

Good morning.

I'm Víctor Moya, from the Library of the University of Seville, and it is a pleasure to be here with colleagues from Italy and from other European countries. Today I would like to invite everyone to reflect on the development of artificial intelligence, on how we use it, and on how it relates to Open Science.



Before getting started, I would like to congratulate the University of Genoa on organising GenOA Week. Devoting a whole week to open access, and opening it to other universities, is exactly what is needed: Open Science is built between institutions, not within each one. My warm thanks to the Library of the University of Genoa and the GenOA Week organisers for inviting me to take part in this event, and in particular to my colleagues Anna Maria and Marcella, whom I had the pleasure of meeting in Helsinki, and to Nuccia, for her help with the presentation.
And so, as you know, this year's GenOA Week theme is «The Cost of Knowledge», which invites us to ask how much it costs to create, share and sustain knowledge. And in the age of artificial intelligence, we cannot ignore how this technology affects the generation of knowledge and all of its associated costs.



We must begin by acknowledging that artificial intelligence is already part of academic work, and that we use it more and more in our daily tasks.
And we have to start from this premise because there is no going back. When we use an AI tool, we save time. But we need to be aware that this use has a cost.

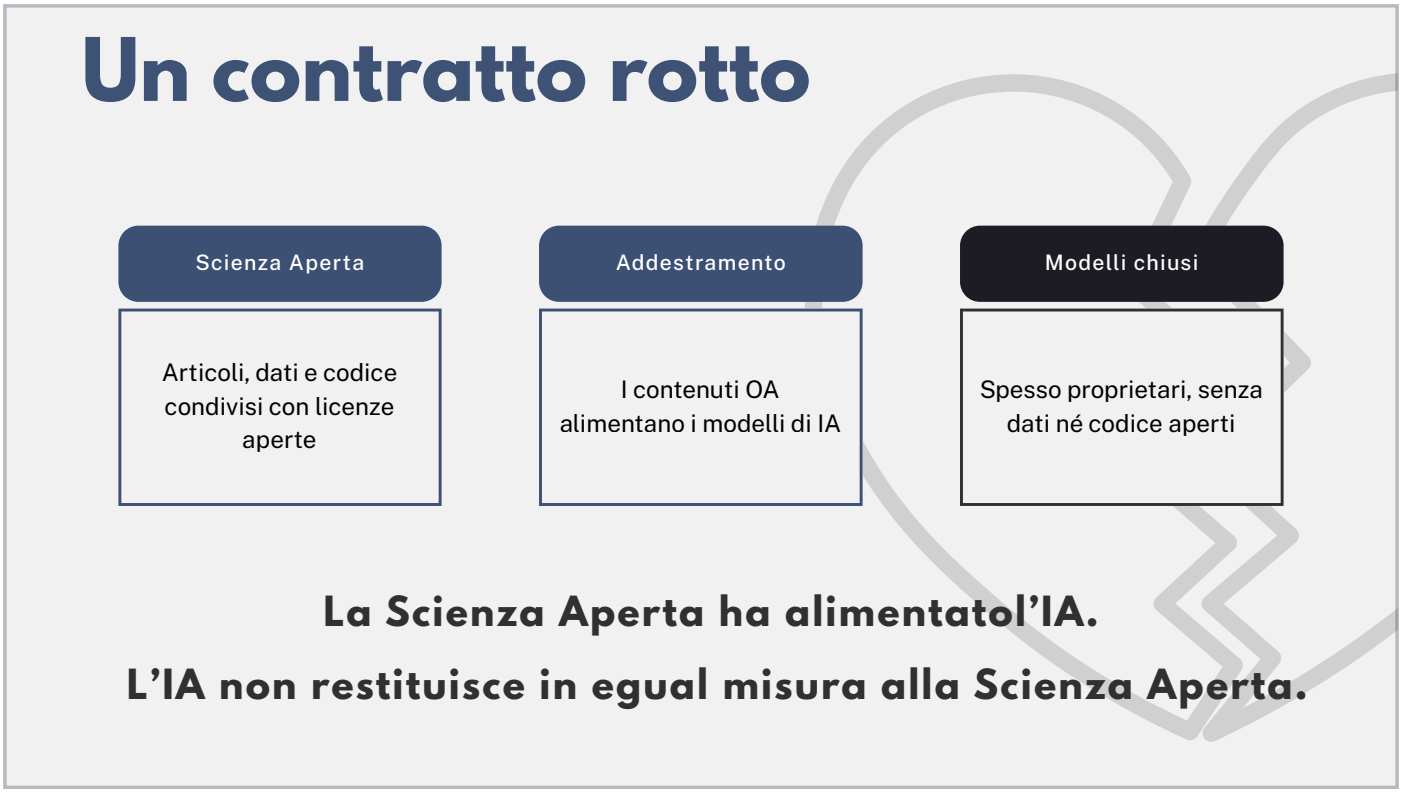


Some of these costs are well known, such as the costs to the planet. The rise in electricity and water consumption is widely known. To give just one figure, the International Energy Agency estimates that data centres will consume around 945 TWh in 2030, more than double today. In just four years. AI is not the only cause, but it is the main driver of that growth.

The second concerns power. Around 91 % of the relevant models of 2025 come from industry, not from academia. In 2021, academia still produced 28 % of those models; in 2025, 2 %. And 91 % of the market share is held by three models: ChatGPT, Gemini and Claude. Three companies hold in their hands everything we do with AI.

The third is knowledge. Of thirteen major developers assessed, eight do not disclose the data they use to train their models. That is, we do not know what data they were trained on, we cannot check where what they say comes from or cite it; we cannot explain their errors and biases, and we cannot reproduce their results. We are using models whose origin we cannot inspect.

And the fourth is rights, understood here as intellectual property. More than 70 % of the most widely used datasets do not specify the licence of use. Models are trained on content protected by intellectual property rights, and there is no way to check that those rights have been respected.



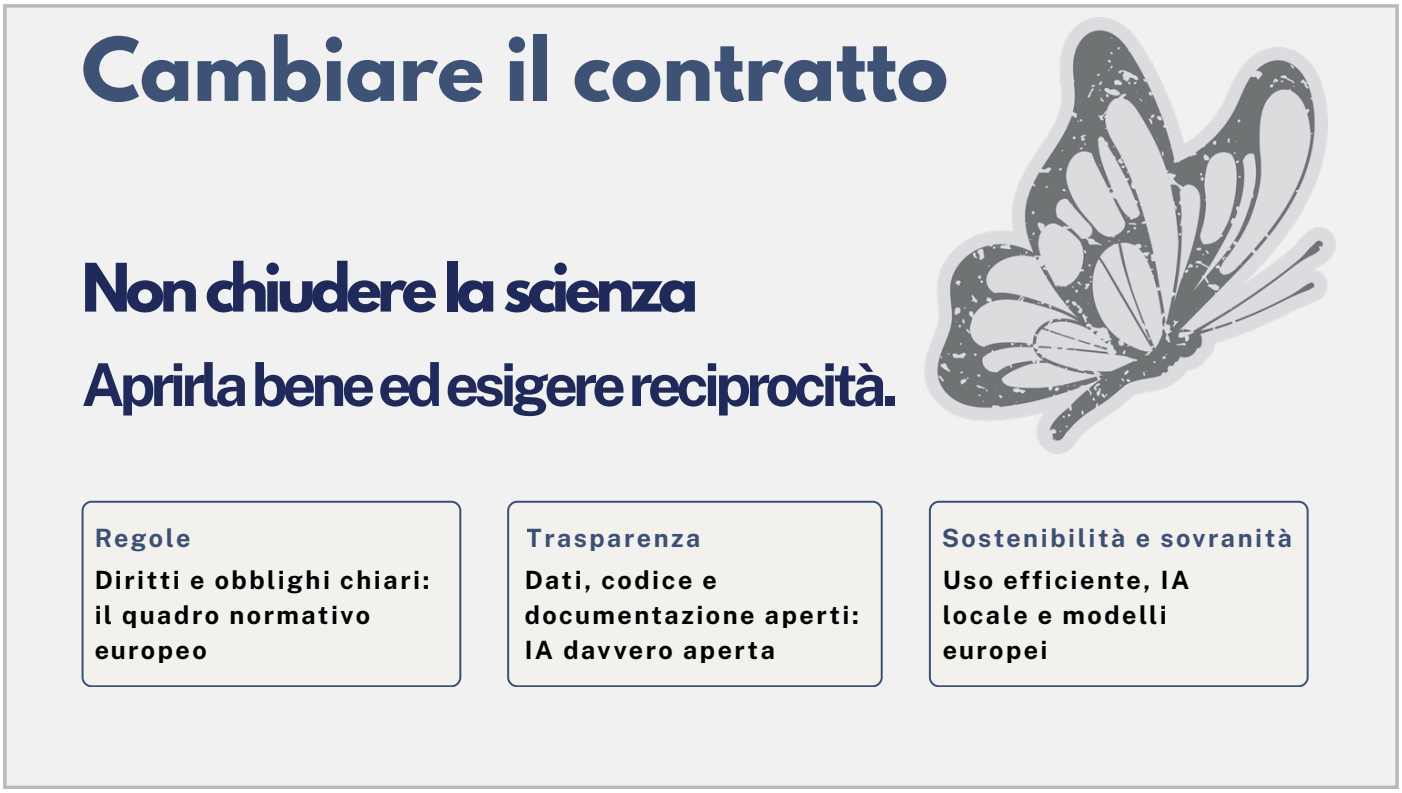
These four costs break a contract: the one we made when we opened up science.
For more than twenty years we have promoted open access publishing, and we have shared data, code and educational resources. All of this so that knowledge could circulate in a fairer and more accessible way, so that we could have a more ethical and safer science, and so that this would feed innovation and growth.
That open knowledge has been one of the pillars of AI training, which has extracted, processed and integrated high-quality scientific knowledge. And large companies have used all this knowledge to develop models that are bringing them great economic benefits.
But in many cases this knowledge ends up inside closed, opaque models that give nothing back to the community that fed them, and which now has to pay to use them.
It is something similar to what used to happen with the publishing industry and open access.
Open science has fed AI, but AI is not giving back to Open Science to the same extent.



We can therefore say that the development of generative artificial intelligence is putting the basic principles of Open Science under pressure.
Transparency and reproducibility clash with opaque, black-box models.
FAIR data clash with training data that nobody can consult.
Integrity clashes with hallucinations and increasingly blurred authorship and attribution.
Equity clashes with biases and with new digital divides.
Collaboration clashes with dependence on a few providers and with fragmentation.
And sustainability clashes with an environmental footprint that is growing out of control.



And here a dilemma for the future arises.
If AI does not give back its share, is it legitimate to ask whether we should keep feeding it with more open knowledge?
My thesis is that closing science to AI would not protect it: it would isolate it.
I see three risks.
1. Science would lose its capacity for discovery, because we would go back to closed, paywalled systems.
2. Researchers would lose visibility and impact, because what is not accessible is harder to find, read and reuse.
And 3. The risk of bias and lack of reproducibility in AI systems would increase, because models would be trained on a narrower and less reviewed portion of knowledge.
So closing science would not protect it. It would isolate it.



So what do we do?
From my point of view, my reflection is that it is not about closing science, but about opening it well and demanding reciprocity.
We have to try to change the contract. And this means three conditions.
1. Demand transparency: open data, code and documentation, that is, truly open AI.
2. Use it with criteria of sustainability and sovereignty: a sober and necessary use of AI, local AI and European models.
3. Set clear rules: clear rights and obligations, which in our case means the European regulatory framework.
Let’s look at this in more detail.



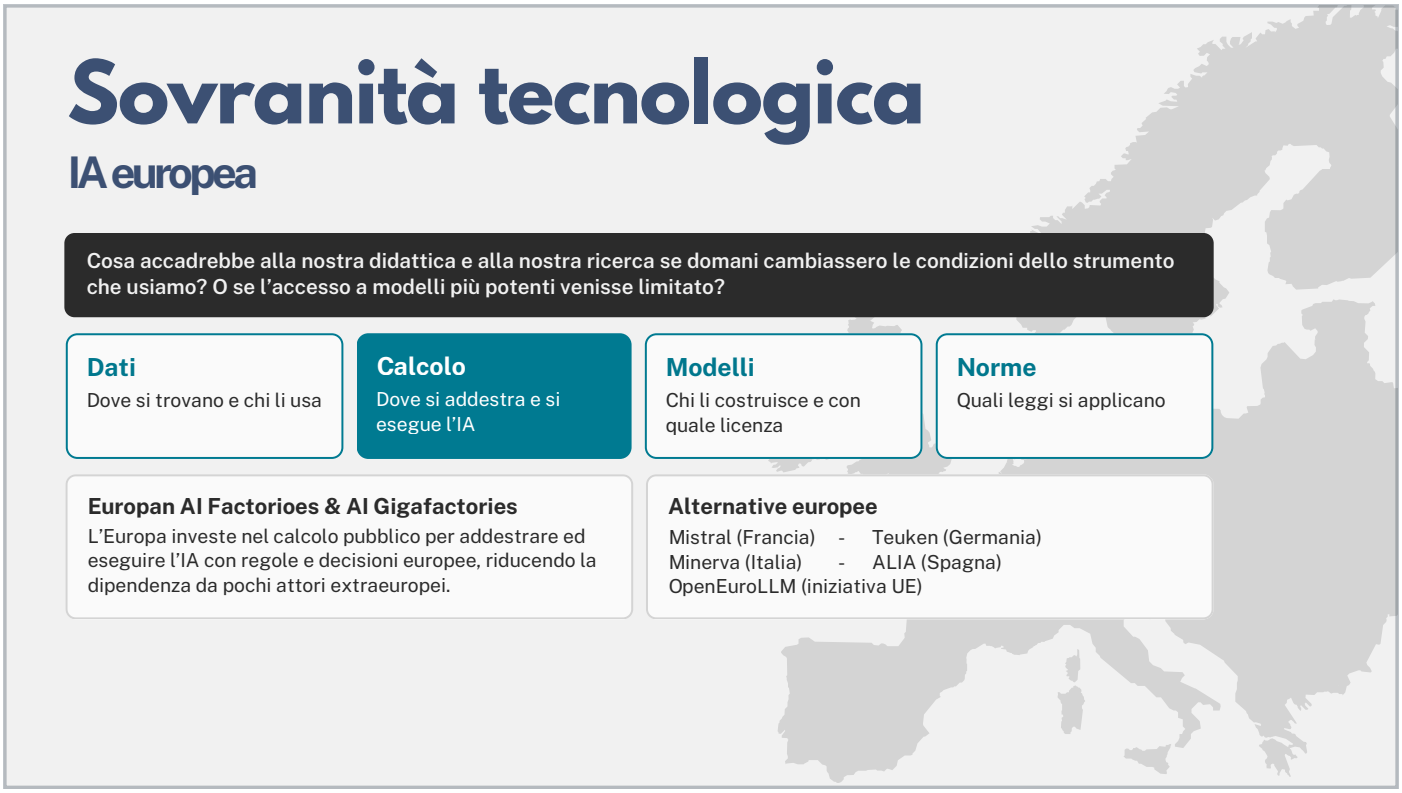
The first condition is to set fair rules of the game, which regulate with judgement and without blocking innovation.
In the European context, we have regulation on this.
The most important is the AI Act, as updated by the digital omnibus on AI, the first comprehensive law in the world to regulate the development, marketing and use of artificial intelligence in the European Union.
The second is the DSM Directive, whose aim is to update intellectual property law for the internet and digital platforms, and which regulates text and data mining, which is precisely what is done to train an AI. And we have other laws, such as the GDPR, on personal data protection, and so on.
We also have the European Commission's ERA Living Guidelines, which are recommendations for researchers, organisations and funders. They are a living document, updated as AI changes.
And at a third level, beyond the rules that countries may set, there must be institutional policies. Each institution must translate the European framework into its own rules, with criteria of openness and sustainability. Living policies, evolving, that protect researchers and science without holding back innovation or scientific progress.



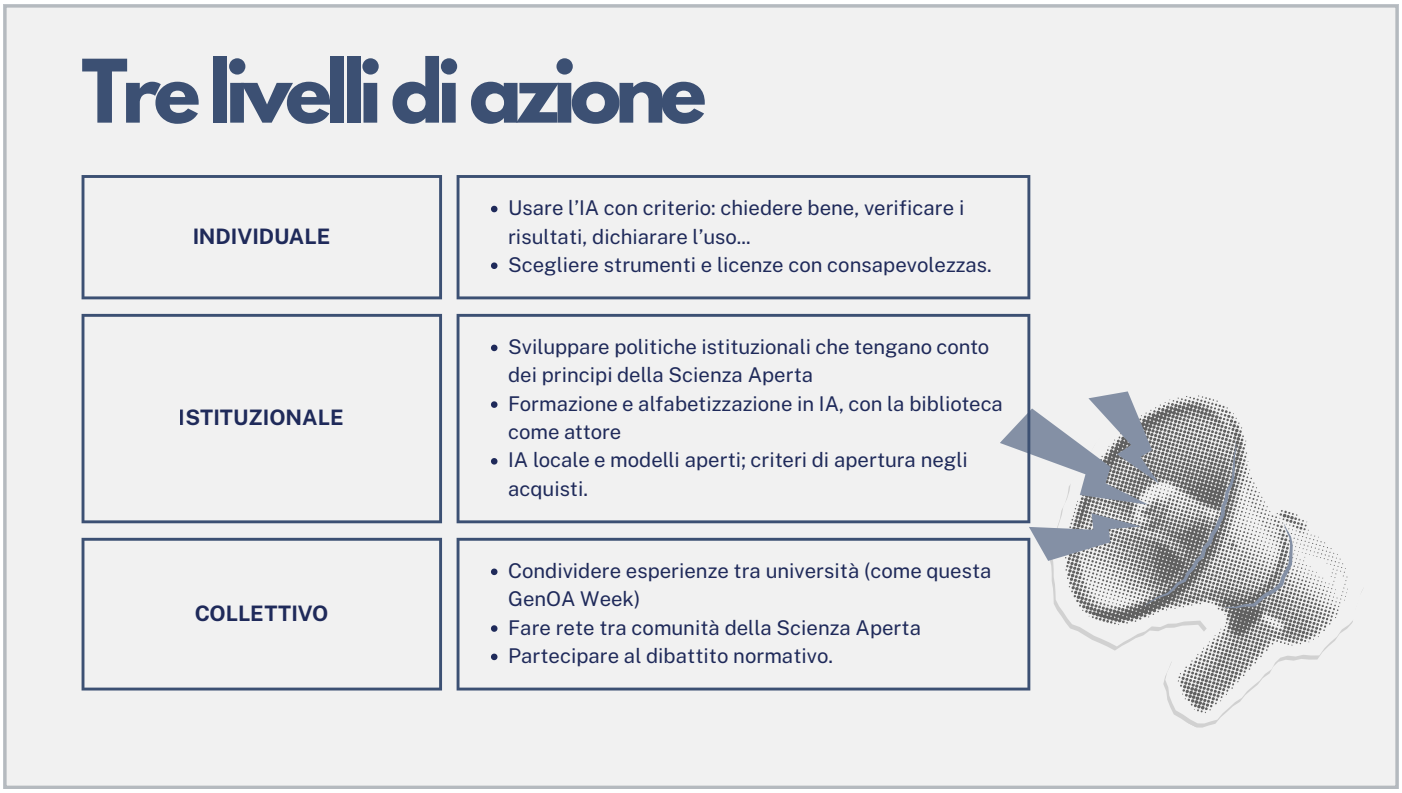
The second condition is to commit to truly open AI, and for that we need to know what «open» means.
On the market there are options advertised as «open source» that are in fact open-weight: the company publishes the model's final parameters so that they can be downloaded and run, but does not reveal the source code of the training algorithm, the full architecture or the data they were trained on. Some examples are Llama (from Meta), Gemma (from Google) or Mistral.
This would be the first level.
The second level corresponds to open source AI according to the Open Source Initiative, whose OSAID 1.0 definition, from 2024, requires publishing the code, the parameters and sufficient information about the data, under an open licence.
The third level is that of the Digital Public Goods Alliance, which starts from the OSAID definition but goes one step further, requiring that the training data also be open.
Beyond the definitions, the OSAID takes up the four freedoms that software must have to be considered free: to use, study, modify and share.
For science, the key is to study: being able to know how a model was built and with what data, in order to verify, explain errors and biases, and reproduce results.
Anything that falls short of this is openwashing: calling «open» what is merely downloadable.



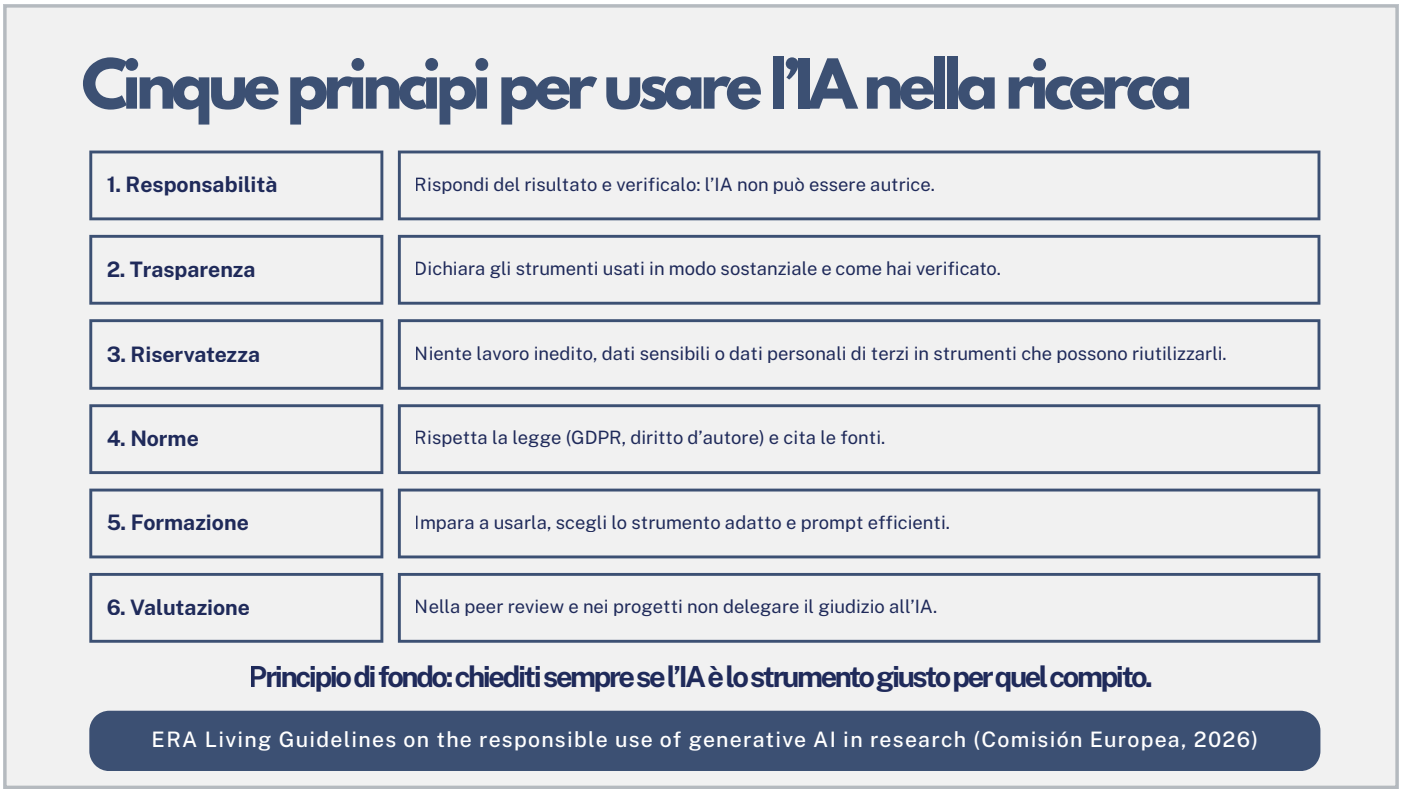
The third condition is a conscious and efficient use of AI.
The idea is to decide whether AI is needed and for what, and to use it well when we do. Trinity College Dublin published, in October 2026, a statement on generative AI and scientific integrity that sets non-use as the default option. It does not deny that some uses have value, but asks that use be a considered decision and that alternatives without AI be explored first. When we are about to use AI, the first question is «is it needed?». If there are alternatives without AI, there is no need to use it.
And when we do use it, let us do so with judgement and efficiently. This is what some people call green prompting: defining the format and length of the answer, being concise, choosing the right model, structuring before generating, and reusing the prompts that work. Keep in mind that we are looking for precision, not brevity: a prompt that is too terse lowers quality and forces retries.
We can also consider using local models, which can be run with tools such as Ollama, LM Studio or Open WebUI. These are models that we download to our own computer, or to institutional servers, and that keep our work from going out to external data centres. Although they also have limits, such as response times, they are worth considering for certain tasks, such as working with sensitive data.



Finally, I also wanted to refer to technological sovereignty: reducing dependence on the few companies that run the models we use in our universities.
One may ask here: what would happen to our teaching and our research if the conditions of the tool we use changed tomorrow? Or if access to the most powerful models were restricted?
But there is also a clear risk: it is about keeping the ability to decide, above all in the face of changes in the policies of countries such as the United States or China on the protection of their models, or of the continuing flow of European data to those countries.
When we talk about sovereignty, we do so at four layers: where our data are and who uses them, where AI is trained and run, who builds the models and under what licence, and which rules apply.
Europe is making a real effort not to miss the train of AI competitiveness. It is investing in the most expensive layer, which is computing capacity. There is already a network of 19 AI Factories in public supercomputing centres, and this summer the call was opened to build up to seven AI Gigafactories.
Although they are not yet as powerful as those coming from the United States or China, increasingly advanced European models are being developed. I would highlight, for example, the Mistral models from France, which are currently the most advanced in Europe, but also those trained on data in European languages, such as Minerva in Italy or ALIA in Spain.
I cannot at present propose replacing the large American or Chinese models, but I do believe we should start considering these seriously, above all for the tasks they are suited to.



After all this, how can we act to change that contract?
We should think of acting at three levels.
The first, of course, is individual: using AI with judgement. Asking well, checking the results, declaring when and how it has been used, and choosing models consciously.
The second level is institutional: committing to institutional policies that take the principles of Open Science into account, offering continuing training in the ethical and efficient use of AI, and opting for local, open and sovereign models whenever possible.
And finally, the collective level: sharing experiences between universities, as in this GenOA Week; building networks among Open Science communities; and taking part in the regulatory debate.
No level is enough on its own. Individual impact is small and systemic impact is large; that is why institutional and collective decisions weigh more.



I will close with what we can do at the individual level in scientific research.
These are the recommendations for researchers from the ERA Living Guidelines that I mentioned earlier.
As a principle, always ask yourself whether AI is the right tool for the task.
One, responsibility: you answer for the result and you verify it; AI cannot be an author.
Two, transparency: declare the tools you used substantially and how you verified.
Three, confidentiality: no unpublished work, sensitive data or third parties' personal data in tools that may reuse them.
Four, rules: respect the law, the GDPR and copyright, and cite your sources.
Five, training: learn to use it, choose the right tool and use efficient prompts.
Six, evaluation: in peer review and in projects, do not delegate judgement to AI.



To close, I would like to leave you with one idea: it is not about closing science to AI, but about changing the contract. Let Open Science open itself well, let it demand reciprocity, and let it do so with clear rules, efficient use and its own capacity to decide.
Thank you again to the University of Genoa for the invitation, and to all of you for attending.

Grazie mille!