tdo trusted data observatory

For a world where trusted data is accessible to all, empowering informed decisions and countering misinformation.

The Trusted Data Observatory will serve as a central hub, making reliable information easier to find, understand, and use.

See the video

Contents

1 Why is trusted data important and how is reliability being challenged?

In today's world, the abundance of information, particularly online, makes it increasingly difficult to maintain focus and differentiate reliable information from misleading or inaccurate content. The internet enables the rapid sharing of content and facilitates the widespread circulation of both information, but also misinformation and disinformation. This misleading or incorrect content can create confusion and hinder international collaboration. While some of it is spread unintentionally (misinformation), in other cases it is deliberately generated to deceive or divide societies (disinformation and / or Fake News).

New technologies and artificial intelligence tools have advanced significantly, enabling the creation of highly plausible but sometimes inaccurate content, which can make it more challenging to distinguish reliable information from misleading material. Experts note that misinformation and disinformation are increasingly recognised as major global challenges. **The World Economic Forum's Global Risks Report** ranks them among the top long-term risks, given their potential to affect public trust in institutions and public administration.

These challenges are particularly relevant in the context of ongoing global conflicts and climate change. Governments therefore need to strengthen the resilience, accessibility and visibility of official information as a foundation for public confidence. Major technology companies play an important role in today's digital ecosystem and are increasingly collaborating with public authorities to support the dissemination and accessibility of trusted information.

At the same time, as digital infrastructures evolve, it is important to ensure that governance arrangements preserve national data stewardship and data sovereignty. This includes clearly defined roles, responsibilities and accountability frameworks, so that collaboration with private-sector partners strengthens, rather than dilutes, national oversight of data and digital infrastructures. Thoughtful partnership models are therefore essential to both innovation and sovereignty.

2 What are the challenges of accessing and using reliable data?

Data is the starting point for information, knowledge and insight. Finding the data you need when you need it, especially from trusted or official data sources, can be difficult. Lexical and AI-enabled search engines can compile such data in a very short time, but they collect information from any text sources of the world wide web, which can be a strength but also a weakness. Moreover, data primarily exists in separated areas, often referred to as data silos. The data is often collected or gathered for a specific purpose that sometimes does not fit with the nature of the query submitted. Lexical and AI-enabled search engines, while providing prompt results as outlined above, often fail to reliably discover trusted or official data sources, even when these sources provide high-quality open data. This is even the case where high-quality metadata is associated with the data but it is not visible, it is not open or machine readable.

3 Why does metadata matter for data discovery?

Metadata is "data about data" – and it is the foundation of discovery, accessibility, and comparability.

In an AI-enabled world, metadata is no longer secondary to data – it is the entry point. If data cannot be discovered, it cannot be understood, trusted, or reused. While official statistics for example are high quality and well-documented, they frequently remain invisible to search engines and AI systems because metadata is fragmented, inconsistently implemented, or not fully machine-readable and AI enabled.

When metadata is open, harmonised, machine-readable, and AI enabled it supports both humans and machines to:

- Identify authoritative data sources
- Assess relevance, timeliness, and provenance
- Understand access conditions and usage rights
- Distinguish trusted data sources from other data sources some of which can be misleading, lack international comparability or of low-quality content

Today, metadata standards can vary across organisations and countries, and implementation is often decentralised. Even where high-quality metadata exists, it may not be visible or optimised for AI-driven discovery. This creates barriers to discovery and increases the risk that inaccurate or unofficial sources dominate search results.

To address this fragmentation, the Trusted Data Observatory will work initially with the statistical community to agree on a **Minimum Viable Metadata (MVM) set**, a subset of the existing metadata holdings of trusted data compilers, aligned with an existing AI-enabled metadata standard. The objective is not to create a new standard, but to select and apply one of the recognised international standards that best supports AI-driven discovery.

Recognising that trusted data compilers currently use different metadata standards (such as SDMX, DDI, Dublin Core and others), the TDO will provide mechanisms to facilitate interoperability across the standards. Through automated extraction, transformation and reclassification processes, metadata can be mapped from existing national or organisational standards to the agreed MVM standard.

Where needed, the platform will include tools – such as metadata editing and conversion functionality – to support this transition. This approach minimises additional workload for data producers, respects existing metadata infrastructures, and ensures that discovery remains harmonised and machine-readable at global level.

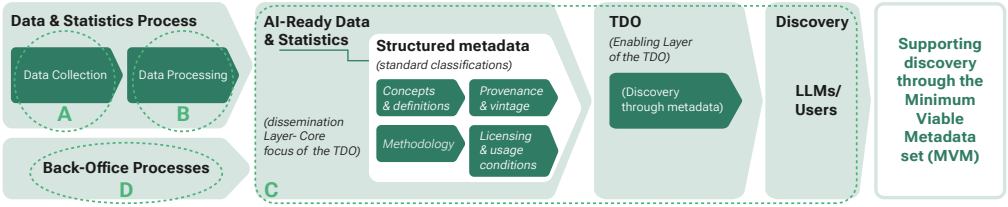
By centralising the systematic implementation of an agreed AI-enabled metadata standard, the TDO reduces fragmentation, enhances consistency, and ensures that AI systems engage with trusted metadata in a consistent, coherent and reliable way.

4 What is AI readiness?

AI-readiness for trusted data and statistics means ensuring that authoritative data and metadata are machine-readable, semantically interoperable, discoverable, and accompanied by clear provenance, vintage, and licensing information so that AI systems can retrieve and interpret them correctly.

It also includes the use of AI-enabled metadata standards – metadata that are structured, enriched, and exposed in ways that allow AI systems not only to discover trusted data sources, but to understand their meaning, context, limitations, and relationships across sources.

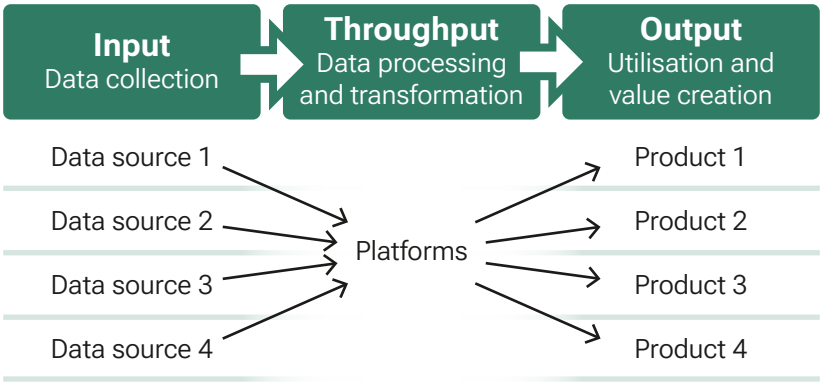
In this context, the TDO focuses specifically on the dissemination layer – strengthening discovery, providing details of access protocols for microdata, AI interfaces, and AI-enabled metadata.



5 How do metadata platforms enable data reuse?

Data collected for one purpose – for example administrative records – can often be reused for other purposes like statistical analyses. But often, the existence of such useful data is not known to those who want to create a data-based product. Metadata platforms make such data discoverable and thus enable the secondary or multiple use of data and the implementation of the "once only" principle, whereby individuals and companies only have to report certain information to the authorities once.

Harmonising metadata and making datasents discoverable and accessible can significantly increase the range of possible insights, products, and uses that can be developed. By enabling the discovery of related datasets across institutions and countries, users can more easily compare and combine data at an aggregate level to identify patterns, trends, and correlations. For example, aggregated data from different domains – such as health statistics and environmental indicators – can be analysed together to explore potential relationships and inform policy discussions. At the same time, harmonised and machine-readable metadata enables more automated data discovery and analysis processes, improving efficiency and reducing costs.



6 What is the Trusted Data Observatory (TDO)?

The Trusted Data Observatory (TDO) is a global metadata platform designed to address the challenge of data discovery in an increasingly AI-enabled world. It will describe trusted data, using the minimum number of metadata components necessary to supports discovery (MVM), from trusted sources using standardized, open, and machine-readable metadata, making this data visible, discoverable, and accessible for humans and machines alike.

The TDO is a metadata platform, not a data platform. Data remains under the control of the original data producers.

Even though trusted data compilers e.g. NSOs and International Organisations already compile high-quality trusted data, in many cases adhering to Linked Open Data standards, with associated metadata, on optimised websites, simple searches using standard search tools, or indeed AI supported search tools (ChatGPT, Gemini, Copilot etc), often provide plausible, yet inaccurate results. These LLMs are not discovering official data even with all of this in place. Metadata (language as distinct from numbers) is the key. LLMs are not currently good at engaging with numbers, but they can engage with metadata (language). Traditionally the focus of trusted data compilers has been on the data, with the organising and structuring of the metadata seen as secondary in some way. Access to the metadata has pretty much been facilitated by first accessing the data. Our paradigm has very much been Data to Metadata. In an AI enabled world, where our first task is to ensure trusted data is discovered (if it's not discovered it certainly won't be understood), that paradigm needs to be inverted to Metadata to Data.

The initiative is driven by Switzerland in collaboration with the Geneva data community and other interested partners, with the ambition to integrate the platform into official international and UN-related processes over time.

7 What are the objectives of the TDO?

- Support the discoverability, accessibility, and reuse of trusted data.
- Enable AI-readability of official data and statistics through machine-readable and AI enabled metadata.
- Bring together existing national and international metadata platforms.
    - Increase transparency and trust in official data.
    - Support evidence-based policymaking, research, and public discourse.
    - Respect data sovereignty by ensuring data remains under the control of their original producers.
    - Provide a neutral, trusted discovery layer for trusted data and statistics.
    - Simplify the discovery of trusted data for both humans and machines

8 From feasibility to implementation

In 2024, a feasibility study commissioned by the Swiss Federal Department of Foreign Affairs and the Swiss Federal Statistical Office and conducted by an independent consultant, concluded that the potential added value of a Trusted Data Observatory is significant.

Since then, the project has progressed into implementation planning, stakeholder engagement, and technical preparation phases. Key challenges identified include metadata quality and heterogeneity, governance and sustainability, and ensuring broad international participation. These challenges are being addressed through a phased, collaborative approach involving national statistical offices, international organisations, and technical experts.

The Phase 1 report has been completed and is ready to be published on the webpage.

9 How the TDO improves discovery of trusted data?

By providing centralized access to harmonised discovery metadata, the TDO will enable users and AI systems to identify who produces which trusted datasets, how frequently they are updated, what their coverage is, and how they can be accessed.

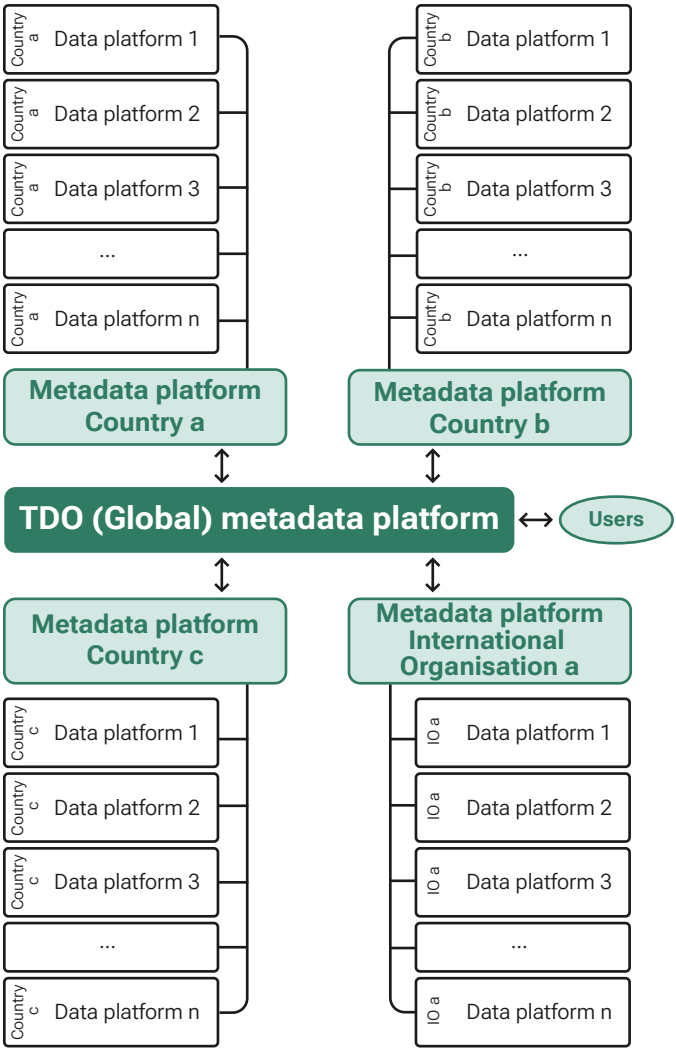
Where aggregate data is open, users can access them directly through existing platforms. Where access is restricted, the TDO will provide clear information on access conditions and points of contact.

By acting as a trusted discovery layer, the TDO will improve the quality of search results produced by search engines and AI tools, helping to counter misinformation, disinformation and misinterpretation.

10 How does the TDO work in practice?

Data on various national platforms and from international organisations will be made visible through a country's or international organisation's metadata catalogue. The data itself remains on its own platform. If the metadata platforms are harmonised, the Trusted Data Observatory can connect to all these platforms. Thus, all trusted data sets could be made visible in a global platform (TDO).

Moreover, by focusing on metadata, the data stays on the data producer's own platform. This means that the Observatory does not modify data ownership and quality assurance responsibilities but creates benefits and synergies for all participating data owners.



11 How the TDO is being developed?

Building on the feasibility study, the Trusted Data Observatory is being developed through a phased, collaborative approach towards full implementation:

Phase 1 Stakeholder identification, governance planning, and requirements definition have commenced in close cooperation with international organisations and national statistical offices.	To initiate this process, targeted stakeholder engagement has been undertaken, including structured discussions and working sessions with key international organisations such as the World Bank, OECD/PARIS21, IMF, UNSD, UNECE, Eurostat, and selected National Statistical Offices. This phase is critical and focuses on establishing a shared understanding of objectives, identifying user needs, and developing initial specifications for the Trusted Data Observatory. The skills supporting this phase include project management, statistical expertise covering data and metadata standards, and systems and business analysis. Initial governance arrangements and seed funding for the Proof of Concept (PoC) are being defined as part of this phase.	**2025 – 2026**
Phase 2 Agreement on a Minimum Viable Metadata (MVM) set, associated standards, technical specifications, and Proof of Concept (PoC) participants.	The objective of this phase is to identify and engage a group of interested partners – a collaboration of willing National Statistical Institutes (NSIs) and International Organisations (IOs) – to participate in the Proof of Concept (PoC). These partners are intended to reflect the global ambition of the project and include organisations with varying levels of metadata maturity. This diversity will help assess the practical workload associated with participation in the PoC and inform the design of potential capacity-building support within the framework of the project. Engagement with universities, academic institutions, and philanthropic organisations is also being considered to support innovation, research, and sustainability.	**2025 – 2026**
Phase 3 Development of a scalable TDO prototype, implementation of the PoC, training of participants, and engagement with AI and technology stakeholders.	Complete the technical specification for the observatory. Observatory goes online, core issues are resolved and information is accessible. Develop a prototype in a manner that is scalable. Encourage the inclusion of the observa- tory in UN strategies. Switzerland to lead with partners. Ensure sustainable funding & governance structures. Commence work on an engagement and communica- tions plan to generate knowledge and interest in the platform.	**End 2026 – Q1 2027**
Phase 4 Review of PoC	Review and refine the TDO post PoC exercise	**By end 2027**
Phase 5 Broader onboarding, integration with post-2030 agenda, and long-term sustainability.	Onboard key partners (NSI's and IO's not involved in PoC) and integrate with the "Post-2030 Agenda". Build communications and engagement plan from phase 3 to drive engagement of data compilers with the observatory and drive usage of the observatory by the user community. Examine a broader suite of trusted data sources within an agreed framework developed by the governance structure put in place for the TDO.	**2027 – 2030**

The phases 1 and 2 of the Trusted Data Observatory are being implemented in close cooperation with international organisations based in Geneva and national statistical offices from all regions of the world. Data for Change – the PARIS21 Foundation is supporting coordination activities in Geneva during the early implementation phases.

12 Where will it be hosted?

Geneva provides a unique and well-suited environment for hosting the Trusted Data Observatory . It is home to many United Nations bodies and international organisations that are major producers and users of official data, and it offers a long-standing tradition of neutrality and multilateral cooperation. Moreover, Switzerland is currently setting up its own national metadata platform (I14Y), which means that know-how can be easily transferred. While Geneva offers a strong institutional ecosystem for the development of the TDO, its long-term governance will be anchored at the global level, reflecting the Observatory's international scope, multilateral character, and worldwide reach.

13 How to engage and collaborate?

Swiss Federal Statistical Office
International and national affairs
international@bfs.admin.ch