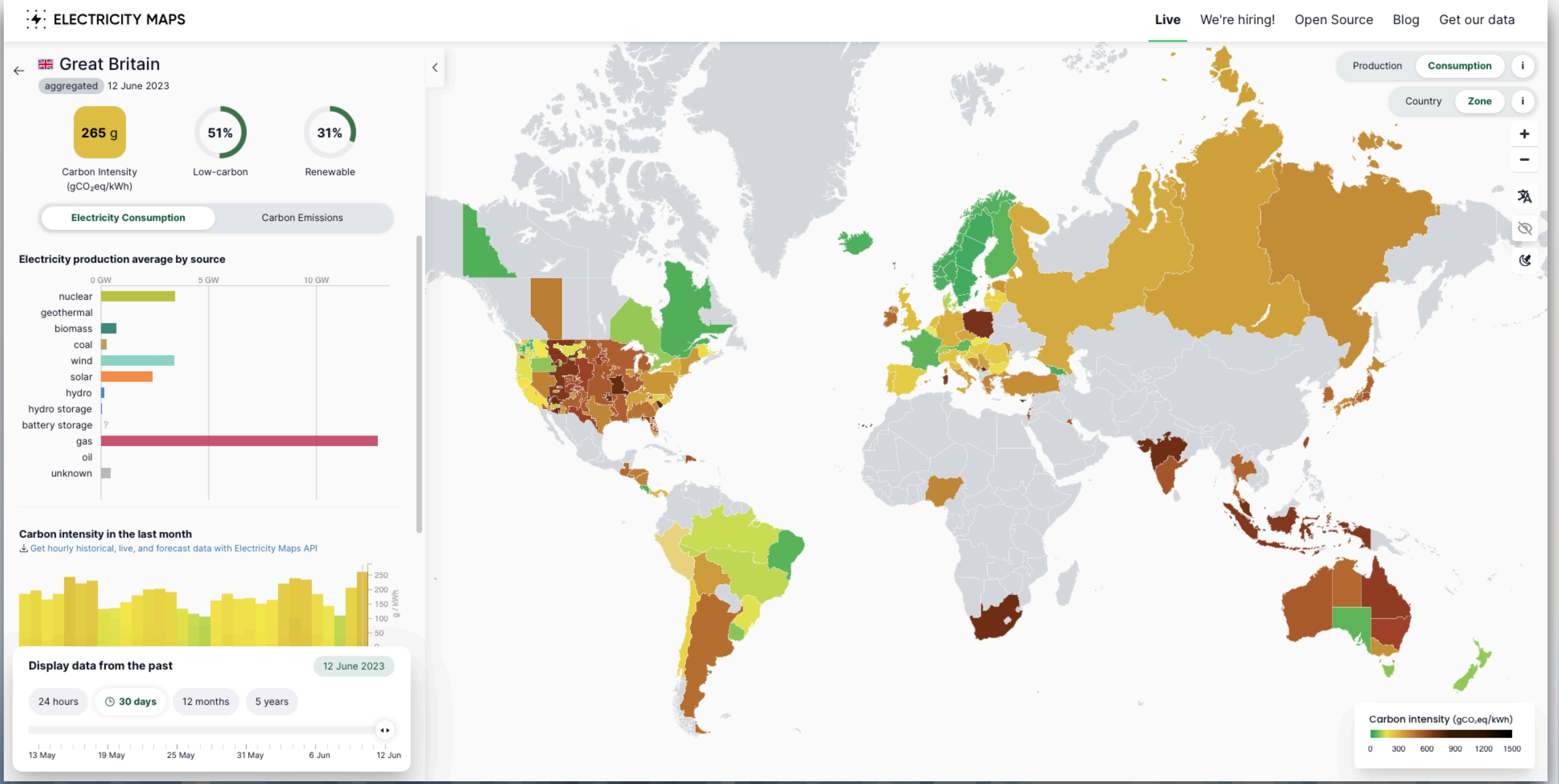


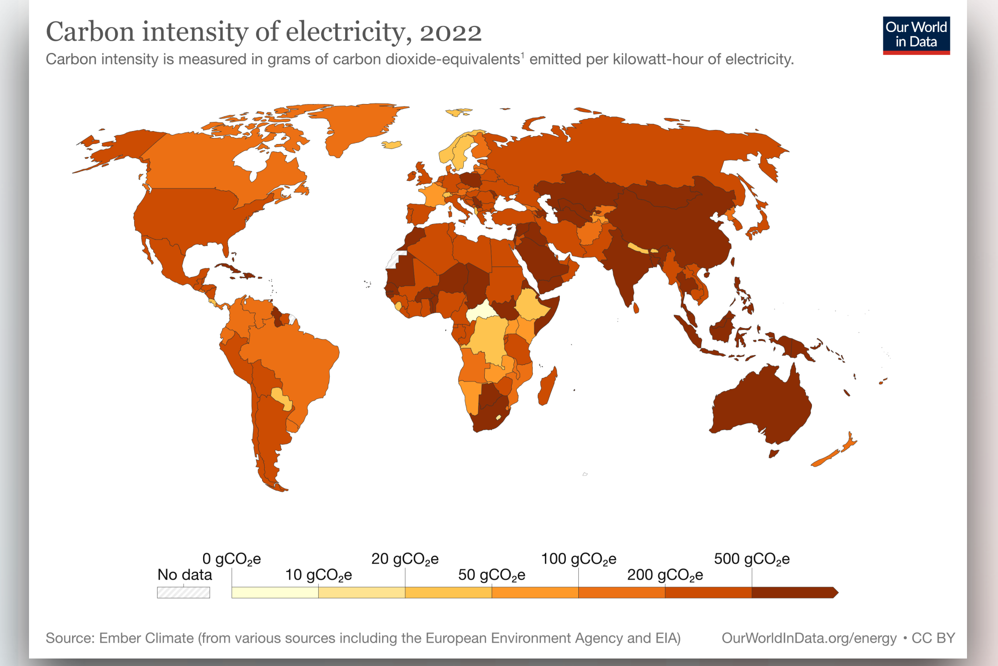
THE CARBON FOOTPRINT OF COMPUTATION: CARBON INTENSITY



THE CARBON FOOTPRINT OF COMPUTATION: **CARBON INTENSITY**

www.electricitymap.org





THE CARBON FOOTPRINT OF COMPUTATION: **CARBON INTENSITY**

https://carbonintensity.org.uk/

About National Regional new Methodology Docs Terms Contact Dark mode beta

Carbon Intensity API

National Grid ESO, in partnership with Environmental Defense Fund Europe, University of Oxford Department of Computer Science and WWF, have developed the world's first Carbon Intensity forecast with a regional breakdown.

The Carbon Intensity API uses state-of-the-art Machine Learning and sophisticated power system modelling to forecast the carbon intensity and generation mix 96+ hours ahead for each region in Great Britain.

Our OpenAPI allows consumers and smart devices to schedule and minimise CO_2 emissions at a local level.

2-Day Carbon Intensity Forecast

+ Today	171	↑ 200	↓ 135
+ Tue	171	↑ 201	↓ 108
+ Wed	100	↑ 111	↓ 92

Values are the average, max, and min Carbon Intensity in gCO_2/kWh for each day

Click on interconnectors to see their names

Key: Very High - High - Moderate - Low - Very Low

BUT…

It doesn’t mean we should stop doing science

Synergies exist

GREENER PRINCIPLES FOR SUSTAINABLE COMPUTATIONAL SCIENCE

nature computational science

Perspective | Published: 26 June 2023

GREENER principles for environmentally sustainable computational science

Loïc Lannelongue, Hans-Erik G. Aronson, Alex Bateman, Ewan Birney, Talia Caplan, Martin Juckes, Johanna McEntyre, Andrew D. Morris, Gerry Reilly & Michael Inouye

Nature Computational Science **3**, 514–521 (2023) | Cite this article

Abstract

The carbon footprint of scientific computing is substantial, but environmentally sustainable computational science (ESCS) is a nascent field with many opportunities to thrive. To realize the immense green opportunities and continued, yet sustainable, growth of computer science, we must take a coordinated approach to our current challenges, including greater awareness and transparency, improved estimation and wider reporting of environmental impacts. Here, we present a snapshot of where ESCS stands today and introduce the GREENER set of principles, as well as guidance for best practices moving forward.

Collaboration with UK stakeholders

UNIVERSITY OF CAMBRIDGE
Department of Public Health and Primary Care

UK Research and Innovation

HDRUK
Health Data Research UK

EMBL-EBI

DARE UK

wellcome

GREENER PRINCIPLES FOR SUSTAINABLE COMPUTATIONAL SCIENCE

Governance — All actors in computational research have a key role to play and can lead the efforts towards sustainable computing.

Responsibility — Embracing both individual and institutional responsibility regarding the environmental impacts of research. This involves being transparent about these and initiating bold initiatives to reduce them.

Estimation — Monitoring environmental impacts to identify inefficiencies and opportunities for improvement.

Energy and embodied impacts — Minimising energy needs of computations and favouring low-carbon energy sources, while also considering the broader environmental impacts (e.g. water usage, mining of raw materials etc.).

New collaborations — Cooperating to leverage low-carbon infrastructures, facilitate equitable access to low-carbon computation and limit wasted resources.

Education — Training all stakeholders to be aware of the sustainability challenges of HPC and to be equipped with the skills to tackle them.

Research — Dedicate research efforts to green computing to improve our understanding of power usage, support sustainable software engineering and enable energy-efficient research.

Cultural change: make environmental sustainability a core element of research

FROM THEORY TO PRACTICE

***Estimating** and **reporting** the carbon footprint of algorithms*

ESTIMATING CARBON FOOTPRINTS **IN PRACTICE**

nature reviews methods primers

Explore content ⌄ **About the journal ⌄** **Publish with us ⌄**

nature › nature reviews methods primers › comment › article

Comment | Published: 16 February 2023

Carbon footprint estimation for computational research

Loïc Lannelongue & Michael Inouye

Nature Reviews Methods Primers **3**, Article number: 9 (2023) | Cite this article

Data analysis relies heavily on computation, and algorithms have grown more demanding in terms of hardware and energy. Monitoring their environmental impacts is and will continue to be an essential part of sustainable research. Here, we provide guidance on how to do so accurately and with limited overheads.

Michael Inouye

EXISTING TOOLS

carbontracker

pypi v1.1.6 | python 3.8 | 3.9 | 3.10 | build passing | License MIT

About

carbontracker is a tool for tracking and predicting the energy consumption and carbon footprint of training deep learning models as described in [Anthony et al. (2020)].

experiment-impact-tracker

The experiment-impact-tracker is meant to be a simple drop-in method to track energy usage, carbon emissions, and compute utilization of your system. Currently, on Linux systems with Intel chips (that support the RAPL or powergadget interfaces) and NVIDIA GPUs, we record: power draw from CPU and GPU, hardware information, python package versions, estimated carbon emissions information, etc. In California we even support realtime carbon emission information by querying caiso.com!

Once all this information is logged, you can generate an online appendix which shows off this information like seen here:

https://breakend.github.io/RL-Energy-Leaderboard/reinforcement_learning_energy_leaderboard/pongnoframeskip-v4_experiments/ppo2_stable_baselines,_default_settings/0.html

CODE CARBON

What it is

A lightweight and easy-to-use Python pip package

Emissions tracked based on your power consumption & location-dependent carbon intensity

Cloud Carbon Footprint

Cloud Carbon Emissions Measurement and Analysis Tool

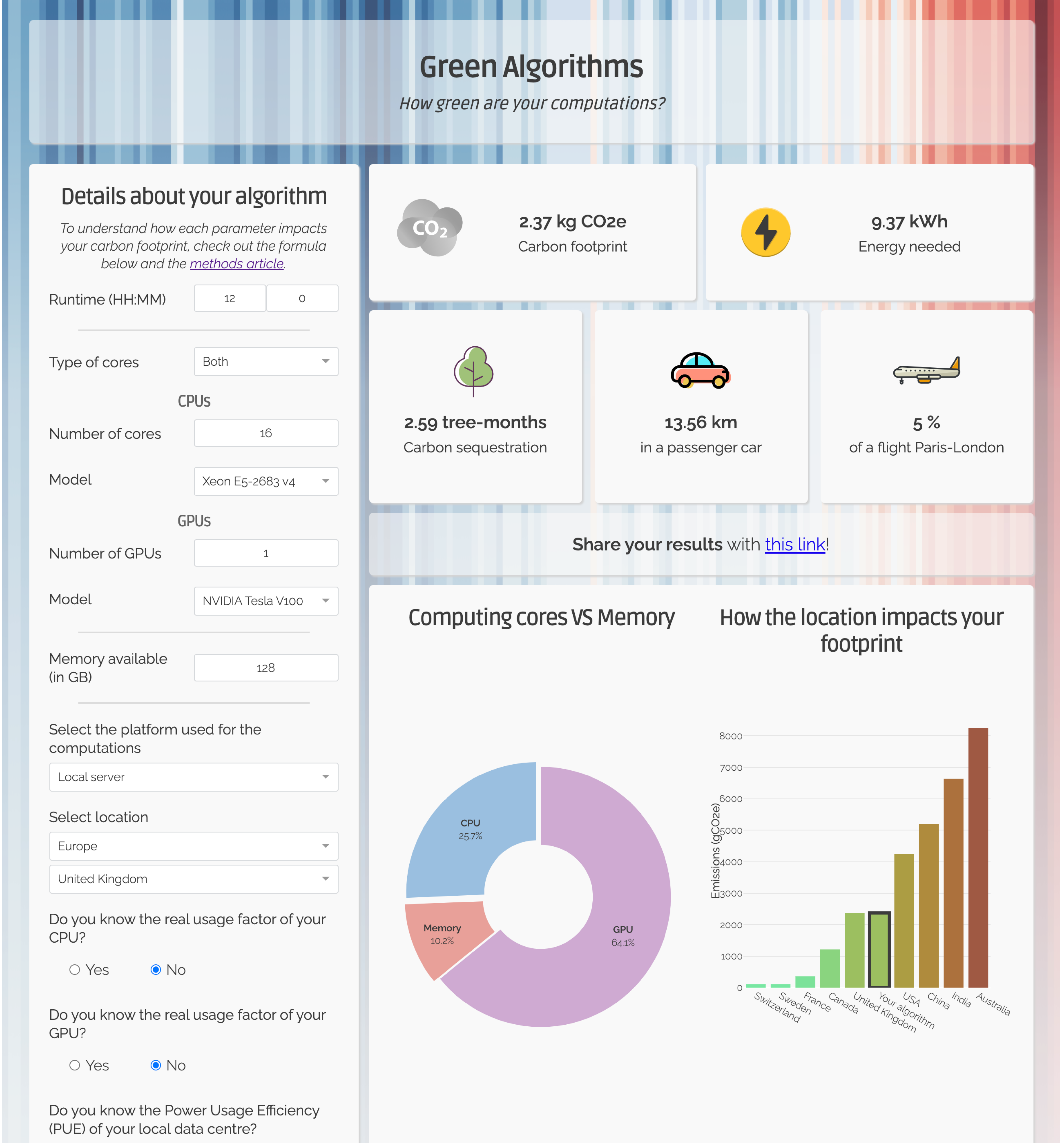
Understand how your cloud usage impacts our environment and what you can do about it

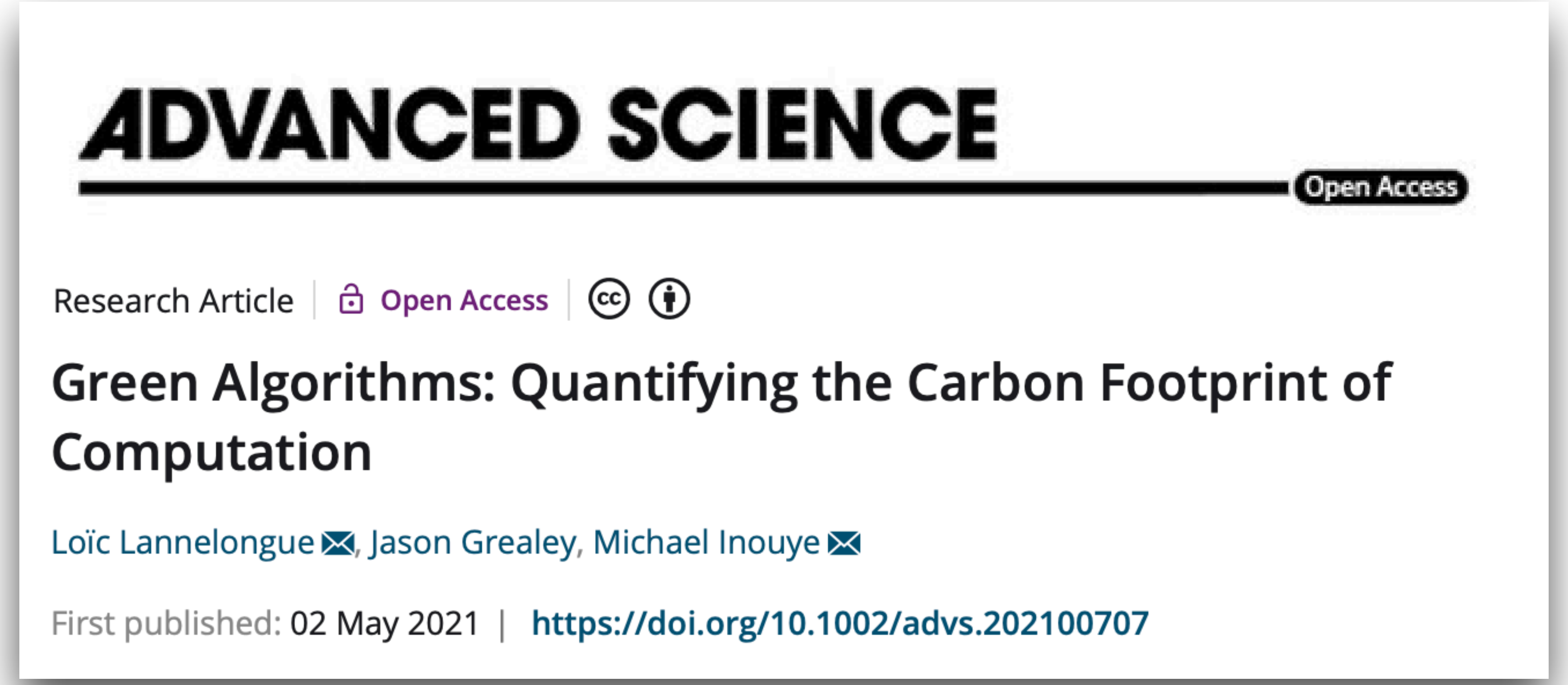
TRACARBON

CUMULATOR — a tool to quantify and report the carbon footprint of machine learning computations and communication in academia and healthcare

Trébaol, Tristan

2020





ADVANCED SCIENCE

Open Access

Research Article | Open Access

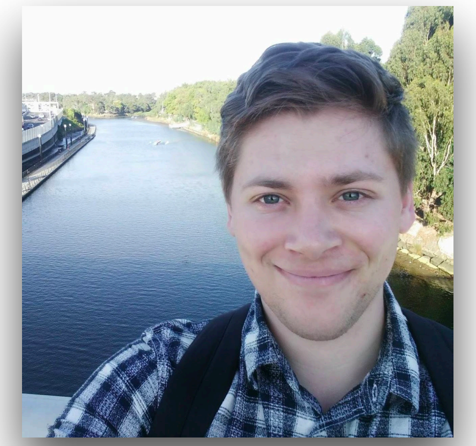
Green Algorithms: Quantifying the Carbon Footprint of Computation

Loïc Lannelongue, Jason Grealey, Michael Inouye

First published: 02 May 2021 | https://doi.org/10.1002/advs.202100707

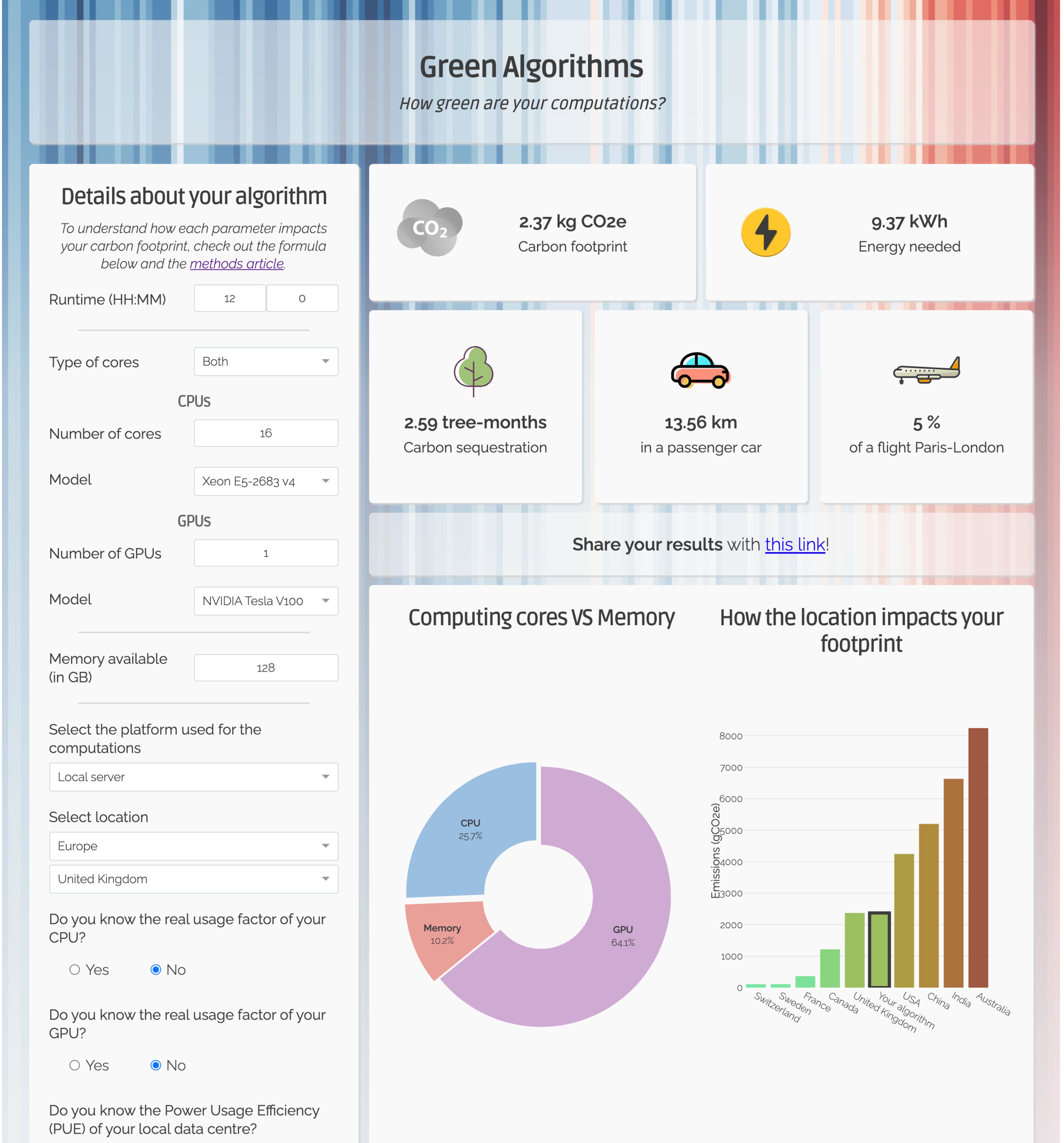
THE GREEN ALGORITHMS CALCULATOR

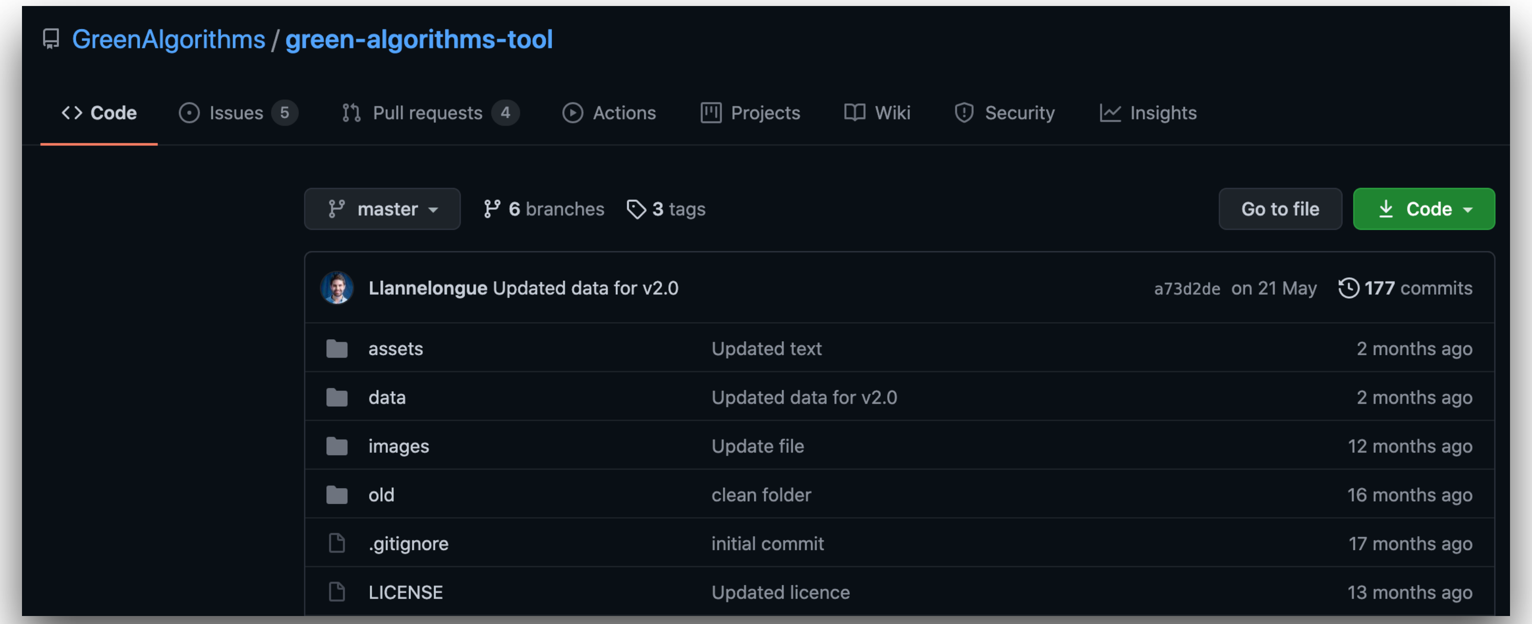
calculator.green-algorithms.org



Jason Grealey

Michael Inouye

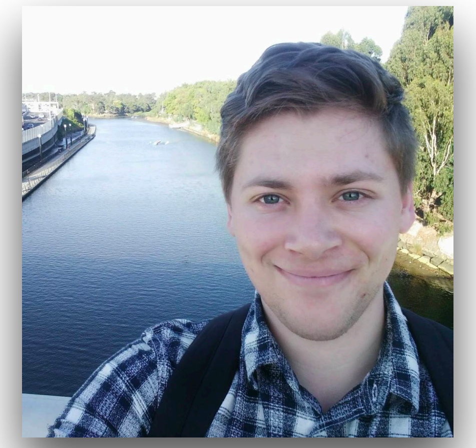




https://github.com/GreenAlgorithms/green-algorithms-tool

THE GREEN ALGORITHMS CALCULATOR

calculator.green-algorithms.org



Jason Grealey

Michael Inouye

GREEN ALGORITHMS 4 HPC

```
GreenAlgorithms4HPC]$ myCarbonFootprint.sh --STARTDAY 2020-01-01 --ENDDAY 2020-06-01
```

GREEN ALGORITHMS 4 HPC

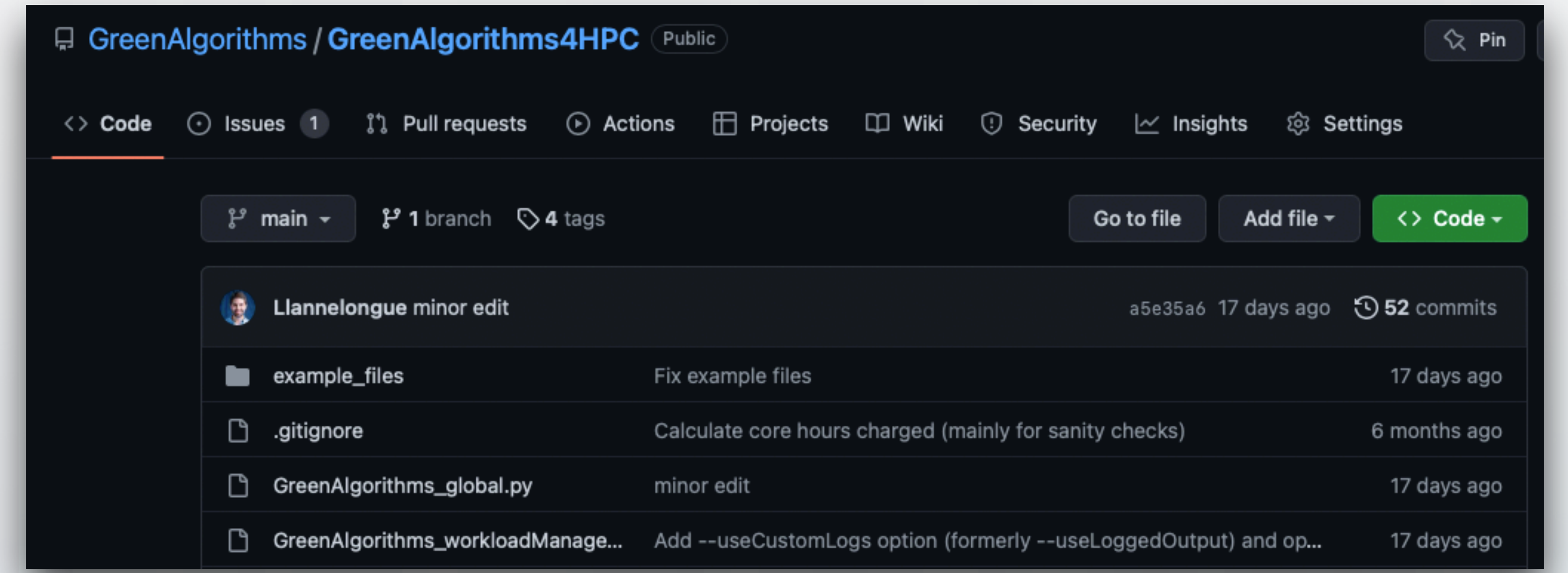
```
GreenAlgorithms4HPC]$ myCarbonFootprint.sh --STARTDAY 2020-01-01 --ENDDAY 2020-06-01
```

```
    ###################################
    #                                 #
    #  Your carbon footprint on CSD3  #
    #    (2021-01-01 / 2021-12-31)    #
    #                                 #
    ###################################

            ----------------
           |   222 kgCO2e   |
            ----------------

...This is equivalent to:
     - 20 tree-years
     - driving 1,268 km
     - 4.44 flights between Paris and London

...26.0% of your jobs failed, which represents a waste of 51 kgCO2e (55.26 tree-months).
...On average, you request at least 1.0 times the memory you need. By only requesting the memory you needed, you could have saved 0 gCO2e (0.00 tree-months).

Energy used: 960.17 kWh
     - CPUs: 88.91 kWh (9%)
     - GPUs: 713.81 kWh (74%)
     - Memory: 32.22 kWh (3%)
     - Data centre overheads: 125.24 kWh (13%)
Carbon intensity used for the calculations: 231.12 gCO2e/kWh

Summary of your usage:
     - First/last job recorded on that period: 2021-01-01/2021-12-08
     - Number of jobs: 1,490 (1,102 completed)
     - Core hours used/charged: 1,302.1 (CPU), 2,852.0 (GPU), 4,154.1 (total).
     - Total usage time (i.e. when cores were performing computations):
        - CPU: 430 days 03:58:39
        - GPU: 118 days 20:01:30
     - Total wallclock time: 132 days 10:49:44
     - Total memory requested: 40,981 GB

Limitations to keep in mind:
     - The workload manager doesn't alway log the exact CPU usage time, and when this information is missing, we assume that all cores are used at 100%.
     - For now, we assume that GPU jobs only use 1 GPU and the GPU is used at 100% (as the information needed for more accurate measurement is not available)
     (both of these may lead to slightly overestimated carbon footprints, although the order of magnitude is likely to be correct)
     - Conversely, the wasted energy due to memory overallocation may be largely underestimated, as the information needed is not always logged.

Any bugs, questions, suggestions? Email LL582@medschl.cam.ac.uk
--------------------------------------------------------------------------------
Calculated using the Green Algorithms framework: www.green-algorithms.org
Please cite https://onlinelibrary.wiley.com/doi/10.1002/advs.202100707
```

GREEN ALGORITHMS 4 HPC

```
GreenAlgorithms4HPC]$ myCarbonFootprint.sh --STARTDAY 2020-01-01 --ENDDAY 2020-06-01
```

https://github.com/GreenAlgorithms/GreenAlgorithms4HPC



Let us know if you try it!

```
####################################
#                                  #
#  Your carbon footprint on CSD3   #
#    (2021-01-01 / 2021-12-31)     #
#                                  #
####################################

                 ----------------
                |   222 kgCO2e   |
                 ----------------

...This is equivalent to:
     - 20 tree-years
     - driving 1,268 km
     - 4.44 flights between Paris and London

...26.0% of your jobs failed, which represents a waste of 51 kgCO2e (55.26 tree-months).
...On average, you request at least 1.0 times the memory you need. By only requesting the memory you r

Energy used: 960.17 kWh
     - CPUs: 88.91 kWh (9%)
     - GPUs: 713.81 kWh (74%)
     - Memory: 32.22 kWh (3%)
     - Data centre overheads: 125.24 kWh (13%)
Carbon intensity used for the calculations: 231.12 gCO2e/kWh

Summary of your usage:
     - First/last job recorded on that period: 2021-01-01/2021-12-08
     - Number of jobs: 1,490 (1,102 completed)
     - Core hours used/charged: 1,302.1 (CPU), 2,852.0 (GPU), 4,154.1 (total).
     - Total usage time (i.e. when cores were performing computations):
        - CPU: 430 days 03:58:39
        - GPU: 118 days 20:01:30
     - Total wallclock time: 132 days 10:49:44
     - Total memory requested: 40,981 GB
```

IT ENABLES **DEEP DIVES INTO PARTICULAR FIELDS**

Molecular Biology and Evolution
smbe
Society for Molecular Biology & Evolution

Article Navigation

The Carbon Footprint of Bioinformatics

Jason Grealey, Loïc Lannelongue, Woei-Yuh Saw, Jonathan Marten, Guillaume Méric, Sergio Ruiz-Carmona, Michael Inouye   Author Notes

Molecular Biology and Evolution, Volume 39, Issue 3, March 2022, msac034, https://doi.org/10.1093/molbev/msac034
Published: 10 February 2022

Molecular Biology and Evolution
smbe
Society for Molecular Biology & Evolution

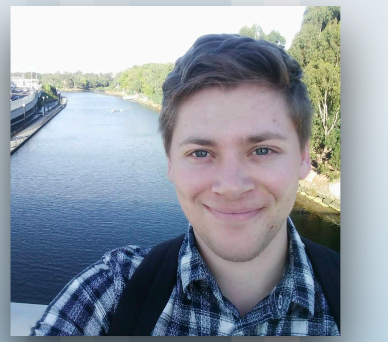
Issues | More content ▼ | Submit ▼ | Alerts | About ▼

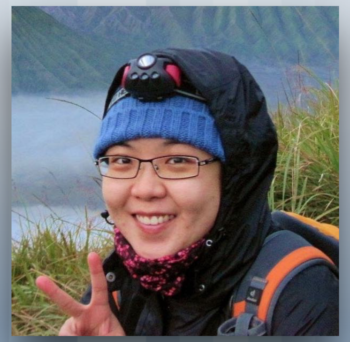
Article Navigation

Embracing Green Computing in Molecular Phylogenetics

Sudhir Kumar

Molecular Biology and Evolution, Volume 39, Issue 3, March 2022, msac043, https://doi.org/10.1093/molbev/msac043
Published: 04 March 2022


Jason Grealey


Woei Yuh Saw


Jonathan Marten


Sergio Ruiz-Carmona


Guillaume Méric


Michael Inouye

IT ENABLES **ENVIRONMENTAL IMPACT STATEMENTS**

GOUVERNEMENT
Liberté
Égalité
Fraternité

Ministère de la Transition écologique et de la Cohésion des territoires
Ministère de la Transition énergétique

FRANCE NATION VERTE
Agir · Mobiliser · Accélérer

Actualités | Politiques publiques | Démarches | Ministères

Accueil → Rendez-vous → Appel à projets Démonstrateurs d'IA pour les transitions écologique et énergétique → Appel à projets Démonstrateurs d'intelligence artificielle

Du 13 juillet au 07 novembre 2022

Appel à projets Démonstrateurs d'IA pour les transitions écologique et énergétique

Elle appuie l'appel à projet « démonstrateurs d'IA frugale dans les territoires pour la transition écologique », doté de 40 millions d'euros sur cinq ans, dans le cadre de la seconde phase de la stratégie nationale pour l'intelligence artificielle (SNIA).

- Estimer la consommation énergétique des services ou des produits numériques développés dans le cadre du projet (algorithmes et composants) exprimée de façon crédible et mesurable. Le porteur de projet s'appuiera sur l'outil en ligne et gratuit ci-dessous. Son code est ouvert et sa méthodologie est considérée comme robuste vis-à-vis de la littérature existante[3] [4]. Il s'agit de ***Green Algorithms (GT)***, Lannelongue et al, https://www.green-algorithms.org/

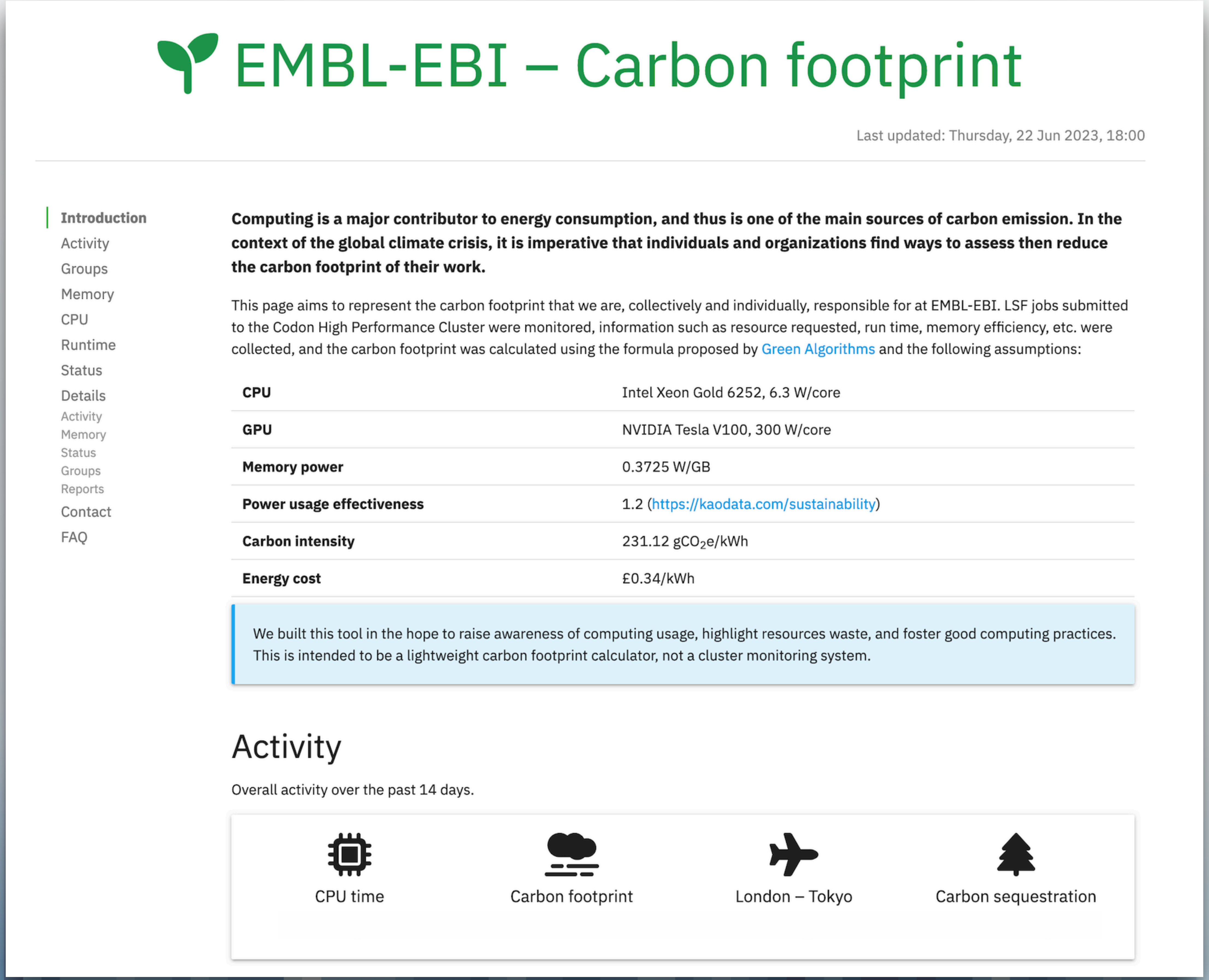
COMING NEXT: GA4HPC **DASHBOARD**

Institutional dashboard to **monitor computing carbon footprint** across research groups, units and departments

COMING NEXT: GA4HPC DASHBOARD

EMBL-EBI

Concept pioneered by EMBL-EBI (and others!)



EMBL-EBI – Carbon footprint

Last updated: Thursday, 22 Jun 2023, 18:00

- Introduction
- Activity
- Groups
- Memory
- CPU
- Runtime
- Status
- Details
  - Activity
  - Memory
  - Status
  - Groups
  - Reports
- Contact
- FAQ

Computing is a major contributor to energy consumption, and thus is one of the main sources of carbon emission. In the context of the global climate crisis, it is imperative that individuals and organizations find ways to assess then reduce the carbon footprint of their work.

This page aims to represent the carbon footprint that we are, collectively and individually, responsible for at EMBL-EBI. LSF jobs submitted to the Codon High Performance Cluster were monitored, information such as resource requested, run time, memory efficiency, etc. were collected, and the carbon footprint was calculated using the formula proposed by Green Algorithms and the following assumptions:

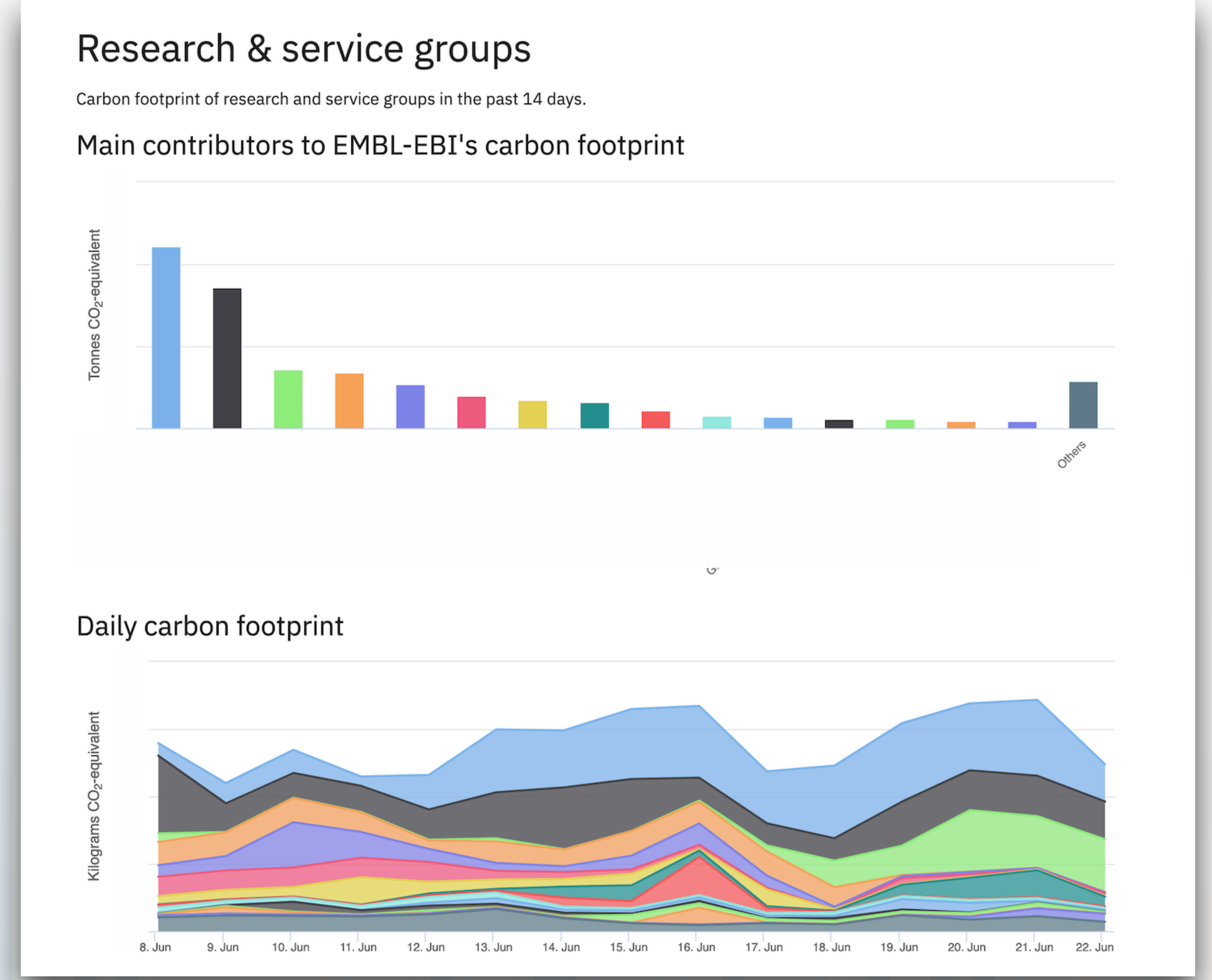
CPU	Intel Xeon Gold 6252, 6.3 W/core
GPU	NVIDIA Tesla V100, 300 W/core
Memory power	0.3725 W/GB
Power usage effectiveness	1.2 (https://kaodata.com/sustainability)
Carbon intensity	231.12 gCO_2e/kWh
Energy cost	£0.34/kWh

We built this tool in the hope to raise awareness of computing usage, highlight resources waste, and foster good computing practices. This is intended to be a lightweight carbon footprint calculator, not a cluster monitoring system.

Activity

Overall activity over the past 14 days.

CPU time | Carbon footprint | London – Tokyo | Carbon sequestration

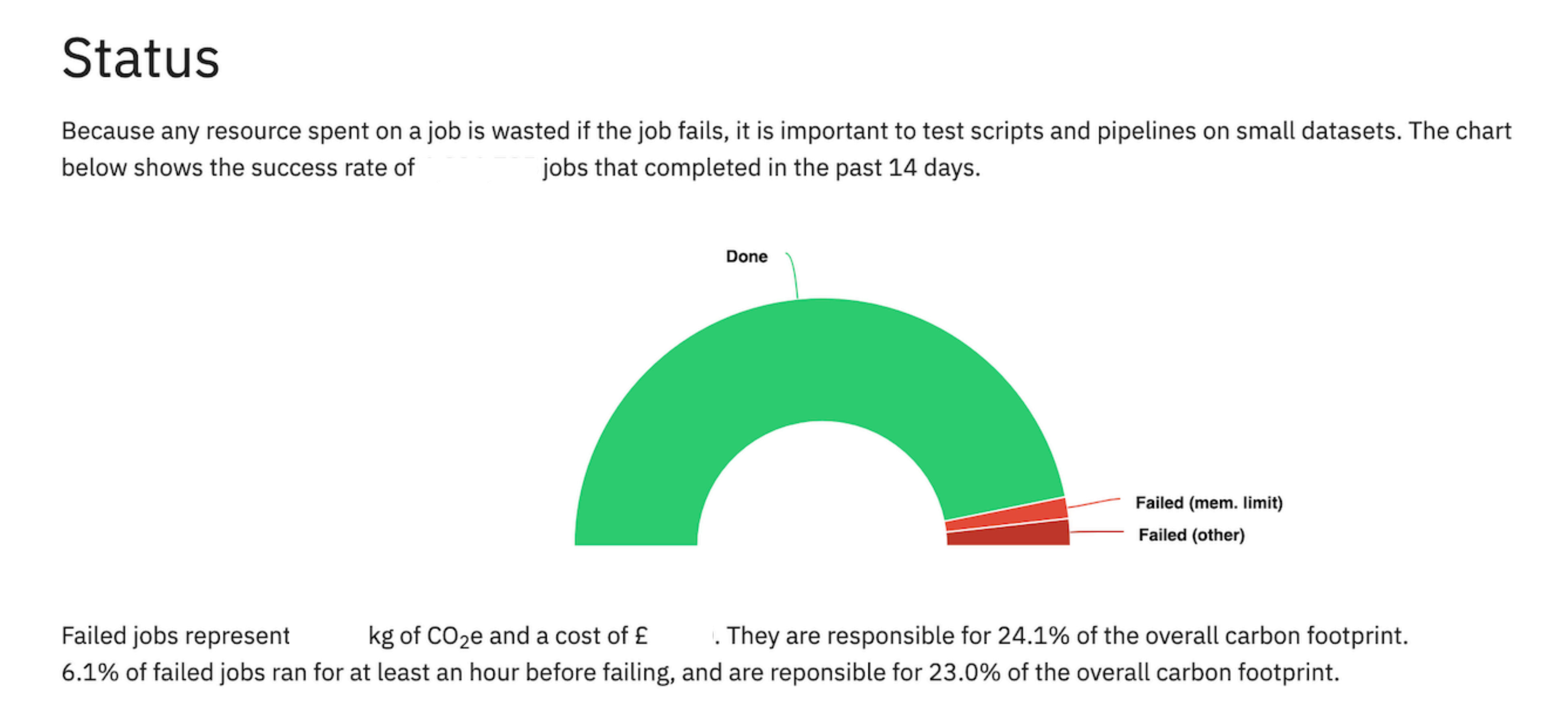


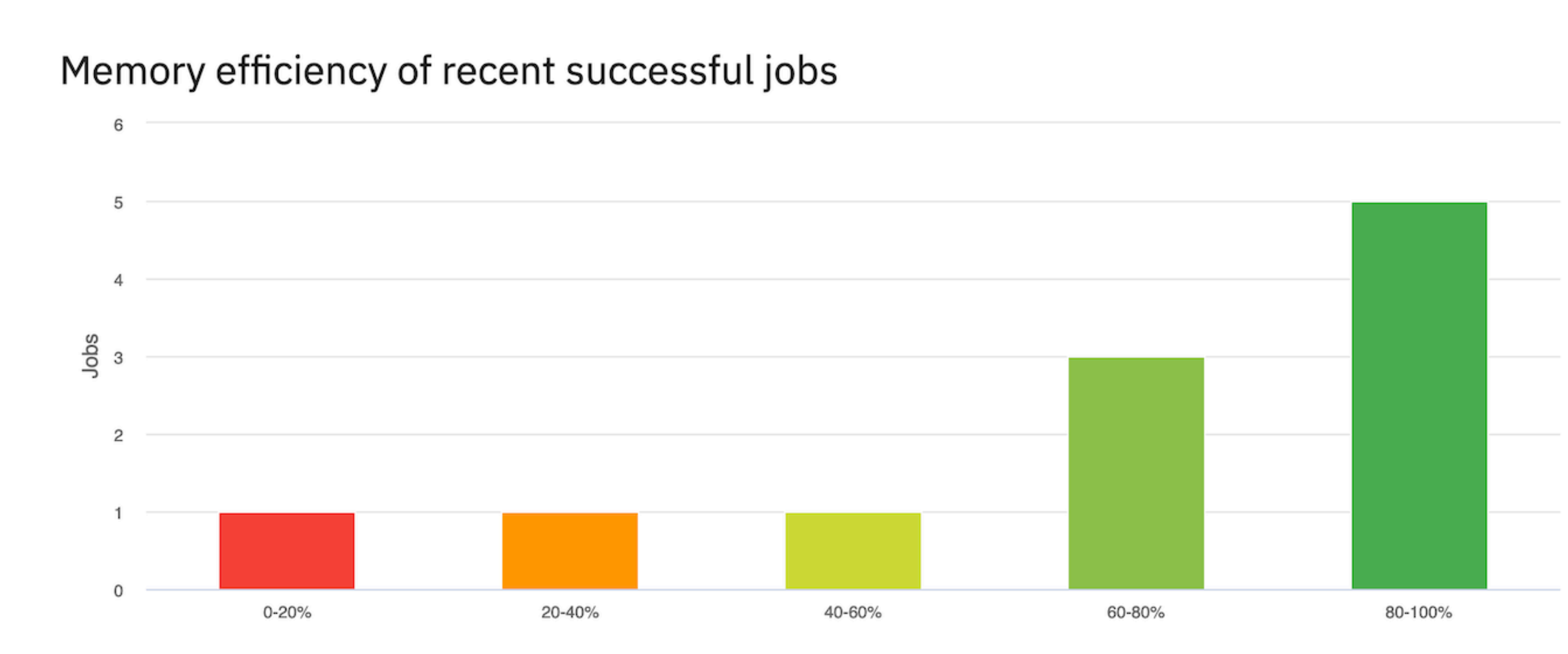
Matthias Blum

Alex Bateman

COMING NEXT: GA4HPC DASHBOARD

Concept pioneered by EMBL-EBI
(and others!)



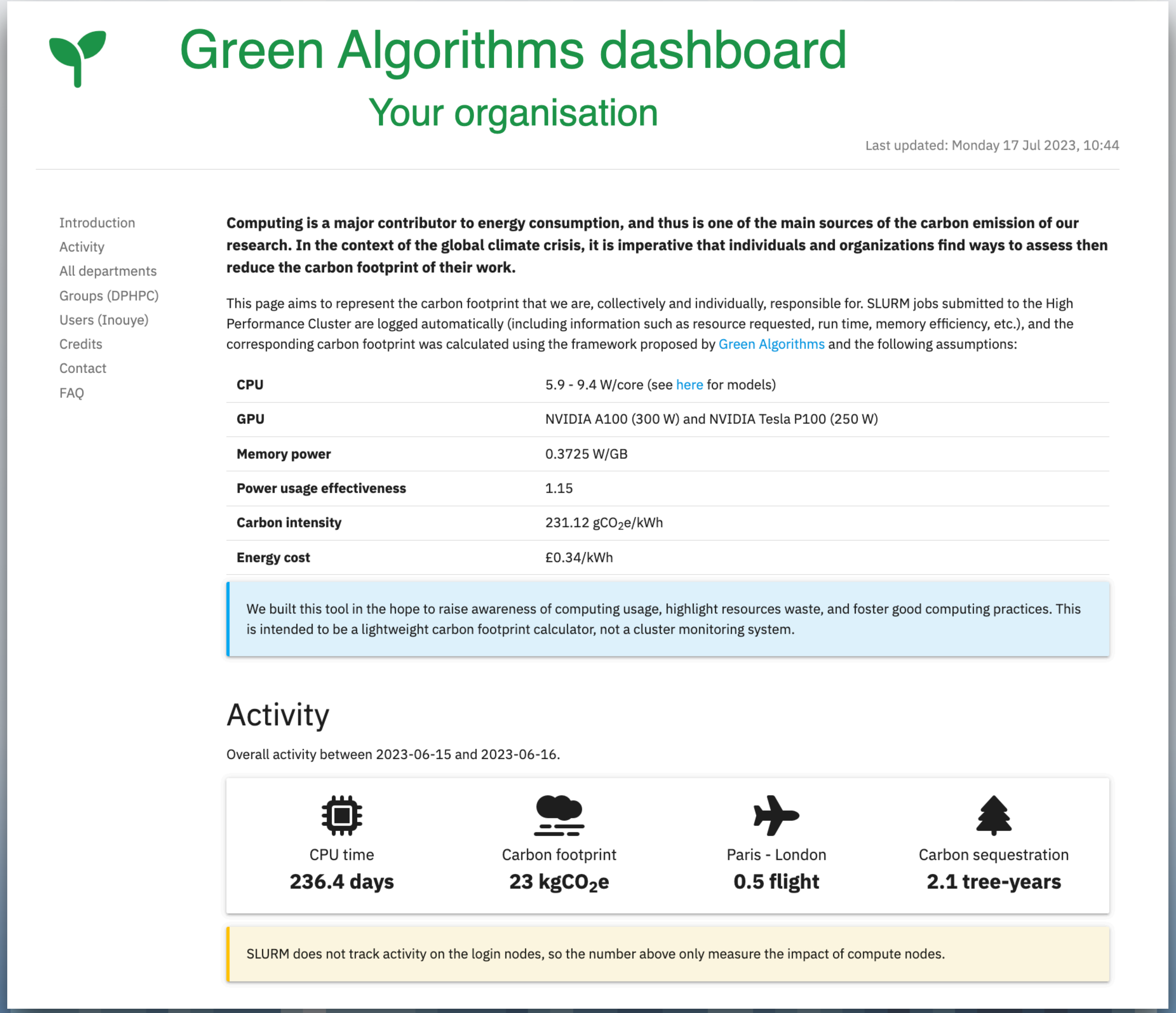


Matthias Blum

Alex Bateman

GREEN ALGORITHMS 4 HPC: **THE DASHBOARD**

Next step: an **open source, easy to deploy, reliable and transparent** SLURM-based dashboard implementing GA4HPC in computing facilities



Green Algorithms dashboard

Your organisation

Last updated: Monday 17 Jul 2023, 10:44

Introduction
Activity
All departments
Groups (DPHPC)
Users (Inouye)
Credits
Contact
FAQ

Computing is a major contributor to energy consumption, and thus is one of the main sources of the carbon emission of our research. In the context of the global climate crisis, it is imperative that individuals and organizations find ways to assess then reduce the carbon footprint of their work.

This page aims to represent the carbon footprint that we are, collectively and individually, responsible for. SLURM jobs submitted to the High Performance Cluster are logged automatically (including information such as resource requested, run time, memory efficiency, etc.), and the corresponding carbon footprint was calculated using the framework proposed by Green Algorithms and the following assumptions:

CPU	5.9 - 9.4 W/core (see here for models)
GPU	NVIDIA A100 (300 W) and NVIDIA Tesla P100 (250 W)
Memory power	0.3725 W/GB
Power usage effectiveness	1.15
Carbon intensity	231.12 gCO_2e/kWh
Energy cost	£0.34/kWh

We built this tool in the hope to raise awareness of computing usage, highlight resources waste, and foster good computing practices. This is intended to be a lightweight carbon footprint calculator, not a cluster monitoring system.

Activity

Overall activity between 2023-06-15 and 2023-06-16.

CPU time	Carbon footprint	Paris - London	Carbon sequestration
236.4 days	**23 $kgCO_2e$**	**0.5 flight**	**2.1 tree-years**

SLURM does not track activity on the login nodes, so the number above only measure the impact of compute nodes.

Interested in piloting it in your organisation? Let's chat!

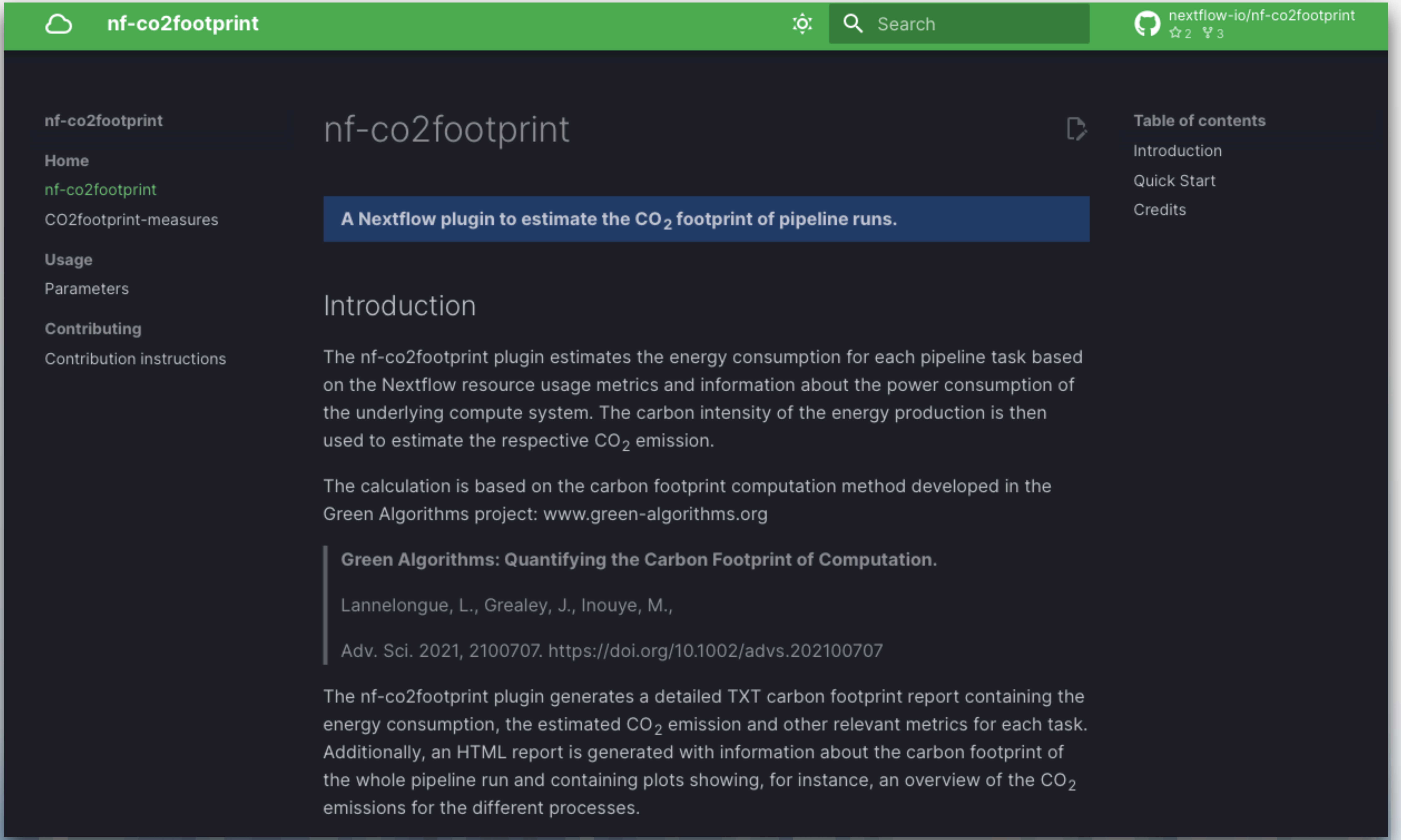
Already have such a dashboard? Let's chat!

Matthias Blum

Alex Bateman

Michael Inouye

COMING NEXT (PART 2): **NEXTFLOW PLUGIN**



Sabrina Krakau

Júlia Mir Pedrol

Phil Ewels

A DETAILED GUIDE FOR **DEEP LEARNING**

ENVIRONMENTAL RESEARCH
COMMUNICATIONS

ACCEPTED MANUSCRIPT • OPEN ACCESS

How to estimate carbon footprint when training deep learning models? A guide and review

Lucia Bouza Heguerte[1], Aurélie Bugeau[2] and Loïc Lannelongue[3]

Accepted Manuscript online 8 September 2023 • © 2023 The Author(s). Published by IOP Publishing Ltd

Lucia Souza

Aurélie Bureau

LABORATOIRE BORDELAIS DE RECHERCHE EN INFORMATIQUE LaBRI

TRANSPARENCY, FROM ALL OF US

Hardware manufacturers

Data centres

Institutions

Cloud providers

Scientists

TRANSPARENCY, FROM ALL OF US

Carbon impact and offsetting

We used GreenAlgorithms v.1.0 (ref. [121]) to estimate that the main computational work in this study had a carbon impact of at least 2,660 kg of CO_2 emissions (CO_2e), corresponding to 233 tree-years. As a commitment to the reduction of carbon emissions associated with computation in research, we consequently funded planting of 30 trees through a local Australian charity, which across their lifetime will sequester a combined estimated 8,040 kg of CO_2e, or three times the amount of CO_2e generated by this study.

Youwen Qin et al., Combined effects of host genetics and diet on human gut microbiota and incident disease in a single population cohort, Nature Genetics, 2022

Carbon impact and offsetting

We used GreenAlgorithms v.1.0 (ref. [84]) to estimate that the main computational work in this study had a carbon impact of at least 1,004 kg of CO_2 emissions (CO_2e), corresponding to 94 tree-years. As a commitment to the reduction of carbon emissions associated with computation in research, we consequently funded the planting of 45 trees through a local Australian charity, which across their lifetime will sequester a combined estimated 12,000 kg of CO_2e, or 12 times the amount of CO_2e generated by this study.

Yu Xu et al., An atlas of genetic scores to predict multi-omic traits, Nature, 2023

Carbon footprint of this project

We did our best to minimise greenhouse gas emissions related to this project and, using the Green Algorithms calculator (v2.1) [35], we estimated that the carbon footprint of this project was 51 kgCO_2e, which corresponds to 4.7 tree-years.

Lannelongue & Inouye, Inference mechanisms and prediction of protein-protein interactions, bioRxiv, 2022

TRANSPARENCY, FROM ALL OF US

Research | Open Access | Published: 19 August 2022

A comprehensive evaluation of microbial differential abundance analysis methods: current status and potential solutions

Lu Yang & Jun Chen

Microbiome **10**, Article number: 130 (2022) | Cite this article

others (146.1s vs 1.2–57.8 s). For large sample sizes, ZicoSeq can complete the analysis at an average of 5 and 25 min for n = 1000 and 5000, respectively (Fig. S22). Based on the Green Algorithms (green-algorithms.org v2.1 [62]) and the geographic location of Minnesota, USA, ZicoSeq has a carbon footprint of 0.06 g CO2e, 0.59 g CO2e, and 3.16 g CO2e for n = 100, 1000, and 5000, respectively.

tel-03726667, version 1 | Theses

en fr

The vehicle routing problem for flash floods relief operations

Florent Dubois [1, 2] Details
1 IRIT-SEPIA - Système d'exploitation, systèmes répartis, de l'intergiciel à l'architecture
IRIT - Institut de recherche en informatique de Toulouse
2 UT3 - Université Toulouse III - Paul Sabatier

To conclude on the ethical aspect of my research, the carbon footprint of my thesis has been evaluated. It has been calculated using green-algorithms.org v2.0 [59]. The calculation considers the large-scale experimentation conducted on the platform Grid5000 located in France. Experimentation took a total of 576 hours of computations on 16 CPUs Xeon E5-2660 v3 and has drawn 207.47 kWh. It represents a carbon footprint of $8.08 kg\ CO_{2e}$.

Equivariant and Modular DeepSets with Applications in Cluster Cosmology

Leander Thiele*
Department of Physics
Princeton University
Princeton, NJ 08544

Miles Cranmer
Department of Astrophysical Sciences
Princeton University
Princeton, NJ 08544

William Coulton, Shirley Ho, David N. Spergel
Center for Computational Astrophysics

[6] Total compute cost is 13.4 (Tesla P100+9CPU) khr (1.09t CO_2e [26]) with a PyTorch [27] implementation.

CSH PRESS GENOME RESEARCH

HOME | ABOUT | ARCHIVE | SUBMIT | SUBSCRIBE | ADVERTISE | AUTHOR INFO | CONTA

Efficient computation of Faith's phylogenetic diversity with applications in characterizing microbiomes

George Armstrong[1,2,3], Kalen Cantrell[2], Shi Huang[1,2], Daniel McDonald[1],
Niina Haiminen[4], Anna
Imran McGrath[2,8], Kris
Guillaume Méric[12,13],
Mohit Jain[2,17,18], Mic
Laxmi Parida[4], Yoshiki

0.5 GB of memory. Additionally, using Green Algorithms (Lannelongue et al. 2021), we estimated the carbon footprint of the scikit-bio reference implementation on the 20,000 sample table to be 12.84 g CO_2e, whereas we estimated the carbon footprint of SFPhD would be 0.04 g CO_2e in the United States, which is a 321-fold reduction in impact on global warming.

FROM ACKNOWLEDGING TO REDUCING IMPACTS

Tackling Energy and embodied impact through New Collaborations

WHAT CAN WE DO NOW?

PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS

EDITORIAL

Ten simple rules to make your computing more environmentally sustainable

Loïc Lannelongue, Jason Grealey, Alex Bateman, Michael Inouye

Published: September 20, 2021 • https://doi.org/10.1371/journal.pcbi.1009324

Keep, Repair, Reuse

Promote efficient data centres

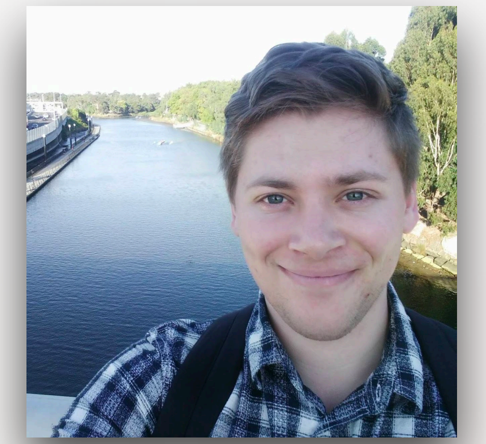
Estimate and report your own footprint for your projects

…and include it in your cost-benefit analysis

Sustainability should be accounted for in renewing policies

Carefully choose your computing facility

Optimise (or use optimised) code



Jason Grealey

Alex Bateman

Michael Inouye

A SUSTAINABILITY STANDARD **FOR DRY LAB**

Pilot phase underway

Coming soon

Want to be kept updated? Email me! LL582@medschl.cam.ac.uk

FIELD-SPECIFIC GUIDANCE: **NEUROSCIENCE**

Ten recommendations for reducing the carbon footprint of research computing in human neuroimaging

AUTHORS
Nicholas Edward Souter, Loïc Lannelongue, Gabby Samuel, Chris Racey, Lincoln Colling, Nikhil Bhagwat, Raghavendra Selvan, Charlotte Rae

Nick Souter

Charlotte Rae

Suggested Action: *Regularly remove files that you do not need.*

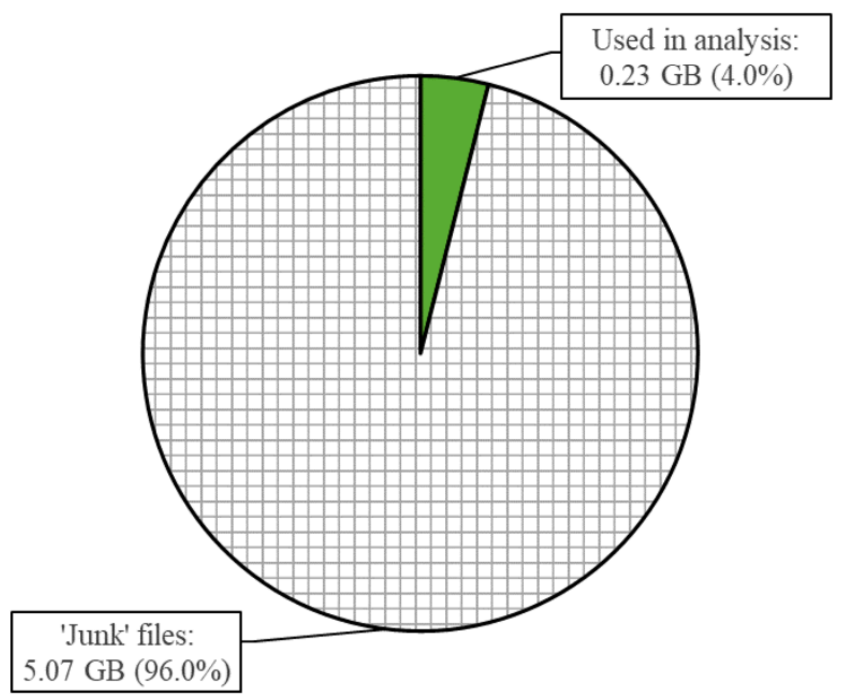


Figure 3. The mean percentage of total data generated by fMRIPrep that is actively used in data analysis (solid green) versus files that can be safely deleted after the completion of

Box 1. Summary of ten recommendations for reducing the carbon footprint of neuroimaging computing

1. *Make use of existing preprocessed data when possible, instead of acquiring and processing new data*
2. *Preregister a study analysis plan in order to avoid repetitions*
3. *Quantify and report the carbon footprint of your computing using available carbon tracking tools*
4. *Only run the preprocessing and analysis steps that you need*
5. *Run your computing at lower carbon intensity times and in lower carbon intensity locations*
6. *Regularly remove files that you do not need*
7. *Plan where, and for how long, you will store files, aided by research technicians*
8. *Advocate for non-commercial and centralised data storage solutions*
9. *Publicly share sufficient data to ensure it is FAIR (Findable, Accessible, Interoperable, Reusable), but consider the extent of what others will actually need or use*
10. *Discuss the importance of greener computing with other neuroimagers and advocate for systemic change*

MINIMISING CARBON INTENSITY THROUGH SMART SCHEDULING

CATS

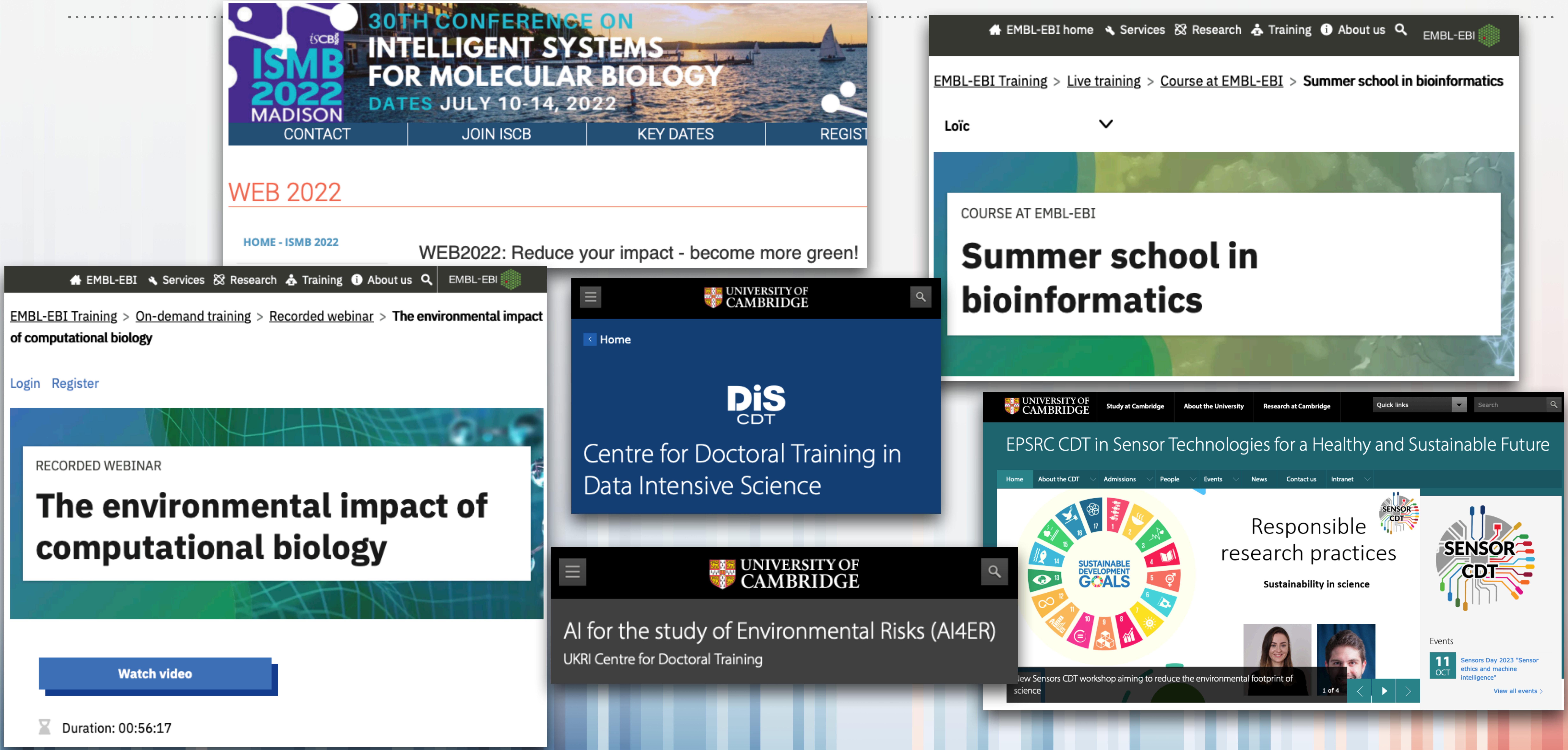
Climate-**A**ware **T**ask **S**cheduler

CATS is a Climate-Aware Task Scheduler. It schedules cluster jobs to minimize predicted carbon intensity of running the process. It was created as part of the 2023 Collaborations Workshop.

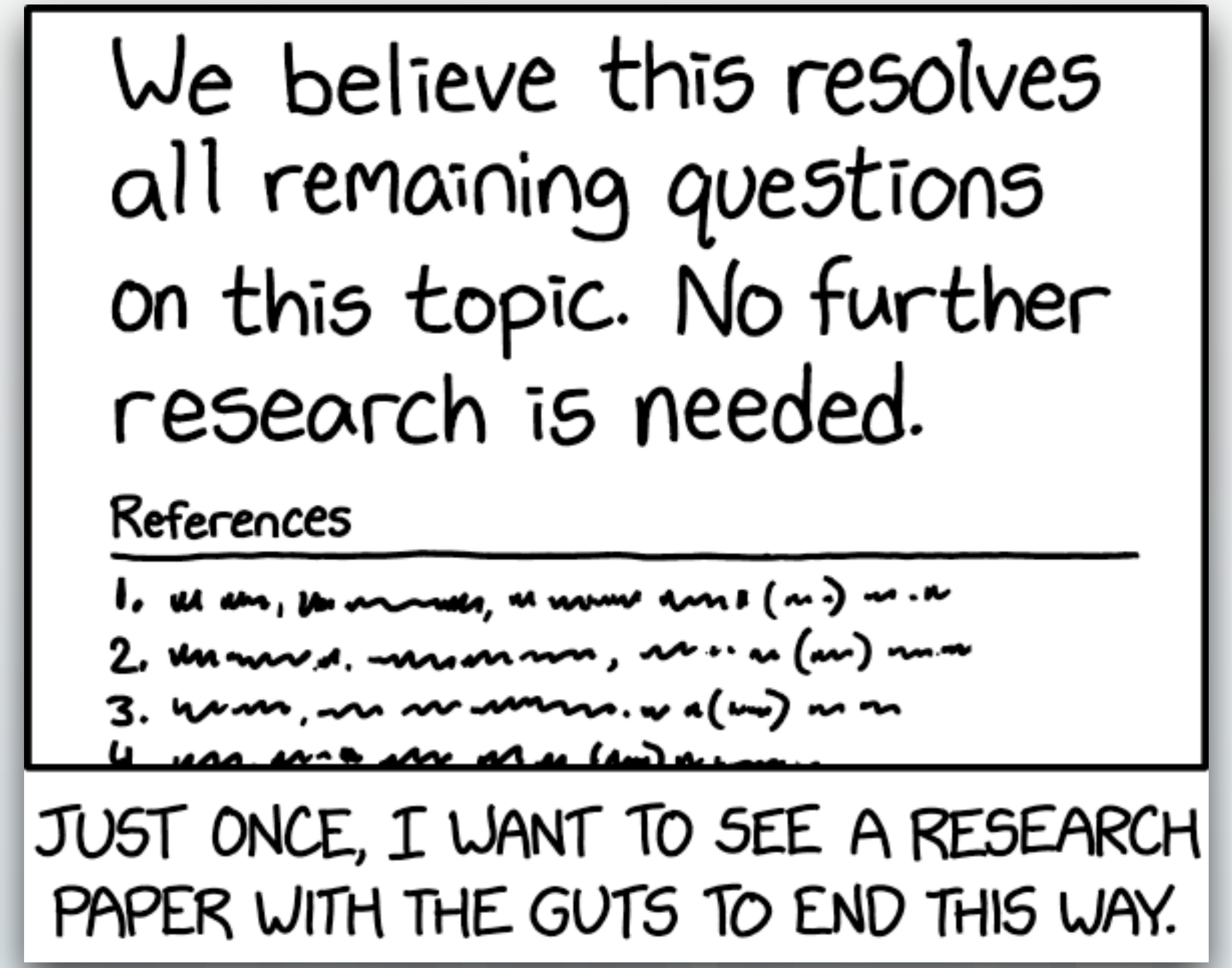
Currently CATS only works in the UK, if you are aware of APIs for realtime grid carbon intensity data in other countries please open an issue and let us know.

MOVING FORWARD: EDUCATION AND RESEARCH

MAKING SUSTAINABILITY PART OF SCIENTIFIC TRAINING



IDENTIFY FURTHER OPPORTUNITIES FOR MORE SUSTAINABLE COMPUTING



Sadly not yet

*So **dedicated** research efforts are needed*

xkcd.com

IDENTIFY FURTHER OPPORTUNITIES FOR MORE SUSTAINABLE COMPUTING

Table 3
Results for binary-trees, fannkuch-redux, and fasta.

binary-trees	Energy (J)	Time (ms)	Ratio (J/ms)	Mb
(c) C	39.80	1125	0.035	131
(c) C++	41.23	1129	0.037	132
(c) Rust $\Downarrow_2$	49.07	1263	0.039	180
(c) Fortran $\Uparrow_1$	69.82	2112	0.033	133
(c) Ada $\Downarrow_1$	95.02	2822	0.034	197
(c) Ocaml $\downarrow_1$ $\Uparrow_2$	100.74	3525	0.029	148
(v) Java $\uparrow_1$ $\Downarrow_{16}$	111.84	3306	0.034	1120
(v) Lisp $\downarrow_3$ $\Downarrow_3$	149.55	10570	0.014	373
(v) Racket $\downarrow_4$ $\Downarrow_6$	155.81	11261	0.014	467
(i) Hack $\uparrow_2$ $\Downarrow_9$	156.71	4497	0.035	502
(v) C# $\downarrow_1$ $\Downarrow_1$	189.74	10797	0.018	427
(v) F# $\downarrow_3$ $\Downarrow_1$	207.13	15637	0.013	432
(c) Pascal $\downarrow_3$ $\Uparrow_5$	214.64	16079	0.013	256
(c) Chapel $\uparrow_5$ $\Uparrow_4$	237.29	7265	0.033	335
(v) Erlang $\uparrow_5$ $\Uparrow_1$	266.14	7327	0.036	433
(c) Haskell $\uparrow_2$ $\Downarrow_2$	270.15	11582	0.023	494
(i) Dart $\downarrow_1$ $\Uparrow_1$	290.27	17197	0.017	475
(i) JavaScript $\downarrow_2$ $\Downarrow_4$	312.14	21349	0.015	916
(i) TypeScript $\downarrow_2$ $\Downarrow_2$	315.10	21686	0.015	915
(c) Go $\uparrow_3$ $\Uparrow_{13}$	636.71	16292	0.039	228
(i) Jruby $\uparrow_2$ $\Downarrow_3$	720.53	19276	0.037	1671
(i) Ruby $\Uparrow_5$	855.12	26634	0.032	482
(i) PHP $\Uparrow_3$	1,397.51	42316	0.033	786
(i) Python $\Uparrow_{15}$	1,793.46	45003	0.040	275
(i) Lua $\downarrow_1$	2,452.04	209217	0.012	1961
(i) Perl $\uparrow_1$	3,542.20	96097	0.037	2148
(c) Swift		n.e.		

Contents lists available at ScienceDirect
Science of Computer Programming
www.elsevier.com/locate/scico
ELSEVIER

Ranking programming languages by energy efficiency

Rui Pereira [a,b,*], Marco Couto [c,b], Francisco Ribeiro [c,b], Rui Rua [c,b], Jácome Cunha [c,b], João Paulo Fernandes [d], João Saraiva [c,b]

[a] C4 — Centro de Competências em Cloud Computing (C4-UBI), Universidade da Beira Interior, Rua Marquês d'Ávila e Bolama, 6201-001, Covilhã, Portugal
[b] HASLab/INESC Tec, Portugal
[c] Universidade do Minho, Portugal
[d] Departamento de Engenharia Informática, Faculdade de Engenharia da Universidade do Porto & CISUC, Portugal

We need more trained Research Softwares Engineers

ALL THIS LEADING TO CULTURAL CHANGE

JEVON'S PARADOX

Rebound effect can ruin all our efforts

BUILDING A COMMUNITY

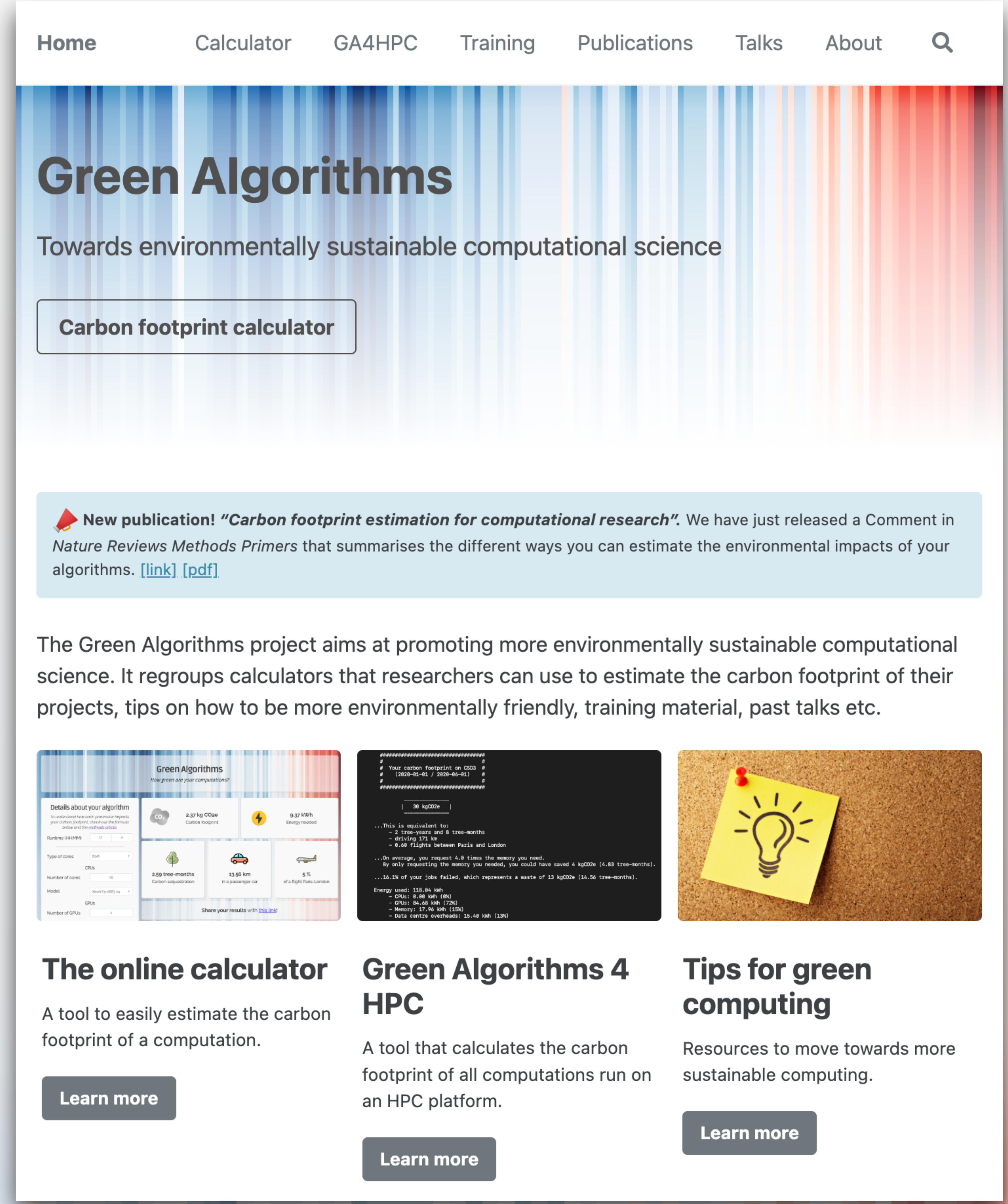
www.eicworkshop.info

An upcoming community of practice

https://forms.gle/rgeqzcpo51gge5Xr6

The Green Algorithms website with all resources

At www.green-algorithms.org



ACKNOWLEDGMENTS

UNIVERSITY OF CAMBRIDGE
Department of Public Health and Primary Care

Cambridge-Baker System Genomics Initiative

Michael Inouye
Jason Grealey
Woei-Yuh Saw
Jonathan Marten
Guillaume Meric
Sergio Ruiz-Carmona
Elodie Persyn
Yang Liu

EMBL-EBI

Alex Bateman
Ewan Birney
Johanna McIntyre
Anna Swan
Sarah Morgan
Matthias Blum

Uni of Sussex

Charlotte Rae
Nick Souther
Lincoln Colling

HDR-UK

Andrew Morris
Hans-Erik Aronson
Gerry Reilly

UKRI

Martin Juckes

LaBRI

Aurélie Bureau
Lucia Souza

KCL

Gabrielle Samuel

Wellcome

Talia Caplan

LEAF

Martin Farley

The CATS team

British Heart Foundation
NIHR | Cambridge Biomedical Research Centre
Baker HEART & DIABETES INSTITUTE
HDRUK Health Data Research UK
Software Sustainability Institute
UKRI Medical Research Council
Jesus College Cambridge

How to follow the project and reach out

www.green-algorithms.org

@Loic_Lnlg LL582@medschl.cam.ac.uk

An upcoming community of practice

https://forms.gle/rgeqzcpo51gge5Xr6