The environmental impacts of neuroimaging, from liquid helium to big data: what's our footprint?

Charlotte Rae

University of Sussex, UK

Slides: https://ohbm-environment.org/

The climate crisis & ecological emergency







The climate crisis & ecological emergency

- Happening now
- We must act urgently
- The causes extend across all domains of human activity:
  - scientific research is one of these
- If we do not act, we will not be able to carry on doing science
- Think about what a public health emergency (Covid) did to science…
- What will the even bigger environmental emergencies do?

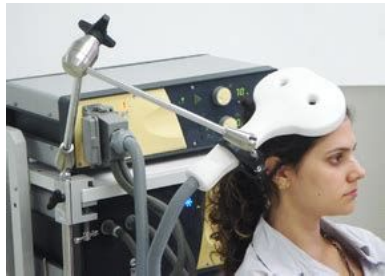
Neuroimaging: what's our footprint?

OSF

hardware

data

open science



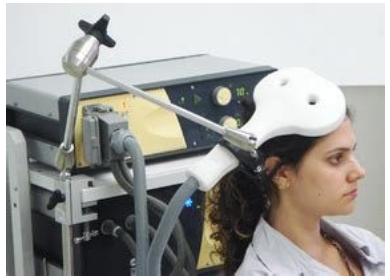
NEUROVAULT

OpenNEURO

Neuroimaging: what's our footprint?

- Raw materials
- Recycling
- Liquid helium

hardware



Liquid helium

NB. OPM-MEG: no helium required

Possible hardware solutions

Future non-helium technologies

- Optically-Pumped Magnetometer (OPM) MEG
- MRI cooled to less extreme temperature with liquid nitrogen

Value our MRI & MEG data more highly

- Especially data with high reuse potential

Install a helium recycling tank

- Captures boil-off: especially important for MEG

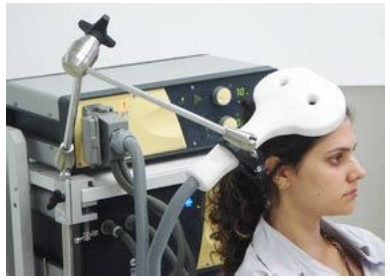
Neuroimaging: what's our footprint?

OSF

hardware

data

open science



NEUROVAULT

OpenNEURO

Data centres: construction

> The building "shell"

Quantified for a 1MW data center

Concrete	
Masonry	Brick, stone, grout
Metals	Steel beams, lead pipes, copper wires, aluminum sheet metal, stairs, railings, floor plates, grates, nails, screws, bolts, aluminum flashing, sheet metal, aluminum ventilation, louver systems
Wood, plastic, composite	Room framing, wire coatings, doors, windows
Thermal/moisture protection	Insulation, vapor barriers
Water	Cleaning, cooling, fire suppression
Chemicals	Glue, glycol, cleaners, water repellants sealants, fire suppression
Glass	
Tar	Roofs, roads, sidewalks, parking lots
Shingles, tiles	

> The data center

Quantified for a 1MW data center

Electricity	177,000,000 kW-hr
Water	60,000,000 gal (227,000,000 l)
Copper	145,000 lbs (65,771 kg)
Lead	21,000 lbs (9,525 kg)
Plastic	33,000 lbs (14,968 kg)
Aluminum	73,000 lbs (33,112 kg)
Solder	12,000 lbs (5,443 kg)
Steel	377,000 lbs (171,004 kg)

Assumptions: 10-year lifetime, high redundancy, two IT refreshes, includes power/cooling/racks/IT, does not include the building

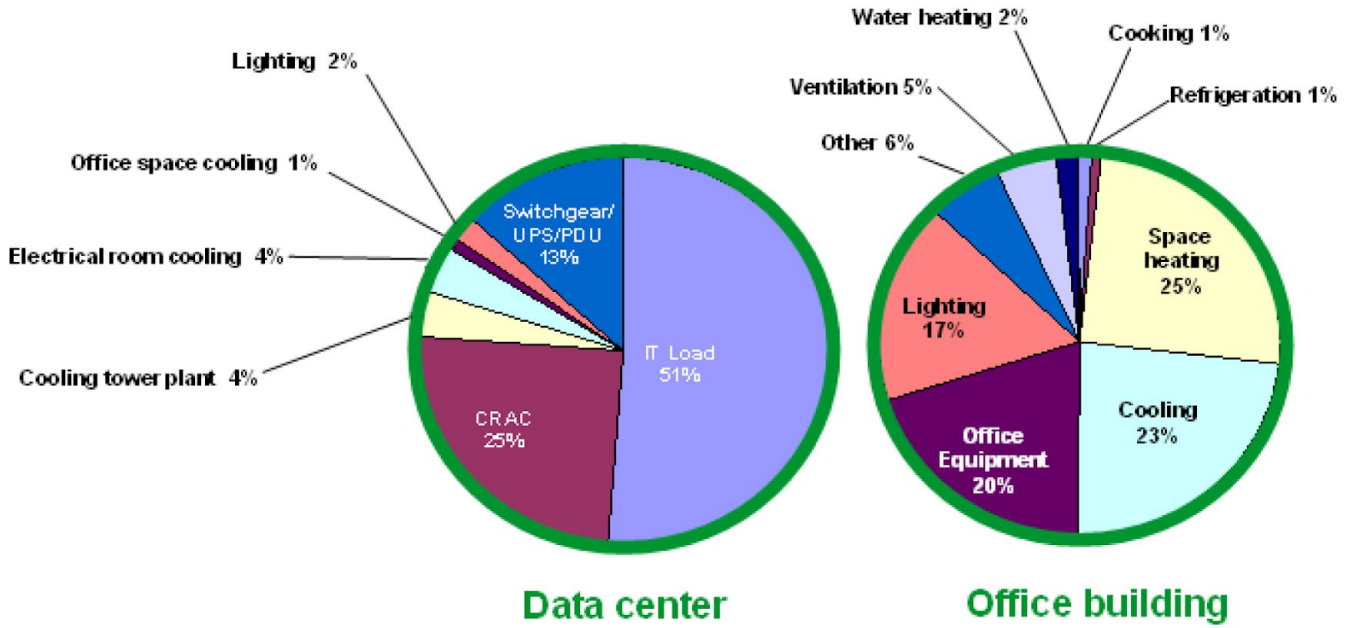
"10 year lifetime"

"2 IT refreshes"

https://www.insight.com/content/dam/insight/en_US/pdfs/apc/apc-estimating-data-centers-carbon-footprint.pdf

Data centres: energy

"Data centres can be **40 times** as energy intensive as an office building"



https://www.insight.com/content/dam/insight/en_US/pdfs/apc/apc-estimating-data-centers-carbon-footprint.pdf

Data centres: energy

3 key factors

- IT load
  - more activity, more energy required
- Location
  - extremes of temperature require more cooling
- Power Use Efficiency
  - centre & system design; local vs cloud

https://www.insight.com/content/dam/insight/en_US/pdfs/apc/apc-estimating-data-centers-carbon-footprint.pdf

Data centres: energy

…where is the energy coming from?





Data centres: energy

…where is the energy coming from?

- UK energy 2020 mix: 43% renewable*
- Renewables require manufacture
  - raw materials
  - lifespan, recycling
- Not enough renewable generation to meet our current energy demands
  - requires **reductions**

*UK Department for Business, Energy & Industrial Strategy

Open neuroimaging repositories


- Uses Google: 100% renewable


- Uses Amazon Web Services
- AWS ~50% renewable; non-renewable storage offset*
- Neurovault looked at switching to guaranteed renewable last year but not yet


- Uses AWS

*offsetting very problematic (Adam's talk)

Data analysis & sharing: suggestions

Consider carefully how much data to acquire, analyse, store, share

- The more we do, the bigger our environmental costs
- ‘Slow science’ (Frith, 2020, *TICS*)

Data analysis & sharing: suggestions

Consider carefully how much data to acquire, analyse, store, share

- The more we do, the bigger our environmental costs
- ‘Slow science’ (Frith, 2020, *TICS*)

‘Reduce, reuse, recycle’

- Reduce data acquisition where appropriate; reduce energy consumption in analysis

Data analysis & sharing: suggestions

Consider carefully how much data to acquire, analyse, store, share

- The more we do, the bigger our environmental costs
- ‘Slow science’ (Frith, 2020, *TICS*)

‘Reduce, reuse, recycle’

- Reduce data acquisition where appropriate; reduce energy consumption in analysis

Carbontracker: Tracking and Predicting the Carbon Footprint of Training Deep Learning Models

Lasse F. Wolff Anthony* [1] Benjamin Kanding* [1] Raghavendra Selvan [1]

https://arxiv.org/abs/2007.03051

Data analysis & sharing: suggestions

Consider carefully how much data to acquire, analyse, store, share

- The more we do, the bigger our environmental costs
- ‘Slow science’ (Frith, 2020, *TICS*)

‘Reduce, reuse, recycle’

- Reduce data acquisition where appropriate; reduce energy consumption in analysis

Data analysis & sharing: suggestions

Consider carefully how much data to acquire, analyse, store, share

- The more we do, the bigger our environmental costs
- 'Slow science' (Frith, 2020, *TICS*)

'Reduce, reuse, recycle'

- Reduce data acquisition where appropriate; reduce energy consumption in analysis
- Reuse existing data: within lab, or meta-analysis
- Recycle popular shared datasets (e.g. Human Connectome Project, Biobank)

! Ensure only 1 institutional download

! File management in general also crucial

Data analysis & sharing: suggestions

Computing resources

- Consider geographical location and time of day
- Cloud usually more efficient than local
- Historical data without ethics to share: using resources but not actively used

  ! Clean up files regularly

- Share preprocessed rather than raw data?

Data analysis & sharing: suggestions

Computing resources

- Consider geographical location and time of day
- Cloud usually more efficient than local
- Historical data without ethics to share: using resources but not actively used

! Clean up files regularly

Open neuroimaging

- OSF ‘green standard’, but not specialised for neuroimaging
- Is sharing summary on Neurovault sufficient?
- Share preprocessed rather than raw data?

Data analysis & sharing: suggestions

Computing resources

- Consider geographical location and time of day
- Cloud usually more efficient than local
- Historical data without ethics to share: using resources but not actively used

! Clean up files regularly

Open neuroimaging

- OSF ‘green standard’, but not specialised for neuroimaging
- Is sharing summary on Neurovault sufficient?
- Share preprocessed rather than raw data?

OSR Panel Discussion!

‘Open science & sustainability’

21st June 17:30 UTC

What's our footprint?

- In the new OHBM **Sustainability & Environment Action SIG**,
- We are quantifying the footprint of human neuroimaging pipelines,
- From data acquisition to analysis & sharing
- To produce tools, advice, and guidance on minimising footprints

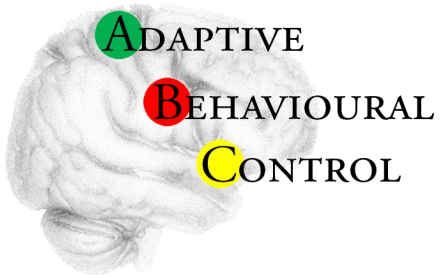
hardware

data

open science

…join us!

…team efforts



c.rae@sussex.ac.uk

@NeuroRae

https://www.sussex.ac.uk/psychology/abc-lab/climate-change

ohbm.sea.sig@gmail.com

@OhbmEnvironment

https://ohbm-environment.org/