



„We Have Plenty of Data. The Question Is: Which Data Can We Trust?"

Georges-Simon Ulrich · Director General of the Federal Statistical Office (FSO)
21st edition of the World Telecommunication/ICT Indicators Symposium (WTIS-26)

8 September 2026

ITU, Geneva, Switzerland

Ladies and Gentlemen,

Standing here, I am reminded of what the British historian Adam Tooze said: *“Welcome to the world of polycrisis.”* One crisis does not simply follow the other; they overlap, interact, and reinforce one another. Energy crises interact with conflicts, geopolitical fragmentation meets economic uncertainty, and misinformation undermines trust at the very moment when clarity is most needed.

In these uncertain times, statistics are more than numbers. They are our compass. They ground us in reality when public discourse may be dominated by fear, speculation, or deliberate distortion. I wish to reflect with you on the progress we have made, the challenges we face, and the commitments we must pursue to ensure that official statistics continue to serve as a trusted foundation for evidence-based decision-making worldwide.

1. Let me take a step back: What Has Changed?

Today, we have access to an unprecedented amount of data. The real question is: which data can we trust?

Twenty years ago, nobody would have asked that question. Data was scarce, difficult to access, and those who had it were considered experts. Today, the situation is exactly the opposite. Data is abundant, and that abundance creates a new challenge.

The data landscape has changed fundamentally, not only technologically, but also socially. Private actors now produce and publish data at a speed and scale that official statistical offices can never match, nor should they at any cost. With many of these data sources, transparency, methodological clarity, and institutional independence cannot always be guaranteed. They may be fast. But are they reliable?

At the same time, the way people access information has changed. Someone who wants to know how unemployment is evolving in Switzerland no longer starts by visiting the website of the Federal Statistical Office. Many do not even begin with a search engine. They ask ChatGPT or another AI system. Directly. Instantly. Conveniently. One answer instead of ten links. That is understandable. In many ways, it is elegant. But it has an important consequence, and it is one we need to discuss.

When people receive information through an AI model, they do not necessarily see the sources, the methodology, or the assumptions behind the answer. The model becomes an intermediary between citizens and the information produced by institutions such as ours. This raises a fundamental question: who determines which information is presented? This is not just a theoretical concern. A recent case in Bavaria, in Germany, illustrates this risk: Google was held liable after its AI-generated search summary falsely linked two publishers to fraudulent business practices and even generated claims that were not present in the underlying sources. The court emphasized that once AI combines and reformulates information into a seemingly authoritative answer, users cannot simply be expected to verify every claim themselves. This illustrates how AI hallucinations can not only distort information, but also obscure the path back to the original, reliable source. The ruling is not yet legally binding, and Google is expected to appeal the decision.

Our responsibility is not simply to produce data. It is to make that data traceable, transparent, comparable and understandable. We need to make sure that our statistics can be found, correctly interpreted, and reliably used, including by the AI systems that are increasingly becoming the interface through which people access information.

The example of AI models also teaches us another lesson. In a world where countries, institutions and companies increasingly depend on a relatively small number of technological platforms, diversity and resilience matter. We can already see why this matters in the rapidly evolving landscape of AI models. The emergence of highly capable models from different parts of the world, including China, Europe and the United States, is a positive sign of technological progress and innovation. We should welcome technological innovation from every region of the world. But we should also avoid replacing one form of dependency with another without understanding the implications. Different models may be trained on different data, subject to different governance frameworks, and optimized according to different objectives. They may also exhibit different forms of bias or alignment. This does not make one model inherently good and another inherently bad. It means that we need to understand the conditions under which an answer

is produced if we are going to rely on that answer. For official statistics, this is particularly important.

For us as a statistical community, this means that technological openness must go hand in hand with methodological independence, transparency and trust.

2.The Real Problem: The Access Path Is Disappearing

If you ask ChatGPT, "How has unemployment evolved in Switzerland?", you will receive an answer. It may even be based on data from the Swiss Federal Statistical Office. But you cannot be certain. You often cannot immediately assess where the information comes from, whether it is based on an official source, a blog post from several years ago, or an inference generated by the model itself.

Large language models (LLMs) combine information from many sources. For users, the origin and quality of that information are often not transparent. As a result, reliable statistics risk disappearing into a growing digital sea of information, not because they are unavailable, but because the direct path to them is being lost. The Federal Statistical Office risks losing direct contact with users because the route to information has fundamentally changed.

At the same time, one principle is non-negotiable: privacy. Large language models must never have access to sensitive individual data. What we are discussing here concerns only aggregated, anonymised statistical results that do not allow any conclusions about individuals or businesses.

3.Why This Is a Democratic Issue

Some may say: so what? That sounds like a problem for the Statistical Offices, not for me. I still get my answers. I would argue the opposite. This is not simply a technical issue. It is a democratic one.

Official statistics provide the foundation for public decision-making. Let me illustrate this with an example that applies in every country. Governments need to know how many children will enter the education system in ten years' time. They need to anticipate how many elderly people will require healthcare services, and where new housing and transport infrastructure will be needed. Those decisions depend on population statistics and demographic projections. If the underlying data are unreliable, the consequences are real: schools may be built in the wrong places, healthcare capacity may be insufficient, and public resources may be allocated inefficiently. This is why official statistics matter. Official statistics are produced according to strict principles: professional independence, quality assurance, sound methodology, transparent processes, and results that remain comparable across regions and over time.

Many of you work with data every day. You know exactly what I mean. The difference between a strong and a weak evidence base is not an academic question. It influences investments worth millions. It shapes legislation. It affects people's lives. Reliable and comparable statistics allow us to assess the consequences of public decisions.

They help us anticipate how many children will attend school in the future, how many people will require healthcare services, and where new housing will be needed. They enable national, regional, and local authorities to plan effectively and allocate resources wisely. They help policymakers design legislation based on real needs and evaluate whether those policies actually work.

If access to reliable statistics is weakened because AI systems bypass direct access to official sources and users can no longer distinguish trustworthy information from unreliable information, then this is not merely a question of convenience. It is a democratic challenge. A fact-based democracy depends on reliable, independent data. And those data must remain accessible to everyone, not only to specialists.

5. How the Federal Statistical Office Is Responding

The Federal Statistical Office could say: "We do our job. Anyone looking for high-quality data knows where to find us." But that would be naive. We cannot simply wait to be asked. We must act proactively. I see four points we have to tackle:

First, this means **transparency**. Our methods, quality standards, and uncertainties must be visible and understandable, not only to experts, but to everyone involved in decision-making. Statistics must not become a black box.

Second, it means **dialogue**. One-way communication is no longer enough. The Federal Statistical Office must be present wherever data is discussed, misinterpreted, distorted, or instrumentalised. And it must contribute. Not politically. But clearly. With conviction.

Third, it means **people**. Not only the institution itself should be visible, but also the experts behind it. The people who explain. Who provides context. And who earn trust.

Fourth, our statistical results must also be **discoverable by AI systems**. Even when the value of official statistics is widely recognised, people will increasingly turn to AI tools when they need information quickly. Our responsibility is therefore to ensure that these systems can find official results directly rather than relying on less reliable alternatives. **The voice of the Statistical Offices remains relevant not because it is loud, but because it is trustworthy.** And all of this rests

on a principle that is not negotiable: the professional independence of official statistics. Facts matter. They must not be softened, delayed, or instrumentalised. That is not merely an aspiration. It is a promise.

5. What We Are Doing in Practice

This discussion is not taking place in isolation. Many of us are facing the same pressures: limited resources, growing expectations, and increasing demands for real-time information. We cannot solve these challenges by working harder alone. We need stronger partnerships, shared standards, and a collective effort to ensure that trusted data remain visible and relevant.

Across the United Nations system, an important shift is underway. The conversation is no longer only about producing statistics. It is increasingly about data governance, data stewardship, interoperability, and trust. How do we ensure that high-quality data remain discoverable, understandable, and reusable in an increasingly complex digital ecosystem? How do we preserve transparency, accountability, and public trust as AI becomes a dominant interface between people and information?

Since 1946, the UN Statistical Commission has played a central role in shaping how the international community understands itself. It was created to ensure that official statistics are not fragmented by national boundaries but harmonized by shared principles. The UN Statistical Commission has always been more than a technical forum. It is a political achievement: it allows 193 countries, with all their differences, to speak a common statistical language. Since the adoption of the 2030 Agenda, however, its discussions have become increasingly strategic and political: statistics are no longer only about measuring reality, but also about supporting major global policy decisions and governing the data ecosystem on which those decisions rely. Even decisions that may appear highly technical can have significant political and legal implications, for example the UN resolution on the System of National Accounts. These are ultimately not merely technical questions, but questions about how the international community chooses to describe and govern reality.

These questions are now being discussed at the highest international level, and not only in our community of statisticians, and it is no coincidence that many of these discussions take place in Geneva. Geneva brings together an exceptional concentration of international organisations, statistical experts, digital governance specialists, humanitarian actors, and standards-setting bodies. It is one of the few places in the world where discussions on data, technology, public policy,

human rights, and international cooperation happen in the same ecosystem. International Geneva has an increasingly important role to play: helping build common principles, trusted infrastructures, and shared standards.

The Statistical Commission must evolve within a changing UN system. The Secretary General's **UN80** Initiative aims to make the UN more agile, coherent and effective by improving efficiency, modernising structures, aligning mandates and strengthening system wide coordination. It brings together reforms across peace, development, human rights, data and technology into a unified action plan. As these reforms move into their political phase, Member States will shape how far they go. The Statistical Commission should position itself as a strategic actor in this transformation ensuring that data, standards and statistical governance remain at the heart of a modernised UN.

Ladies and gentlemen, let me turn now to another pressing issue: The Global Digital Compact, adopted in 2024 as part of the UN Pact for the Future, has made clear that equitable and interoperable data governance is a global priority. This is not an abstract debate. Artificial intelligence systems are now shaping the information that billions of people see and use every day. Yet too often, these systems do not draw on official data, or they fail to distinguish authoritative sources from unverified ones. Our responsibility is to ensure that official data is accessible, open, interoperable, and - very importantly - AI-ready.

If standards are the grammar of our statistical language, metadata is its dictionary. Without metadata - without information about what our numbers mean, how they were collected, and under what definitions - they lose their power to be compared and trusted.

This is precisely the context in which we developed an ambitious initiative. The challenge we face today is not that trusted data do not exist. In fact, vast amounts of high-quality official statistics are produced every day across national statistical offices and international organisations. The challenge is **discoverability**.

The **Trusted Data Observatory**, or TDO is a global initiative launched by the Federal Statistical Office and the Swiss Federal Department of Foreign Affairs together with partners including the World Bank, the OECD, the IMF, and Eurostat.

In the digital age, metadata acquires an even greater role. AI systems cannot interpret data without clear, machine-readable metadata. Without harmonized standards, these systems risk misrepresenting or overlooking official data, thereby undermining trust.

The idea is simple, even if the implementation is not. The Trusted Data Observatory helps ensure that when someone asks ChatGPT or another AI system a question, reliable official statistics can be found and identified as such. The TDO does not host data. It describes them. In a machine-readable, standardised, and discoverable way. In essence, it rebuilds the access path. Directly into the world of large language models. This is Swiss pioneering work. Serving the global information infrastructure.

Metadata is often described as "data about data". It provides information about a dataset's content, quality, origin, methodology, and conditions of use. In an AI-enabled world, metadata is no longer a technical afterthought. It has become the gateway to **discoverability**. If data cannot be found or interpreted by either humans or machines, they cannot effectively inform decisions.

The Trusted Data Observatory addresses exactly this challenge. It works with open, standardised and machine-readable metadata that make trusted data sources easier to discover, understand and reuse. Importantly, the data remain with their original producers. The TDO focuses on metadata, creating a trusted discovery layer rather than a new data repository. The challenge before us is not to produce ever more data. It is to ensure that trusted data remain discoverable, interoperable and reusable across institutions, countries and technologies. This is not a task for one country or one organisation. It is a collective responsibility.

In practical terms, the objective is simple: when people, search engines, or AI systems look for information, trusted official statistics should not be invisible. They should be identifiable, discoverable, and clearly distinguished from less reliable sources.

6. Conclusion

Colleagues, the road ahead is not simple. We must continue to invest in capacity development. We must ensure that national data and statistical systems, including the NSOs in every country, regardless of resources, have access to the tools, training, and platforms they need. We must strengthen partnerships with international organizations, academia, and the private sector. Above all, we must remember that data is not an end in itself. Behind every number are human lives, human struggles, and human progress. New technologies give us new opportunities, but we should invest in capacity development to leave no one behind. Again, investing in data is investing in development.

Thank you. I look forward to your questions.