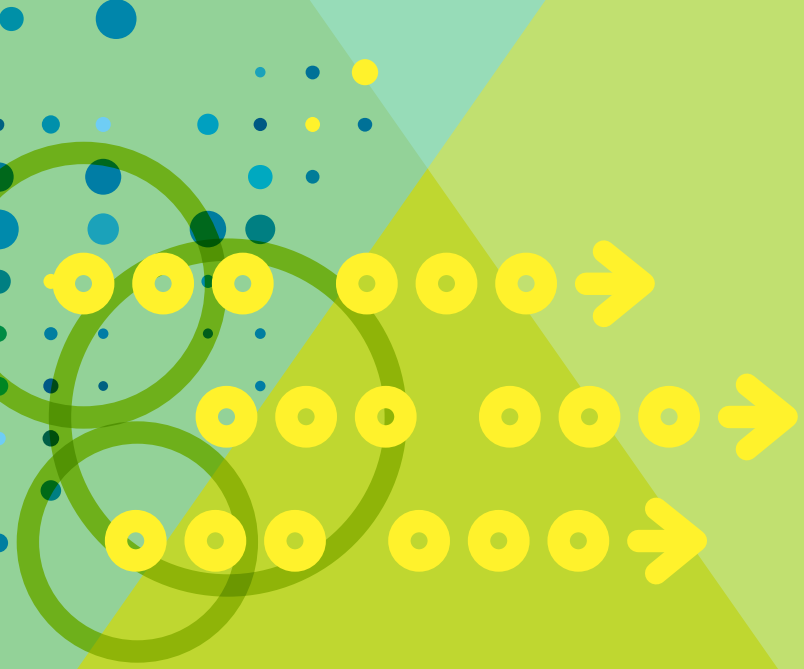
Driving Progress with Data

A strategy for more and better data for new, effective and forward-looking data use



The
Federal Government

Driving Progress with Data

A strategy for more and better data for new, effective and forward-looking data use

Preface

While in the late nineteen-eighties 99 percent of all data was still analogue, 99 percent of data today is digital. We all deal with data every day, generating and using it on our smartphones, computers or at our workplaces. The amount of data generated worldwide each year has increased fiftyfold since 2010, rising from 2 zettabytes (= 2 billion terabytes) to over 100 zettabytes in 2022. And in just three years, between 2019 and 2022, the data volume per stationary broadband connection has more than doubled from 135 GB to 274 GB. Data maps the world in a wide variety of ways. The ability to 'read' it can fundamentally improve human lives.

Using data helps improve healthcare, delivers new findings to the scientific community and creates new opportunities for the education sector, enhances production workflows, promotes innovative government action, helps preserve resources and makes everyday life easier in many ways. Data is the key to the digital and ecological transformation of society, the economy, the scientific community and the public sector.

In particular, the highly dynamic development of digital technologies and applications like artificial intelligence (AI) makes enhancing access to (especially high quality) data as well as the effectiveness and the efficiency of responsible data use increasingly important. As a foundation for machine learning, data facilitates pattern detection, probability calculation, development of relevant models, forecasting and preparation of decisions on this basis, production of text, images, audio and video, as well as problem-solving.

For example, AI can help diagnose cancer more reliably and faster, translate languages in real time or optimize machine maintenance.

A lot of valuable data goes unused in Germany – and this includes open data. For example, roughly 80 percent of data generated by industry is not currently reused.

Some areas are still collecting too little data, unsuitable data or data of insufficient quality.

In addition, a lot of data cannot be found, is not accessible, is not interoperable or cannot be re-used as the required licence conditions are not in place.

In a survey by the Association of German Chambers of Industry and Commerce (DIHK), 42 percent of companies reported a lack of in-house expertise on data use as an obstacle to its use.

There remains a lot of uncertainty regarding data use and sharing: For instance, 56 percent of companies surveyed by the Association for Information Technology, Telecommunications and New Media (Bitkom) reported that they do not provide any data because they assume that this is not permitted for reasons of data protection. 35 percent were not sure whether sharing is legally possible.

We want to improve this situation and better harness the potential of data.

At the same time, data can contain both sensitive personal information and business secrets or represent intellectual property. Since the 70s, strict data protection standards have taken the need for security and protection of personal data into account. The protection of commercial interests in data processing is also legally recognized. Data use and data protection are two sides of the same coin. Data policy is influenced by both objectives, which must be balanced consistently and carefully: How can we bring about the necessary increase in data access and use in compliance with the fundamental right to informational self-determination? How does the data protection law principle of collection for a specific purpose and data minimization or effective protection of business secrets and intellectual property fit with the declared objective of sharing more data? And how can we resolve the conflicting objectives between the EU aim of digital sovereignty and the concept of free movement of data and information?

Technological progress and data-driven innovations must appropriately take into consideration fundamental rights and property rights. We want to find appropriate and innovative solutions in each case that represent the greatest common factor, not the lowest common denominator: by putting into place intelligent frameworks and responsible data use.

The new Federal Government Data Strategy describes pathways to responsible, effective and future-proof data use and formulates a roadmap for the years to come.

It aims to improve the innovativeness, transformational capabilities and competitiveness of industry, the scientific community, the public sector and society. It permits more innovative and responsible data use for public benefit and therefore promotes societal progress. With this Data Strategy, we want to show how data use in Germany will be improved to benefit our citizens and the environment, as well as the opportunities this entails. It is intended to help form a society where clear rules effectively protect the right to informational self-determination. We believe that a transparent, understandable, participative and proficient use of data is necessary to clearly illustrate the importance of data for citizens. In our opinion, this includes ensuring that citizens are aware of the importance of data and have the knowledge they need to use it, and that data is accessible for all. We want a society in which citizens can simply decide to whom they want to provide specific personal data and choose the purposes for which they do so. We also need a healthcare system in which people can be sure that the medical staff has all relevant information and can use it to provide the best possible care. We are ensuring that the scientific community has access to data for research purposes. We are laying the groundwork for a productive and future-proof working environment, supported by responsible data use on the part of employees and companies. We are working towards a digital and interconnected economy that helps improve products and boost sustainability.

Effective use of data delivers a key contribution to the timely achievement of the UN 2030 Agenda and its sustainability goals.

We want public institutions to be able to base their decisions on the solid foundation of relevant, high-quality datasets. And we are striving to be a government that provides digital services in an innovative and user-friendly way, driving a reduction in red tape. We are committed to ensuring that data reflects the diversity of our society and is free of stereotypes, discrimination and imbalances. This includes bridging the gender and age data gaps.

European and national law, as well as the Federal Government's Digital Strategy (2022), as an overall strategy, form the foundation and framework. The National Digital Strategy contains the strategically relevant projects of the individual government departments in the digital sector.

We are building on the measures adopted in the 2021 Data Strategy, some of which are already being implemented, and going one step further in terms of data provision ('More data'), data quality ('Better data') and data use. In addition, establishing and enhancing data literacy across all sectors are key to successful and responsible data use.

The Data Strategy helps us work towards the targets of the Federal Government's Future Research and Innovation Strategy.

It contributes to strengthening the digital and technological sovereignty of Germany and Europe. In this way, the Data Strategy helps add value and boost competitiveness. The new Data Strategy is accompanied by the AI Strategy, the Cyber Security Strategy, the Open Data Strategy and the Gigabit Strategy.

The new Data Strategy is to initiate an open, progressive and opportunity-driven implementation of the existing legal requirements to enhance data use for public benefit and commercial purposes, and create planning certainty. To this end, the valid relevant regulatory framework should be reviewed to determine whether it meets these objectives.

Where necessary, we are adapting regulatory frameworks. As we do so, we ensure that data protection and IT security standards, security concerns to safeguard public safety, security and order and the protection of business secrets and intellectual property remain unaffected.

In order to reach our goals, we are focusing on scalable data use and a low administrative burden. The ways in which data can be accessed and used in an efficient and responsible manner are to be expanded. At the same time, public trust must be strengthened when it comes to using data. For these reasons, the new Data Strategy focuses on the opportunities of data use for the upcoming transformation processes and the responsible handling of data in our interconnected society.

To this end, it is important that we break down data silos, use and develop data standards and work together to make effective generation, collection, use and management of data second nature in Germany.

The Strategy will also promote concerted action as its vision. Effective and future-proof data use is a challenge that can only be overcome if industry, the scientific community, the public sector and civil society join forces and work together.

Contents

Preface 4

1 More Data 10

1.1 Public sector and research 11
1.1.1 We are making data easier to find 11
1.1.2 We are expanding access to government data assets 12
1.1.3 We are opening access to data from Federal funding projects and research contracts 13
1.1.4 We are putting in place the framework conditions for more data for public benefit 14
1.1.5 We are opening data for research 17

1.2 Private individuals and companies 18
1.2.1 We are supporting the sharing of model contract clauses and best practices 18
1.2.2 We are putting in place a framework for more investments in the data economy 18
1.2.3 We are using competition law to make data sharing easier 18
1.2.4 We are making data easier to access in the healthcare sector 19
1.2.5 We are simplifying data protection and facilitating practical implementation 20
1.2.6 We are creating data security with security by design 21

2 Better Data 22

2.1 We are relying on standardized data descriptions (data labelling by design and by default) 23

2.2 We are supporting competitiveness with interoperable standards and quality assurance 24

2.3 We are supporting the development of standards for trust-based handling of data 24

2.4 We are increasing the quality of government data 25

3 Data Use and Data Culture ... 26

3.1 The added value of data use ... 27

3.1.1 We are focusing on data-based government action at all administrative levels ... 27

3.1.2 We are supporting the sharing of sample use cases ... 28

3.1.3 We are simplifying the use of unstructured data for Large Language Models (LLMs) by the public sector ... 28

3.2 Data networking and data exchange ... 28

3.2.1 We are supporting data spaces and their networking as a data use instrument ... 28

3.2.2 We are founding a data institute ... 30

3.2.3 We are committed to a reliable exchange of data across borders (Data Free Flow with Trust) ... 30

3.3 Acceptance of data use ... 31

3.3.1 We are strengthening the resource-efficient use of data and its use for sustainability goals ... 31

3.3.2 We are ensuring data-driven innovation while protecting intellectual property and business secrets ... 31

3.4 Comprehensive data literacy ... 31

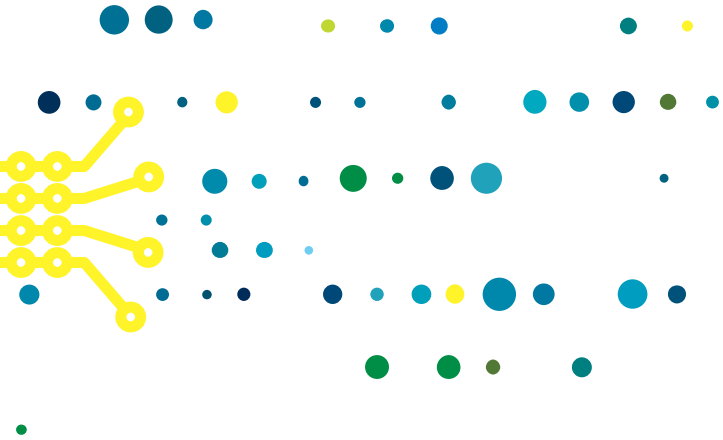
3.5 Courageous and responsible data culture ... 33

4 Implementation ... 36

Data Strategy Roadmap for the period to Q4/2024 ... 37

Publication data ... 38

1 MORE DATA



1.1
Public sector and research

1.1.1 We are making data easier to find

We are compiling a data atlas of the Federal administration to show data from all ministries and their agencies at a metadata level. In this way, we are creating transparency on the existing data assets. In addition, we will drive forward the establishment of a Federal administration data pool for machine-readable data.
The data atlas and data pool will become the foundation for data-driven processes and decisions in Federal authorities. They will facilitate more effective knowledge management and support the preparation and provision of data as open data. On this basis, public authorities can also share data across departmental boundaries for effective and future-proof administrative governance.

The data atlas uses and supplements existing administrative data overviews like the Administrative Data Information Platform (VIP) of the Federal Statistical Office, the Register Map of the Federal Office of Administration or the GovData Metadata Portal for open data of the Federation, federal states and local authorities. The data labs in the ministries are responsible for the data atlas.

We will create a publicly accessible, curated, online overview of existing public and private data portals (a landing page as a central starting point). We will strive to have data assets catalogued in a machine-readable form at the sources (for open government data, this includes incorporation in GovData.de), to make the data easier to find. The description of the data in data catalogues includes in particular standardized metadata. This can result in applications for data of a variety of formats, thus driving shared use of data, also between different user groups and data spaces.

1.1.2 We are expanding access to government data assets

The government is setting a good example in data provision. Official information and administrative data can already be accessed in many ways. However, the requirements for this are governed in different acts, and vary widely in some cases. The publication practice to date has also been inconsistent.

We are going to harmonize the legal bases for access to official information and administrative data. To this end, we are going to evolve the freedom of information laws into a Federal Transparency Act. Of course, the individual right of access will remain in place. However, we will also merge and extend the Federation's obligations to proactively publish certain categories of data.

We are also introducing a legal entitlement to open data. This will make the existing obligation to publish specific administrative data under open data requirements (re-usable, machine-readable, interoperable, open licence) legally enforceable.

Through these legislative projects, the Federation is taking big steps towards openness and transparency. This will take the information access options enshrined in law to a new level, improve cooperation within the government, promote scrutiny of the administrative and governmental action and strengthen participation options for citizens.

In the future, we will make far more public sector data accessible at a higher quality level, which will benefit industry, the scientific community, civil society as well as the public sector itself, as its open re-use can unleash creative potential to tackle the challenges of tomorrow. This is particularly true for rural areas. Of course, particular needs for protection and justified, proportionate access restrictions, which must be defined precisely, will be taken into consideration in the implementation.

We are providing more data on public infrastructure. The public sector holds a lot of data on traffic flows, the transport infrastructure (e.g. road works) and its occupancy levels. Harnessing this data is key for sustainable mobility. The Mobility Data Act is to enable more and better travel and transport infrastructure data to be made available and reused on fair terms. The main objectives of the Act will therefore be data availability across all modes of transport, better data quality and clear rules on data use. In addition, the Mobility Data Act will create a structural framework to fulfil the data provision obligations and strengthen the mobility data infrastructure institutionally and organizationally. Rules on data provision will address government and private sector data stakeholders equally.

In the area of development cooperation, we are making the foreign development contributions funded by the Federal Government available in a transparency portal in a user-friendly format and as open data.

In their data-based procedures, the Federal authorities are integrating processes that ensure that open data is made available and can be found. The entire life cycle of data, from creation to archiving or deletion, must be taken into consideration in this context.

In order to be able to publish the data in accordance with the FAIR criteria ('findable', 'accessible', 'interoperable' and 're-usable'), a repository is being built to allow the allocation of persistent identifiers, comprehensive metadata and registration as well as indexing in metasearch engines.

Research data centres promote data sharing between the public sector and external stakeholders (especially in the research community). Centres of this kind have already been successfully established in some areas. All ministries and executive agencies are currently assessing the creation of additional research data centres.

1.1.3 We are opening access to data from Federal funding projects and research contracts

The Federal Government will consider an amendment to the ancillary provisions for project funding grants, with the objective of making the data collected or generated in research and scientific projects (raw data, processed data and metadata) publicly accessible in principle, taking into account justified protected interests and property rights (e.g. privacy and personal dignity, business and commercial secrets, intellectual property), while also allowing its re-use. In the art and culture sector, we will take into account the particular interests of the stakeholders. In the future, where possible, data and metadata is to be provided in accordance with the FAIR principles.

We are introducing a uniform EU-wide and internationally compatible standard for handling data generated or collected in Federal funding projects, and thus creating the prerequisites for largely automated metadata collection for analysis of the Federal Government's involvement in Germany and other countries (country profiles).

We will assess whether and how the datasets generated as part of government research contracts, which are or have been published in the form of studies, surveys and monitoring reports, for example, can also be made accessible for research (e.g. secondary analyses) via scientific use files.

To this end, a uniform standard could be introduced in contract award procedures and in the drafting of contracts for the provision of research data, taking into consideration property rights, when awarding research contracts. Access to this research data for scientific use can be facilitated via the existing research data infrastructure in Germany (research data centres and data archives).

1.1.4 We are putting in place the framework conditions for more data for public benefit

Our goal is to ensure fair and, where possible and expedient, open access to data. The government also needs more and better data for its activities. Public sector access to private sector data must be effective, proportionate and target-driven, like the provisions on emergencies in the Data Act (see also the box on the Data Act on p. 16).

Within the guidance of the Data Act, data generated by private sector enterprises as part of their business activities can be used for public benefit purposes in order to promote innovation and reduce red tape when discharging statutory functions. In this context, regulations on issues of disclosure duties, protection of business secrets, investment protection, liability and remuneration must be considered. In accordance with the moratorium on additional burdens for the German economy adopted by the coalition government, protection of investments and a lack of red tape must be observed irrespective of the company's legal form and without regard to the company size. Whether and when this data is made available as open data after its collection is determined by the applicable open data regulations. The commercial use of data generated by consumers must be equitably balanced with their need for protection.

EU legal acts with data relevance

The **EU Regulation on European Data Governance** – abbreviated to **Data Governance Act** (DGA) – envisages that specific protected data of public bodies may be made more accessible and usable to the general public above and beyond the open data regulations and may be exchanged across borders within the EU for universal benefit. What is more, it creates legal bases for data transmission services, through which data is to be made accessible and usable in a secure processing environment. The DGA includes provisions for data altruism organizations, which facilitate data use for public policy purposes. The public policy purposes expressly mentioned are healthcare, combating climate change, improving mobility, facilitating the development, production and distribution of official statistics or improving the provision of public services, government decision-making processes or scientific research in the general interest.

The **Directive on open data and the re-use of public sector information** – abbreviated to **Open Data Directive** – of 2019 governs the re-use of public sector information for commercial and non-commercial purposes. Besides what are referred to as high-value datasets (HVD), this also includes research data. The Implementing Regulation on high-value datasets (2023/138), adopted based on the Open Data Directive, requires that selected datasets of particular significance for industry and society from the geospatial, earth observation and environment, meteorology, statistics, companies and company ownership and mobility categories be provided for re-use free-of-charge in a machine-readable format, via an API and, where appropriate, as a bulk download.

The **Digital Services Act** (DSA) is an EU Regulation that governs platforms in particular and aims to create a safe, secure and responsible online environment. To this end, the Regulation includes rules on removal of illegal content and bans on using data for advertising and the protection of minors. Very large online platforms and search engines are subject to special due diligence obligations, such as risk analysis and risk mitigation obligations. In particular, the Act envisages that the research community is to be able to access the core data of larger platforms and search engines to monitor the spread of online risks. In addition, combating of illegal contents on platforms is to improve.

The **EU Regulation on harmonised rules on fair access to and use of data** (**Data Act**) will contain entitlements to data access, in particular data generated during use in the IoT segment. There will be provision entitlements of users vis-à-vis companies (B2C) and companies vis-à-vis companies (B2B) and public sector bodies vis-à-vis companies (B2G). Data holders must provide data to public sector bodies if there is an exceptional need, such as in the case of a public emergency (for example in a natural disaster). While the Data Governance Act creates the processes and structures for the provision and exchange of data, the Data Act governs who can use data under specific conditions.

The **Digital Markets Act** (DMA) is an EU Regulation that supplements competition law and is intended to restrict the power of digital corporations that are dominant on the market. The Regulation also envisages access and transparency obligations for companies dominant on their markets (gatekeepers). Accordingly, gatekeepers are obliged to give end users access to their activity data in real time, among other things. Gatekeepers are also prohibited from combining end user data collected across multiple platform services without their explicit consent.

The **European Health Data Space Act** (**EHDS**) will regulate the use of data in the healthcare sector (primary use) and the secondary use of data to promote public health, research, innovation and development of healthcare, and put in place the conditions for establishing a cross-border health data infrastructure.

1.1.5 We are opening data for research

We will simplify and improve data access for public and private sector research as well as linking of different datasets with a Research Data Act. In this context, we will also use the potential of data protection within the framework afforded by constitutional law and Union law to create scope that benefits research.

We will regulate sector specifics separately while also ambitiously implementing European requirements derived from sector-specific legislation such as the EHDS.

In this context, we will facilitate the use of pseudonymized data in protected processing environments and sharing of anonymized and non-personal data in the public interest. To this end, we are ensuring effective protection of intellectual property and business secrets as required under the rule of law, implementing appropriate liability regimes and establishing equitable compensation schemes.

1.2 Private individuals and companies

1.2.1 We are supporting the sharing of model contract clauses and best practices

We will make it easier to find model contracts and model contract clauses as well as best practices for legally watertight application of regulations, especially on data protection and IT security. In this way, we are helping companies, civil society, public bodies and consumers securely share data and reduce transaction costs.

We will work to ensure that not only data can be shared, but also methods and processes for data handling and data access (best practices).

We will also support the development, distribution and usability of model contract clauses, model contracts and model general terms and conditions at EU level.

1.2.2 We are putting in place a framework for more investments in the data economy

The generation, processing, finishing, quality assurance and provision of data incur costs, as do the management of access rights and monitoring of compliance with the use conditions, technical infrastructures and also contract management for non-open data. These investments must also be worthwhile and easy to budget for companies (including municipal companies). That is why we are going to make Germany an attractive location for investments in the data economy.

1.2.3 We are using competition law to make data sharing easier

Access to and shared use of data can leverage significant potential to promote competition and develop new products, services and business models, including those of public sector companies. We are helping companies to implement data cooperation projects that comply with antitrust legislation, in cases where there is a significant legal and economic interest. The national and European competition law framework must promote and guarantee data access, data portability and interoperability.

We will evaluate the new regulations and develop the national legal framework if necessary. The competition authorities are also assessing the extent to which closed data silos are created on data markets and the extent to which they impair data sharing and competition on data markets.

1.2.4 We are making data easier to access in the healthcare sector

We are setting ourselves the goal of improving the availability and usability of health and care data to support individual provision of healthcare and research. We are also making government healthcare data and statistics easier to find centrally (in the epidemiology sector, among others) and increasing the availability of health-related open public authority data (open government data).

In particular in the healthcare sector, we are learning from the successful approaches to leveraging the scope afforded by the GDPR in other EU Member States like Austria or the Scandinavian countries. We want to make it easier to use data and findings generated from data processing (for example via opt-out approaches), improving provision of healthcare and health in this way. Data should also be accessible and linkable in a standardized way for research purposes or specific secondary uses.

In this way, we want to improve the use of health data with a decentralized, highly interconnected research data infrastructure for health data on the basis of and within the scope of the European Health Data Space (EHDS). Among other things, we are building on the experience of the Medical Informatics Initiative, the University Medicine Network, the Medical Register and the Health Research Data Centre at the Federal Institute for Drugs and Medical Devices (BfArM).

We are lowering access barriers to enhance the distribution and use of the electronic patient record (ePA) by introducing a graduated opt-out procedure to improve provision of healthcare for insured persons with better data availability.

Data is also to be accessible and linkable in a standardized way, including for retrospective use for the purpose of developing the provision of healthcare or promoting public health. Among other schemes, we are already working on this as part of the Federal Government's Digital Strategy and the Healthcare Digitalization Strategy.

In order to improve the options for linking data in the healthcare system, we are aiming to introduce a sector-specific research pseudonym for the health sector.

We want uniform standards for data processing in the health research sector, and will therefore put in place uniform federal regulations for the use of health data in this area in the health data use act. We are supporting the federal states' standardization efforts in the federal state hospital regulations.

In addition, we are supporting uniform regulations for the use of health data at European level, for example by establishing the European Health Data Space.

1.2.5 We are simplifying data protection and facilitating practical implementation

We want to reach an innovative, user-friendly and responsible data use that is outcome-driven and ensures the protection of personal data. Among other ways, we will achieve this with facilitating data protection and consistent implementation of technical and organizational data security. Means to achieve this goal can include anonymization or pseudonymization of data, the use of data spaces and data trustees.

Furthermore, we want to make data protection simpler, more coherent and more practicable. To this end, we will improve the options for homogeneous application and enforcement of data protection law, nationally and at EU level.

This includes our strengthening of the conference of independent data protection authorities of the Federation and of the federal states as a coordinating body by institutionalizing it in the Federal Data Protection Act. At the same time, we welcome the measures already taken by the Data Protection Conference to make the cooperation between data protection supervisory authorities of the Federation and the federal states more binding and effective, among other things by majority resolutions.

In order to further improve the uniformity of data protection supervision, we will scrutinize the feasibility of a sole competent data protection supervisory authority for nationwide data processing under joint responsibility. We will also want to make this option available for multi-federal-state research projects under joint responsibility.

We are also assessing additional potential solutions with the objective of bringing about a more uniform application of the law and enforcement of data protection.

Facilitating data protection makes use of the scope and flexibility clauses of GDPR to enhance data processing, for example via provision of information through technical and organizational guidance, via tools such as data protection cockpits, data trustees, Personal Information Management Systems (PIMS), by implementing principles like ‘Privacy by Default and by Design’ or via controlled ‘Experimental and Opportunity Spaces’ to develop good practices such as AI sandboxes in particular. Privacy-Enhancing Technologies (PETs), including anonymization, pseudonymization, masking, zero-knowledge encryption and data synthetization, which we are promoting with research projects, will also play an important role. We are assessing the establishment of a sandbox for testing PETs and the extent to which they can be incorporated in the federal architecture guidelines.

We will also make use of the flexibility clauses of GDPR to create legal clarity for employers and employees with a modern, practicable Employee Data Protection Act, and effectively protect the privacy and personal dignity rights of the employees.

Facilitating data protection can be driven forward within the framework of what is permissible under Union law with legal authorization conditions, regulatory examples and clarifications as well as opt-out approaches or, in the academic research sector, via broad consent.

We are working to ensure that companies, researchers and civil society stakeholders receive practical assistance for implementation of statutory data privacy standards and can use them, and we are encouraging companies to make more use of the support offered by data protection supervisory authorities and company data protection officers in order to further increase the legal certainty when processing data.

1.2.6 We are creating data security with security by design

As data security by design has proven itself when designing technical systems that securely process data, we are championing data security by design universally, across all technologies and applications, from product development until the end of a product's life cycle. Security aspects are to be taken into consideration as early as the design process. In particular for government products and services, we will implement this as a voluntary commitment and assess regulations for the private sector. Instruments used will include the federal IT architecture guidelines.

We will support the reliability, comprehensibility and transparency of data processing with audits and conformity attestations that are easy to implement, yet reliable, with a low degree of risk. We will support trust in the fulfilment of minimum security requirements in data processes with easy-to-understand, recognized labels and automated security updates. Neutral, private or government auditing organizations can make a useful contribution with their international experience and expertise in relevant standards.

2 BETTER DATA

2.1
We are relying on standardized data descriptions (data labelling by design and by default)

For regulated data exchange, data providers and data users expect that digital interactions can be completed without complex loopbacks or multiple confirmations. We want data labelling by design (labels are integrated in the data as standard) and by default (the standard settings require labels). For universal processing of standardized data descriptions, they are being driven forward in the measures of the Federal Government and in conjunction with current legislative proposals at national and EU level.

In standardizing the data description, we will refer to the degree of processing of the data (raw data, processed data, curated data), the conditions of data collection (context) to identify distortion (bias), the completeness of data (i. e. missing values) and the granularity of the data (level of detail and frequency of availability). In this context, we are taking into account international standardization and the experiences various stakeholders have had with existing approaches.

In order to obtain and manage consent to decisions on the use of cookies and similar tracking technologies online, we are drawing up a proposal for a regulation for a user-friendly consent procedure in conformity with competition laws. By integrating recognized services to manage their decisions on the use of cookies, users are to be given an instrument that enables them to view, trace and modify their decisions. If it becomes apparent that this pilot project for cookies and similar tracking technologies works, we will extend the procedure to other areas (e. g. Internet of Things [IoT]).

2.2
We are supporting competitiveness with interoperable standards and quality assurance

High data quality is a success factor for data use in all areas. We are supporting open specifications and the application of international standards as a key requirement for a broad-based use of data as well as for dealing with data and technologies, such as AI or distributed ledger technology (DLT). Standards create transparency, set application-specific conditions for data quality, operationalize practical implementation of statutory requirements and ensure interoperability and portability between digital applications. We are working to ensure that standards can be provided in digital, machine-readable and practicable formats, in order to facilitate process chains without media discontinuities and with seamless data integration. Furthermore, we are committed to ensuring that representatives from Germany are involved in the relevant bodies that design these standards.

Specifications and standards may not be misused for forcing players out of the market. They must be designed openly so that data markets are marked by competition and Germany remains competitive and does not fall behind internationally. At international level, we are working to close the digital gap (infrastructure, access, evaluation, data inequality). Developing countries must have a fair chance to generate their own data, use available data and create value from data.

2.3
We are supporting the development of standards for trust-based handling of data

We are supporting the development of standardized data, quality and test standards to create transparent, easy to market instruments that strengthen responsible data use and trust in the data economy in addition to the interoperability of digital applications. Standards are a basis for a data ecosystem that can be used in the long term and are therefore a core component of international measures and promotion schemes. To this end, we will first assess the data standards already available, both horizontally and vertically in the individual sectors. Building on this, we will call for the establishment of missing standards at international level and in the international standardization organizations. Where there are multiple standards in practical use (internationally) for the same use case, we will assess whether these standards are interoperable.

Key issues in this area are in particular the scaling of existing approaches and the creation of broad acceptance for trustworthy and legally watertight sharing of data in all sectors.

We want to accelerate the development of technologies and standards for legally watertight anonymization and pseudonymization. We are calling on the European Data Protection Board (EDPB) to submit guidelines on legally watertight anonymization. They are to include both the process of anonymization itself and the steps the relevant stakeholders are to comply with subsequently to ensure that anonymized data cannot be re-identified in the long term. We are also supporting the recognition of a code of conduct for pseudonymization and anonymization applicable EU-wide.

We will include the criminal liability of unlawful de-anonymization enshrined in the coalition agreement in legislation.

2.4
We are increasing the quality of government data

We will publish high-quality datasets from the mobility, statistics, earth observation and environment, agriculture and food, meteorology, geodata, company and company ownership data sectors in a standardized, quality-assured and machine-readable form with metadata in the European framework. In order to avoid redundancies, we will rely on existing infrastructure for data publication or take interoperability into consideration where this is legally and practically possible.

We will make open public sector data usable in high quality as standard. That means that the data is complete, up-to-date, transparent and clearly described as well as regularly maintained.

We see data as a key resource in the increasing use of automated decision-making systems. We consider it important to identify and prevent discrimination at the data level, e. g. via test mechanisms or data literacy building measures. Another quality feature of the data we publish will be the inclusion of gender and age-specific data.

Where additional statutory bases are required for this, we will take the necessary steps to set this process in motion.

3 DATA USE AND DATA CULTURE

3.1
The added value of data use

3.1.1 We are focusing on data-based government action at all administrative levels

We want our government action to be up-to-date, citizen-friendly and transparent. To this end, we are consistently focusing on data-driven government action at all administrative levels. To use data effectively, we need data and IT infrastructure solutions for the interfaces between the public administration and both citizens and businesses. This calls for taxonomy-based standardization of data use and infrastructure sharing.

To this end, we want to create sustainable organizational structures and processes as well as common tools. In particular, these include:

Organizational structures: efficient, permanently established, interconnected data laboratories in the federal ministries including chief data officers and/or chief data scientists as well as open data coordinators.

In the area of open administrative data, more datasets are to become available and linkable in the long term (Linked Open Data).

Processes: the use of common data classification criteria (data taxonomy), checks whether required data is available in sufficient quality, especially for the preparation of Federal Government regulatory projects (data checks), agility and a positive error culture support the necessary continuous adaptation of (digital) administrative processes.

Tools: Use of an overview of the data available in public authorities (linking data atlases or data maps, administrative data information platform, register map, etc.) and their sources, provision of up-to-date IT infrastructure shared across multiple authorities and data trustee models.

3.1.2 We are supporting the sharing of sample use cases

To enable the different stakeholders to familiarize themselves with data-based solutions, possible sample use cases must be published. In this way, successful concepts and sample use cases can have a broader impact. This will contribute to reducing uncertainties and enable stakeholders to tap the opportunities arising from the use of data under equal competitive conditions ('level playing field').

We as the government are leading by example in this field and are sharing best practices at research and development levels by way of networking events, accompanying communication measures and research activities.

3.1.3 We are simplifying the use of unstructured data for Large Language Models (LLMs) by the public sector

We want to enable also the public sector to harness the opportunities of artificial intelligence. For this purpose, we are examining whether and to what extent LLMs could be used by the public sector in a meaningful way in compliance with data protection requirements.

We will identify what is needed and required to use LLMs in the public sector (for example, with regard to data security, data protection, digital sovereignty and open source applications) to harness the potential of this technology for the public sector. For this task, we will be supported by the Advisory Centre for Artificial Intelligence in Public Administration (BeKI) and the Algorithm Evaluation Centre for Authorities and Organisations with Security and Safety Tasks (ABOS), among others, as well as the Federal Commissioner for Data Protection and Freedom of Information (BfDI) and the data laboratories.

LLMs are trained initially on a very large amount of data ('pre-training'). Usually, this data is available in unstructured form only. To use LLMs internally or train them for specific use in the public sector, data silos must be broken down. Using Privacy-Enhancing Technologies (PETs) is to become possible so that the data can also be used for training purposes.

3.2 Data networking and data exchange

3.2.1 We are supporting data spaces and their networking as a data use instrument

For a secure exchange of data, we are driving forward the development of sectoral data spaces (for example, Mobility Data Space, Mobilithek, Culture Data Space, Catena-X, Manufacturing-X, EHDS, NFDI, European Open Science Cloud (EOSC), Agricultural Data Space, European Green Deal Data Space, etc.) and their dynamic networking.

We want to improve the interoperability of secure and self-contained data spaces with research data infrastructures, open data portals and data infrastructures from specific fields (e. g. geospatial, environmental, quality infrastructure, biomedical/health, agricultural or mobility data) in Germany, Europe and beyond.

Data spaces enable the networking of decentralized data assets in compliance with data protection laws to protect the rights and interests of the data suppliers. To tap the potential inherent in the cross-sectoral availability of data and the resulting possible use cases for society, the economy and public benefit purposes, it is also important to make the currently developing sector-specific data spaces interoperable and network them in a federated data ecosystem. With Gaia-X, there is already an open framework for building a decentralized, sovereign European data infrastructure. This system enables a broad range of stakeholders to use existing European services by means of common rules, open source code and interoperability standards. Taking into account the results of Gaia-X and other national and European initiatives, the research and development project National Initiative for AI-based Transformation to the Data Economy (NITD) is developing the organizational and technological foundations for the provision and use of data across data spaces as well as its pilot implementation in practice.

By networking the relevant stakeholders, the resulting data ecosystem will lay the foundations for the development of competitive and scalable data and AI applications.

The National Research Data Infrastructure (NFDI) and its effective European and international networking (e. g. through connection to the European Open Science Cloud (EOSC)) are characterized by their cross-disciplinary approach and ensure that Germany remains internationally attractive and competitive as a research location in the global science market.

Quality infrastructure (QI) cloud solutions based on data spaces allow the creation of additional value through ongoing digitalization of processes, for example the assurance of product and service quality. A QI ecosystem for public administration stakeholders, testing facilities, companies and standardization organizations that is based on norms, open standards and free licences is promoting the digital transformation and innovation in these areas, some of which are highly regulated.

By establishing Platform Analysis and Information Systems (PLAIN) in the Federal Government's Digital Strategy, we have created a standard for the sovereign and protected processing of big data problems in the Federal Government, with the goal of improving the information on which political decisions are based.

3.2.2 We are founding a data institute

It was decided in the coalition agreement to found a data institute to drive forward the availability and standardization of data and establish data trustee models and licences.

The Federal Government has taken up the recommendations of an independent founding commission for the development and operationalization of the data institute, with the aim of improving the availability and usability of data for society as a whole within the existing legal framework in Germany. In cooperation with existing partners, the data institute is to form a powerful national actor that pools know-how, provides targeted assistance – especially in intersectoral exchange – and takes into account the special importance of research data. The data protection authorities are to be closely involved in the activities of the data institute.

A demand-driven, agile process is to enable the data institute to begin its work soon. The first pilot projects (use cases) implemented are to demonstrate the challenges of data access and data sharing, data availability and use as well as data standardization. These findings are to help derive the specific tasks, the required expertise and the structure of the data institute.

3.2.3 We are committed to a reliable exchange of data across borders (Data Free Flow with Trust)

We will work towards an international framework that is conducive to the transfer and sharing of data across borders while respecting the values and principles of the EU. This framework is to enable cross-border data exchange for citizens, civil society, industry, the scientific community and government while safeguarding the interests of stakeholders. In particular, we will advocate for the creation of European and international data spaces. We are supporting the Data Free Flow with Trust concept suggested by the G7 and G20 – also beyond economic data. In particular, we want to facilitate joint research projects that use artificial intelligence. In this context, we are advocating for legally watertight regulations on the basis of the EU-US Data Privacy Framework.

3.3 Acceptance of data use

3.3.1 We are strengthening the resource-efficient use of data and its use for sustainability goals

The Data Strategy is intended to promote resource-efficient use of data. We are supporting solutions that focus on long-term, sustainable data management to establish economical processes across all sectors. We are also supporting digital data collections that contribute to achieving the sustainable development goals of the 2030 Agenda for Sustainable Development. This can be the evaluation of services, processes or products by means of life cycle assessment data, the recording of the state, modification and forecast of our ecosystems or the transformation of our economy to a circular economy by means of product data management. We also want to open up scope for opportunities and experimentation, strengthen research and promote initiatives.

3.3.2 We are ensuring data-driven innovation while protecting intellectual property and business secrets

We want to make sure that a modern data economy develops.

Within the framework of the statutory provisions, rights holders in works that are part of data assets must have the certainty of being able to control the use of their self-produced works and being in principle adequately remunerated for making them available. In a modern data economy, integrating intellectual property into the sharing and exchange of data in a regulated way can add value and boost innovation. We want to create incentives to better harness this potential for additional value creation from intellectual property.

3.4 Comprehensive data literacy

Data and information are the basis of our knowledge society and economic system. Literacy in these areas is a prerequisite for the future viability of every individual, society and government in times of digital transformation and exponential data growth.

Providing Germany's citizens with the necessary skills is indispensable for this transformation. Our goal is for almost everyone to be able to judge the use of data confidently and for as many people as possible to use data competently and innovatively by themselves. Transparent data use, with an added value that is comprehensible for everyone, can contribute to this. We will create suitable framework conditions for a user-friendly and participatory design of the technical tools with regard to data availability and data use.

We are also supporting the development of the breadth and depth of people's data literacy. This also includes accessibility of data, especially for people with disabilities.

Data literacy is the ability to identify data sources, collect and organize data, understand, analyse and interpret data, and present and communicate data in an understandable way. This includes understanding data structures and data formats. Data literacy also covers a basic understanding of the right to informational self-determination and applicable data protection rules.

Data literacy allows for critical thinking and reasoning. It helps to recognize patterns, trends and contexts and make evidence-based decisions. Recognizing uncertainties, prejudices, misinformation and disinformation is also part of data literacy.

Using data-based procedures, such as visualizations, analyses and algorithmic systems, requires technical and professional skills. The choice of standards, interfaces and licences, for example, is highly relevant here. Data law, data protection, data and IT security, ethical aspects and discrimination prevention also play a major role, especially with regard to the use of data in AI systems.

Data literacy needs to be learned. It must therefore be taught in schools, universities, vocational training, (vocational) further education and lifelong learning programmes/into old age.

Introducing computer science as a subject in schools across Germany could be useful, and we are encouraging the federal states to do so. Data literacy should also play an important role in all areas of general education.

In addition, we are strengthening extra-curricular STEM education as a basis for gaining data literacy. The STEM Action Plan 2.0 pools all STEM measures along the entire education chain, from kindergarten to university.

As the Data Strategy focuses on responsible data use and the opportunities data use provides for society, data literacy also plays a crucial role in the successful implementation of the Data Strategy. With the Data Strategy, the Federal Government is supporting stakeholders along the entire data value chain.

To promote the collection, use and re-use of data by civil society and the public sector, we are creating services for building up data literacy and for supporting public-benefit focused data projects. Examples include:

- The Civic Data Lab, one of the anchor projects of the interdepartmental Civic Coding – Innovation Network AI for the Common Good initiative. The initiative aims to strengthen the data and AI literacy of civil society by pooling interdepartmental funding programmes and support measures.

- The Data Literacy Toolbox, which provides comprehensive access to learning opportunities and tools for better handling of big data for different learning levels along the entire education chain.

- The Advisory Centre for Artificial Intelligence in Public Administration (BeKI), with which we help the public sector to use AI technologies competently and initiate and support the sustainable building of data literacy in public administration.

3.5
Courageous and responsible data culture

Comprehensive digital transformation processes and new technologies, such as artificial intelligence in particular, have brought us to a new era of data use and on the way to a digital and sustainable society. While this new era raises many questions, it also permits and requires a cultural shift towards increased data provision and use, collaborative approaches for cooperation and interdisciplinary and cross-institutional exchange. In this context, the government can provide support in establishing governance structures and facilitate the user-friendly design of data-based applications.

The potential inherent in data use can only be fully tapped if citizens, civil society, industry, the scientific community and government recognize the benefits of data use and develop trust in the responsible use of data. For this purpose, data and consumer protection must be enforced to enable everyone to explore the digital world independently, self-determinedly and safely and securely.

To enable people to recognize the benefits of data use, these benefits must be made visible. All stakeholders can contribute to this. Not every potential abuse and not every conceivable risk should be decisive in the assessment.

To realize the goal of achieving a higher level of data use, it is not sufficient to initiate a cultural shift in public authorities. The goal of increasing data use can only be achieved if all stakeholders are open-minded, become more transparent and also consider data use in the processes/systems when collecting or generating data.

A courageous and responsible data culture is important in all parts of society. In addition to government agencies, civil society groups, commercial enterprises and research institutions also have an important role to play in supporting the voluntary sharing of data above and beyond data sharing obligations. Individuals should be also empowered to share their data in the public interest in a self-determined way. We are supporting the active sharing of data by individuals or institutions (data altruism). In this context, it is important to recognize and leverage the opportunities and potential of one's own data while handling it more responsibly and mindfully in everyday transactions (e.g. in apps).

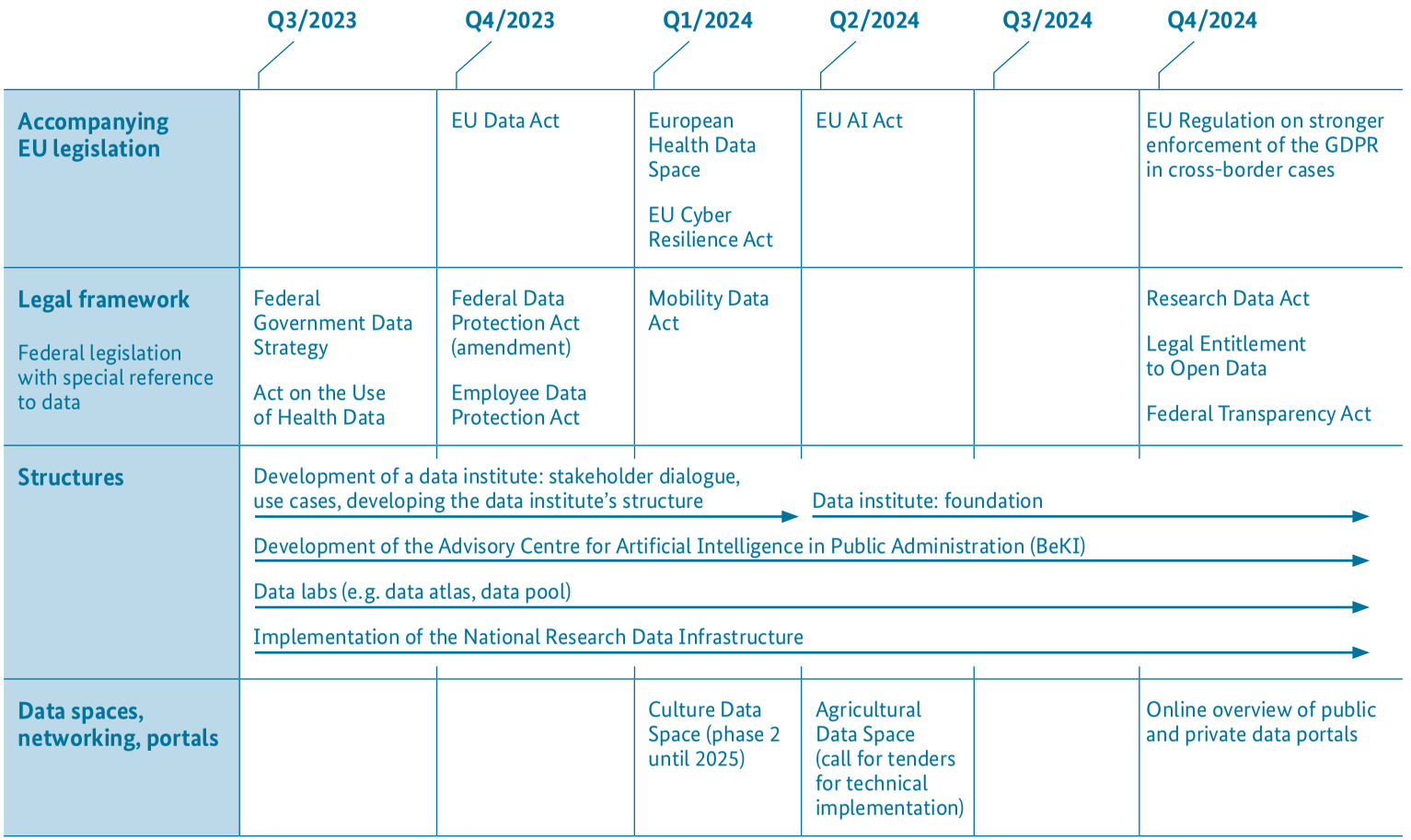
A new data culture will only emerge through joint efforts by the citizens, civil society, industry, the scientific community and government. This is a matter for society as a whole that we have to tackle together. Only in this way can the conflicting goals mentioned in the preface be overcome. We want to join forces to make the responsible and effective generation, collection, use and management of data second nature in Germany.

4 IMPLEMENTATION

Data Strategy Roadmap for the period to Q4/2024

Roadmap of EU and federal legislation (cabinet decision) with special reference to data including important projects of the coalition agreement

Measures of the Data Strategy adopted by the Federal Government that have an impact on the budget will be implemented by each ministry on its own responsibility and within the framework of the funds available for the respective purpose in the budget and financial planning.



Already completed
Important EU legal acts with special data relevance:

2016: GDPR
2019: Directive on open data and the reuse of public-sector information
2022: Digital Services Act
2022: Digital Markets Act
2022: Data Governance Act

Due to the complex national and international procedures, the roadmap can only present estimates. This means that there may be delays in the schedule. The roadmap will be updated regularly.

Publication data

Published by:
Federal Ministry for Digital and Transport
Invalidenstr. 44
10115 Berlin
Internet: www.bmdv.bund.de
E-Mail: Poststelle@bmdv.bund.de

Federal Ministry for Economic Affairs and Climate Action
Scharnhorststr. 34-37
10115 Berlin
Internet: www.bmwk.de
E-Mail: Poststelle@bmwk.bund.de

Federal Ministry of the Interior and Community
Executive Staff; Planning and
Communication Directorate-General
Alt-Moabit 140
10557 Berlin
Internet: www.bmi.bund.de
E-mail: presse@bmi.bund.de

As at:
August 2023

Designed by
MedienMélange: Kommunikation!
Goetheallee 6, 22765 Hamburg
www.medienmelange.de

www.bundesregierung.de

facebook.com/Bundesregierung
twitter.com/RegSprecher
youtube.com/bundesregierung
instagram.com/bundeskanzler