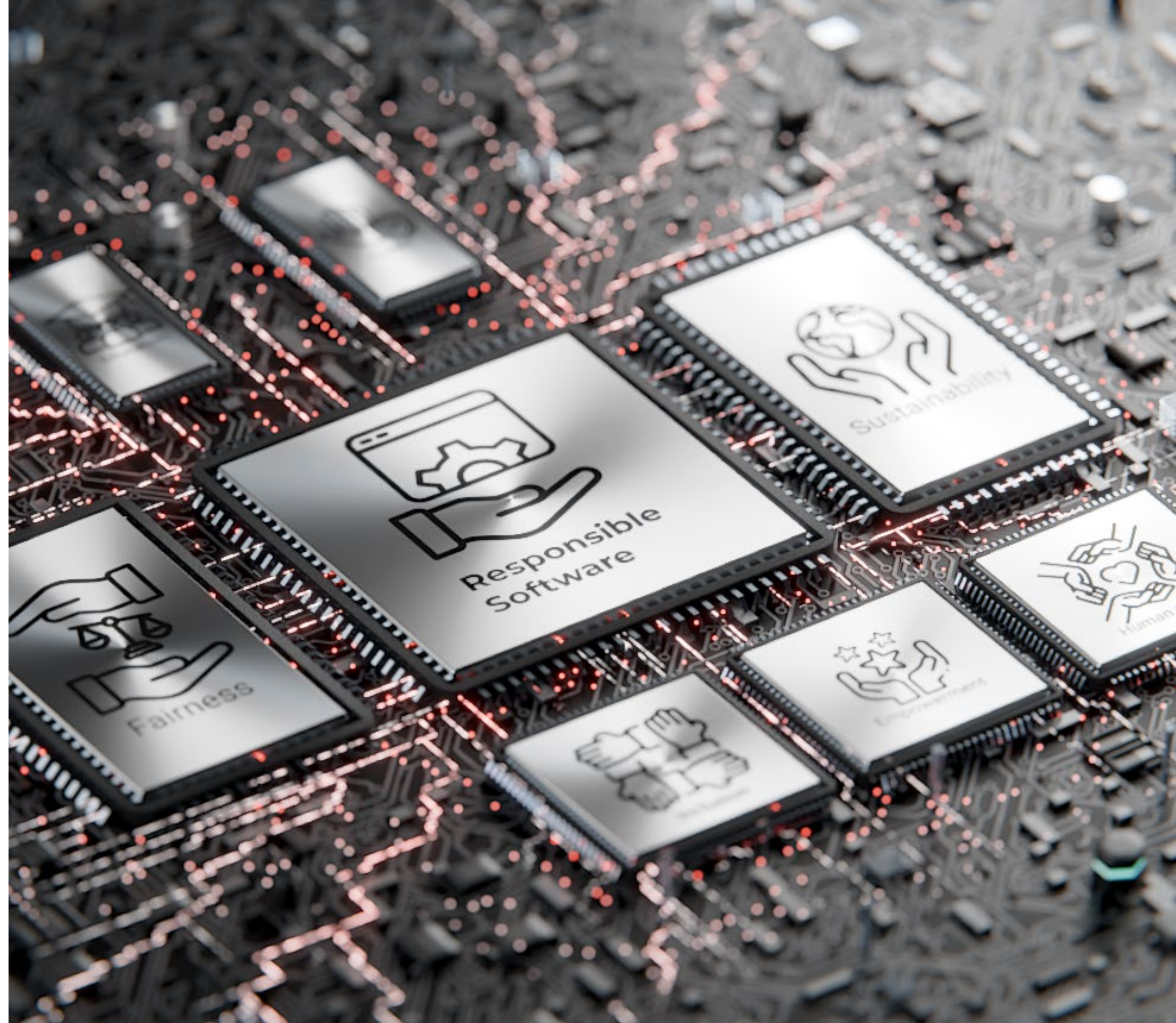


Sustainability 2 Review & Case studies 18 nov.

Cécile Hardebolle

**Responsible
Software**



The footprint of training - 1

URL: ttpoll.eu
Session ID: cs290

What are the 3 most important elements in the carbon footprint of ML training?

Rank them **by decreasing impact** (i.e. most impactful first) :

- a. The training time
- b. The power consumption of the CPU
- c. The power consumption of the GPU
- d. The PUE of the datacenter
- e. The carbon intensity of the electricity

The footprint of training - 2

URL: ttpoll.eu
Session ID: cs290

Let's consider the training of the model SupChat-7B. The computing node has 2 GPUs of the model Nvidia A100 80GB, which consume 400W each. Our datacenter, which has a PUE of 1.2, is located in Germany (carbon intensity: 381g CO₂e / kWh).

The training time is 80 000 hours of total GPU computation time.
What is the carbon footprint for the training of SUPMOD-7B?

- a. 14,63 tonnes CO₂e
- b. 29,26 tonnes CO₂e
- c. 14 630,4 tonnes CO₂e
- d. 29 260,8 tonnes CO₂e

The footprint of inference - 1

URL: ttpoll.eu
Session ID: cs290

What are the 3 most important elements in the carbon footprint of ML inference?

Rank them **by decreasing impact** (i.e. most impactful first) :

- a. The number of user queries
- b. The electricity consumed per query
- c. The PUE of the datacenter
- d. The carbon intensity of the electricity

The footprint of inference - 2

The model SupChat-7B is now deployed in production. It is hosted on the same computing node with 2 GPUs of the model Nvidia A100 80GB, which consume 400W each. Our datacenter, which has a PUE of 1.2, is located in Germany (carbon intensity: 381g CO₂e / kWh). Our model is able to serve 120 token per second of computation time. It has an average of 5000 users daily and generates an average of 5000 tokens per user per day.

1. What is the total GPU computation time used over 1 day (in h)?
2. What is the power consumed by the model for inference (in W)?
3. What is the total electricity consumed over 1 day (in kWh)?
4. What is the carbon footprint over 1 day (in kg CO₂e)?

Total carbon footprint

URL: ttpoll.eu

Session ID: cs290

We have obtained the carbon footprint of SupChat-7B at training and at inference time. What is its total carbon footprint?

- a. Training
- b. Inference
- c. Training x Inference
- d. Inference – Training
- e. Training + Inference
- f. Other

Hardware renewal

URL: ttpoll.eu

Session ID: cs290

We want to optimize the energy consumption of SupChat-7B at inference time. We decide to upgrade our hardware platform and to replace our A100 GPUS with H100 GPUs. The H100 are 4 times more performant than the A100 in terms of computation speed. Their power consumption is 700W at maximum use.

What effect(s) are we likely to observe (select all that apply)?

- a. A decrease in the energy consumption
- b. An increase in the energy consumption
- c. A decrease in the overall carbon footprint
- d. An increase in the overall carbon footprint

Water Usage Effectiveness

URL: ttpoll.eu

Session ID: cs290

The datacenter hosting SupChat-7B consumes an average of 1 MW. This means annually a total of 8 760 MWh of electricity. It consumes approximately 15.8 million liters of water each year. What is the WUE of the datacenter?

- a. 0,18
- b. 0,55
- c. 1,8
- d. 18,03
- e. 55,44