

Responsible by Design:

Every algorithm carries a worldview.
The question is: whose worldview will
shape our collective future?



Giuseppe Aceto
AI Ethicist



AI Is Never Neutral

94% of foundation models are controlled by 5 tech companies.

When we talk about "AI ethics," we're really asking three questions:

Power: Who decides what problems deserve solving?

Knowledge: Whose expertise counts as "ground truth"?

Values: What trade-offs are we willing to accept?

"The master's tools will never dismantle the master's house." — Audre Lorde

The Problem: Ethics Theater

Ethics Washing

Superficial frameworks that legitimize business-as-usual

Black Box Decision-Making

You can see inputs and outputs, but the logic inside? Trade secret.

Technical Gatekeeping

Complex jargon excludes the communities most affected by these systems

We're debugging symptoms while the architecture remains fundamentally broken.



Free Software Gave Us a Blueprint

"Free software is a matter of liberty, not price. Think free as in free speech, not as in free beer."

Richard Stallman

The Four Freedoms:

1. Run the program as you wish
2. Study how it works
3. Redistribute copies
4. Modify and improve it

But AI needs four more freedoms:

- Understand the training data
- Audit for bias and harm
- Govern deployment contexts
- Contest automated decisions

The Co-optation of "Open"

Open Washing in Action

"Open Source" → Readable code, proprietary training data

"Open Access" → Paywalled after 12 months

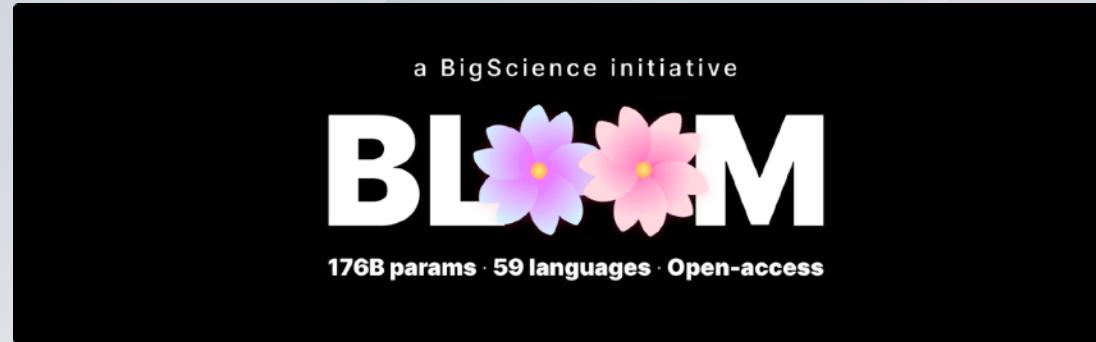
"Open Weights" → Download the model, but no documentation on how it was trained

True Ethical Openness:

Transparency + Governance + Community Control

Reading the recipe doesn't help if you can't access the ingredients or challenge the chef.

Competitive Advantage



What makes it different?

- **Governance:** 1,000+ researchers from 70+ countries voted on key decisions
- **Data transparency:** Full documentation of training corpus with consent processes
- **Responsible AI License:** Explicit restrictions on harmful uses (surveillance, disinformation)
- **Multilingual by design:** 46 languages, not just English-centric

The result? A 176B parameter model built collaboratively, not by corporate fiat.

The catch? Training cost ~\$7M. Even "open" AI requires massive resources.

Three Provocations

1. What if AI systems ran locally by default?

Privacy-preserving, no surveillance capitalism, community-owned infrastructure

2. What if data contributors got royalties?

Your photos trained Stable Diffusion. Did you get paid? Should you?

3. What if affected communities governed AI?

Not shareholders. Not just academics. The people whose lives are actually shaped by these systems.

These aren't fantasies—they're design choices we're not making.

From Individual Ethics to Collective Action

The problem with "ethical AI" discourse:
It asks individual developers to be ethical inside unethical structures.

The alternative:
Build community-governed infrastructures with ethics embedded at every layer—governance, architecture, datasets, licenses, culture.

Free software principles offer a blueprint: **technology created by and for communities, not corporations.**

What You Can Do Tomorrow

1. Support community AI projects

- Contribute to Hugging Face community spaces
- Join EleutherAI or BigScience working groups
- Use and improve tools like LocalAI for privacy-first deployment

3. Join the policy fight

- Support digital rights orgs: EFF, AlgorithmWatch, Access Now
- Comment on AI regulation proposals
- Bring affected communities into tech policy discussions

2. Demand transparency where you work

- Ask: "Who audited this model for bias?"
- Ask: "Can we explain this decision to affected users?"
- Push for documentation standards (Model Cards, Data Sheets)

Key Takeaways

Ethics must be structural, not cosmetic

Embedded from day one, not audited at the end

Free software principles are necessary but insufficient

We need new governance models for the AI era

Building ethical AI is a political project

Every technical choice reflects and reinforces power relationships

The question isn't whether we embed values in AI. It's whose values we embed.



Thanks

giuseppe.aceto@cern.ch

Let's build AI that distributes power rather than concentrates it.

