

ESMA Data Strategy 2023-2028



Table of Contents

1	Executive Summary	3
2	Drivers, stakeholder expectations, and current situation	4
2.1	Background	4
2.2	Driving factors for ESMA Data Strategy 2023-2028	5
2.2.1	New or evolving ESMA mandates and activities:	5
2.2.2	Technological developments/digitalisation:	5
2.3	Stakeholders' expectations	6
2.4	Current ESMA Data Situation	7
3	Data Vision and Strategy	7
3.1	Data vision	7
3.2	Data strategic objectives	8
3.2.1	Objective A – Enhanced data hub: enhance ESMA's role as data hub, focusing on improved data, information accessibility, interoperability and usability, along with data harmonisation and standardisation.	8
3.2.2	Objective B – Access to data of public interest: contribute to providing relevant, useful and understandable information to the market in machine-readable form, and facilitate its use, including by retail investors.	9
3.2.3	Objective C – Data-driven supervision: enable cutting-edge, smart and effective data-driven supervision	10
3.2.4	Objective D – Thought leadership: pursue thought leadership and collaboration on data standards, technologies, and reporting innovations	11
3.2.5	Objective E – Efficient data policy: promote efficiency, transparency, and cooperation in data policy, and reduce reporting burdens	12
3.2.6	Objective F – Systematic data use: systematically use data for evidence-based policy development, supervision and risk assessment	13
4	Target capabilities	15
5	Implementation plan	18
5.1	Roadmap	18
5.1.1	Objective A – Enhanced data hub	18

5.1.2	Objective B – Access to data of public interest	18
5.1.3	Objective C – Data-driven supervision.....	19
5.1.4	Objective D – Thought leadership	19
5.1.5	Objective E – Efficient data policy	20
5.1.6	Objective F – Systematic data use	20
5.2	Resources.....	20
6	Communication plan.....	23
7	Annexes	24
7.1	Annex I – Acronyms and definitions.....	24

1 Executive Summary

Reasons for publication

This document supplements the ESMA Strategy 2023-2028 by elaborating on how ESMA data assets will be mobilised to best serve and help deliver on its strategic and thematic objectives in the period covered.

Contents

The ESMA Data Strategy 2023-2028 first lays out the internal and external motivations for this report, which serve as a baseline for subsequently articulating ESMA's data vision and objectives. It then provides details on how ESMA intends to scale up its data capabilities to ensure the successful delivery of the listed projects and tasks within the proposed timeline, and to achieve its strategic objectives:

- enhance ESMA's role as data hub, focusing on improved data, information accessibility, interoperability and usability, along with data harmonisation and standardisation;
- contribute to providing relevant, useful and understandable information to the market in machine-readable form, and facilitate its use, including by retail investors;
- enable cutting-edge, smart and effective data-driven supervision;
- pursue thought leadership and collaboration on data standards, technologies, and reporting innovations;
- promote efficiency, transparency, and cooperation in data policy, and reduce reporting burdens; and
- systematically use data for evidence-based policy development, supervision and risk assessment.

Next Steps

ESMA has already started with the scaling up of its data capabilities and the implementation of some of the first planned deliverables, and it will continue to do so. This document, and in particular the implementation plan section, is meant to be revisited over time as new legislative, technological or any relevant types of development emerge that need to be addressed in a way that would require adjustments to the ESMA Data Strategy 2023-2028.

2 Drivers, stakeholder expectations, and current situation

2.1 Background

The mission of the European Securities Market Authority (ESMA) is to enhance investor protection, promote orderly financial markets and safeguard financial stability. Established in January 2011 as part of the European System of Financial Supervision (ESFS), ESMA's central objectives are achieved by assessing risks to investors, markets and financial stability; completing a single rulebook for EU financial markets; promoting supervisory convergence; and directly supervising several key actors in financial markets.

In October 2022 ESMA announced its new strategy for 2023-2028, which focuses on three strategic priorities (fostering effective markets and financial stability, strengthening the supervision of EU financial markets and enhancing the protection of retail investors) and two thematic drivers (enabling sustainable finance and facilitating technological innovation and effective use of the data).



The existing ESMA Data Strategy, which was developed in 2017, was instrumental to establish ESMA as an Authority that relies regularly on the use of evidence-based analyses; that is a key and reliable data hub for its stakeholders; that facilitates data access, distribution, and analysis for NCAs achieving EU-wide synergies; and that is the reference for EU financial markets data regulation and the development and adoption of international data standards.

It now needs to be reevaluated to account for the many changes that occurred between 2017 and 2023, as well as to tackle the new challenges ahead, in order to best support the ESMA Strategy 2023-2028.

2.2 Driving factors for ESMA Data Strategy 2023-2028

ESMA has become an established regulator and supervisor playing a significant role in the well-functioning of the EEA financial markets. Meanwhile, the strategic role and importance of data has evolved significantly since 2017, as new legal mandates, new regulatory frameworks and new technologies – such as Big Data and Artificial Intelligence (AI) technologies – increase the ability for authorities to obtain better and broader insights. These developments result in increased expectations from various stakeholders and ESMA must strive to meet them. To keep up with a fast-changing landscape and uphold the appropriate level of regulatory and supervisory oversight, ESMA must consider the following driving factors in its new Data Strategy:

2.2.1 New or evolving ESMA mandates and activities:

- Recently established supervisory mandates (e.g., Critical Benchmark Administrators, DRSPs, Tier 2 CCPs);
- New regulatory work on Digital Finance (e.g., DORA, MiCA, DLT Pilot Regime) and Sustainable Finance (e.g., CSRD, SFDR);
- The European Commission Strategy on supervisory data in EU financial services¹;
- The further progress on the Capital Markets Union agenda, including the establishment of a European Single Access Point (ESAP).

2.2.2 Technological developments/digitalisation:

- The possibilities created and challenged posed by the rapid growth of new technologies suited for supervision, reporting, data collection or data use, and fast-evolving digital developments (e.g., in the areas of Cybersecurity, Big Data, DLT, Artificial Intelligence, RegTech / SupTech);
- The lack of sufficient resources and expertise across ESMA and NCAs to closely follow technological innovation, while grasping at the opportunities offered.

¹ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021DC0798>

2.3 Stakeholders' expectations

The ESMA Data Strategy 2023-2028 objectives consider the expectations of and added value for ESMA's stakeholders. Specifically, 25 interviews were conducted with a sample of representatives of different groups: ESMA staff, NCAs, European Bodies, International Organisations, entities supervised by ESMA and market participants. This section summarises the key findings from the interviews, as regards the stakeholders' expectations of the ESMA Data Strategy 2023-2028.

While ESMA invested a lot of effort in improving data capabilities, which is recognised by its stakeholders, some areas for improvement and expectations for the future have been voiced moving forward, which are summarised in the paragraphs below and have been considered when establishing the objectives of this new Data Strategy. In that respect, stakeholders consider that to create additional value ESMA should:

- increase the efficiency of data reporting processes by realising economies of scale in collaboration with the other ESAs and other EU bodies through **enhanced governance and cooperation**;
- further reduce the reporting burden and increase data quality by using simplified, standardised, non-duplicative, and unambiguous requirements achieving **data harmonization and a common data dictionary**;
- be a thought leader and reference point in modernising data reporting technologies, to reduce compliance cost and provide high quality data to stakeholders through the **optimisation of reporting and dissemination flows**;
- contribute further to the improvements in data quality and usability, through the promotion of **more efficient data quality responsibilities and governance**, in collaboration with NCAs and other stakeholders;
- promote effective and efficient collaboration among authorities by improved **sharing of data, information, analytical tools** (i.e., the data hub) **and expertise** to provide analytical insights. This should be done through harmonized and interoperable data sharing processes, including improved machine-to-machine interactions to reduce interpretation issues and processing costs, thus contributing to an overall reduction of reporting burdens by, e.g., reducing duplicative data requests;
- have organisation-wide capabilities for advanced **risk-based, data-driven supervision**, thereby improving the availability and quality of data for direct supervision and increased supervisory convergence;
- be a thought leader on the regulation of new financial products and services for the purpose of supervisory reporting and regulation, with an increased focus on information related to **sustainability and investor protection**; and

- increase **communication, transparency, and feedback** on why and how the data is used by the regulatory and supervisory community through, among others, closer cooperation with market participants, and other authorities
- improve access to data and analytics by market participants, by extending the scope of shared data and ensuring easy access, in particular in a machine-readable form.

2.4 Current ESMA Data Situation

An assessment of the ESMA data capabilities was performed at the end of 2021 based on an industry-leading data capability maturity framework.

Overall, the assessment concluded that:

- **Data is a strategic driver at the core of ESMA:** ESMA has developed consistent data reporting standards in all areas where ESMA has a regulatory/supervisory role and has processes in place to access and use data, create analytical content, as well as plan and recruit data talents; however,
- **ESMA is mostly on an “opportunistic²” maturity level:** initiatives and efforts to exploit data are not systematically coordinated internally, and capacity to analyse, integrate, organise and access data could be more advanced. ESMA data governance is also too complex and not clearly defined enough, which strains resources in coordination efforts.

3 Data Vision and Strategy

The below ESMA Data vision and objectives are based on the overall ESMA Strategy 2023-2028 recently published, the ESMA stakeholders’ expectations detailed in a previous section, as well as the limitations exposed during the capability assessment mentioned above.

The ESMA Data Strategy complements the ESMA’s Strategy by reflecting on the data challenges posed by the latter and translates them into a vision and actionable objectives that will help ESMA to deliver on its mission during the 2023-2028 period; focusing on its three strategic priorities and being driven by the two key themes.

3.1 Data vision

The ESMA Data vision is to “contribute to ESMA’s mission with high quality data and analytical information”

This vision has three underlying drivers:

² “Opportunistic” to be understood in this context as ESMA not having a fully consistent and coordinated approach to data initiatives across the different work areas and departments.

- **Data efficiency:** Reducing compliance costs of market participants and streamlining the use of data
- **ESMA and NCA data capabilities:** Sharing of information, knowledge and tools
- **Data at the core of ESMA:** Providing evidence based on data to support ESMA decisions and activities (including outcome-driven supervision)

A first step undertaken by ESMA that translates this vision into a tangible outcome is the creation of the Data Intelligence and Technology department. This department will cover the full set of competences related to the entire data lifecycle by the establishment of centres of excellence supporting internal and external stakeholders. The new organisation will be one of the key enablers for the implementation of ESMA's Data Strategy by setting up a close cooperation and understanding of the stakeholders needs, a comprehensive overview of data, processes and technology, efficiency, synergies, and economies of scales. The combined competences will enable exploiting the full potential of data by internal and external stakeholders.

3.2 Data strategic objectives

For the ESMA data vision to take shape, six strategic objectives have been defined. For each, a short explanation of the backdrop (where ESMA is, where it wants to be, and why) serves as an introduction to the main actions ESMA will undertake in the coming years to deliver on these objectives.

3.2.1 Objective A – Enhanced data hub: enhance ESMA's role as data hub, focusing on improved data, information accessibility, interoperability and usability, along with data harmonisation and standardisation.

As part of its previous data strategy ESMA has developed a standardised access to data for stakeholders via the creation of centralised systems providing access to supervisory data for competent authorities (e.g., the single access point to trade repositories data, transaction reporting exchange mechanism), to reference data on which market participants can rely for their day-to-day operation (e.g., the European Rating Platform, the Financial Instruments Reference Data System, the Financial Instruments Transparency System, the Prospectus database), as well as to public registers³ accessible on the ESMA website via a unified search engine.

ESMA will continue the development of the data hub to provide, in addition to data exchange functionalities, a **shared platform where stakeholders⁴ and ESMA can work on the same data in a unified manner and without duplicating it**. This will also help reducing the overall

³ E.g., the list of MiFID authorised investment firms, UCITS and AIFMD fund managers, third-country benchmarks.

⁴ Stakeholders may include regulators and supervisors at national and European level, market participants, investors and academics.

cost by avoiding building multiple systems for the same purpose and having the possibility to use the same platform for different projects, including joint projects with ESAs and NCAs.

The specific actions envisaged by ESMA are:

- Where appropriate, launch new joint projects to achieve economies of scale:
 - Explore the possibilities to **provide NCAs with common data and supervisory tools**⁵;
 - On the basis of the ongoing project to standardise order book data, develop a **mechanism facilitating order book data exchange**⁶;
 - Set up **data sharing arrangements** to facilitate data exchange between ESMA, NCAs and other authorities⁷
- Continue developing the **ESMA Data Platform based on big data technologies** towards a shared analytical platform:
 - Enable **advanced analytics** (Artificial Intelligence, Machine Learning, etc.) across datasets on the ESMA Data platform;
 - Provide high quality consolidated **master data**⁸ to NCAs;
 - Develop **dashboards** with analytical functions and ready to use data aggregates **to serve internal and external stakeholders**;
 - Allow **collaborative work with NCAs** on ESMA available regulatory and supervisory datasets (e.g., EMIR, SFTR, Securitisation, MiFID) as well as commercial ones (when contractually permitted) **via scripting, visualisation, and communication tools**.

3.2.2 Objective B – Access to data of public interest: contribute to providing relevant, useful and understandable information to the market in machine-readable form, and facilitate its use, including by retail investors.

Data, provided it is made available to the general public in a user-friendly format and an easily accessible way, can also serve ESMA's strategic priorities by enhancing transparency in the markets and reducing barriers to accessing regulatory information, thus contributing to building trust in the EU financial markets from all its participants, especially the (retail) investor

⁵ E.g., ICT-related incident data under DORA, retail investor data under existing reporting regimes, data sources to support the supervision of ESG disclosures, supervisory tools for crypto assets service providers or for exploring and analysing fund data.

⁶ As envisaged in the "Capital Markets Union: clearing, insolvency and listing package".

⁷ E.g., with central banks to contribute to developing an integrated reporting system as envisaged by the proposed AIFMD and UCITS directives.

⁸ Master data to be understood as a unique source of reference data that NCAs can reuse in their processes and analytics in order to ensure a consistent approach and comparability of results.

community. This has become all the more relevant over the last years with the European Capital Markets Union (CMU) topic back on the regulatory agenda, as well as the growing role played by retail investors in financial markets and the risks they face.

ESMA has identified several areas of action to successfully deliver on this objective for which it will **leverage on its role as a data hub** and the foreseen improved capacities listed previously:

- Ensure **public data and information on securities markets is easily accessible and usable by all its stakeholders** (e.g., via centralised data access points and interfaces suitable for manual or automated access with intuitive search functionalities), in particular via the development of **ESAP**⁹;
- Provide **access to anonymised or aggregated supervisory data**¹⁰ otherwise not publicly available to facilitate, e.g., innovation and research¹¹;
- Share **data and analytics with the market** in an intelligible, structured and user-friendly manner by leveraging the ESMA Data Platform, in order to grant retail investors easier access to public information¹².

3.2.3 Objective C – Data-driven supervision: enable cutting-edge, smart and effective data-driven supervision

Data plays an increasingly important role in ESMA's and NCAs' supervisory activities, with new types of data and growing volumes available to authorities. On the other side, innovative technologies also drive changes via new types of activities becoming subject to supervision. Therefore, ESMA will continue to use and **make available to supervisors the necessary data, information and tools** enabling data-driven supervision and harnessing the synergies between the complementary ESMA and NCAs supervisory activities.

In particular, ESMA intends to:

- Strengthen the cooperation with NCAs on the use of data in supervision, to identify and exploit synergies between the supervisory activities of ESMA and NCAs, by **sharing expertise and performing joint undertakings** benefiting ESMA and NCAs¹³;
- Ensure efficient use of modern analytical tools¹⁴, for the ESMA direct supervision mandates and by NCAs, in particular by carrying out **projects using novel technologies**¹⁵, such as artificial intelligence and web scraping, that will allow

⁹ ESMA will also contribute to this objective by facilitating the establishment of Consolidate Tape Providers.

¹⁰ E.g., EMIR, SFTR, MiFIR or CRA data.

¹¹ E.g., by contributing to the EU Digital Finance Platform Data Hub.

¹² E.g., by exposing key metrics and interactive dashboards on financial entities and instruments.

¹³ Past examples include emission allowances, trading of sanctioned instruments, Article 25 of AIFMD and ESMA assessment, short selling opinions.

¹⁴ Including monitoring dashboards, with early warning indicators, alerts, triggers and scoring to support an efficient prioritisation of supervisory and convergence work.

¹⁵ E.g., advanced machine learning, natural language processing.

gathering knowledge and experience and the **joint development of supervisory tools**¹⁶.

- **Modernise the approach to data quality** to a more automated, outcome-focused, data-driven and risk-based approach, based on EU-wide monitoring, common KPIs, benchmarks and dashboards; and **expand it to a wider range of datasets**¹⁷.
- Develop **common methodologies** and models to be applied for the detection of potential breaches and for enabling a **risk-based approach** and more targeted ESMA and NCAs supervisory actions.

3.2.4 Objective D – Thought leadership: pursue thought leadership and collaboration on data standards, technologies, and reporting innovations

Over the past years, ESMA has been a driving force to strengthen the development and consistent application of data standards, as demonstrated by its efforts to foster the use of common identifiers and standards such as the Legal Entity Identifier (LEI) and other ISO standards in reporting regimes under its remit and by other jurisdictions.

ESMA's ambition is to reinforce its role as a **global thought leader** in the development, coordination, implementation and promotion of international data standards and best practices and to **increase collaboration with NCAs and other stakeholders** to explore new ways of funding and delivering joint projects for the benefit of the whole community.¹⁸

At the same time, the advent of new technologies has created new opportunities in the regulatory and supervisory landscape that ESMA seeks to seize by building up subject matter expertise in order to position itself as a **reference point on RegTech / SupTech for NCAs**, international regulatory and supervisory authorities, and the broader financial sector.

To deliver on the above, ESMA plans the following:

- **Facilitate pilot projects and experimental environments** in order to increase collaboration with and between NCAs and to develop joint expertise around new data-related technologies such as RegTech / SupTech and facilitate their adoption in day-to-day processes;¹⁹

¹⁶ E.g., existing projects on the use of artificial intelligence in market abuse monitoring (to be run with a group of 16 NCAs) and on the detection of anomalies in CRA data.

¹⁷ Such an approach is already successfully applied to the work on EMIR and SFTR data quality.

¹⁸ E.g., projects developed by a few NCAs and made available to all NCAs under joint governance.

¹⁹ E.g., the study on how on-chain data can be retrieved that is conducted by ESMA in the context of the DLT Pilot regime.

- **Explore the use of alternative funding sources** beyond ESMA and NCAs budget²⁰ for projects related to data and new technologies, performed jointly by ESAs²¹, by ESMA and NCAs, or by NCAs, and coordinated by ESMA, for the benefit of all NCAs.
- **Increase collaboration with public society** to gather feedback, gain new insights and strengthen expertise on data hosted by ESMA, and on new technologies and standards related to data. This should be performed via, e.g., research projects in partnership with academics, and by organising conferences or workshops to develop concrete proposals with field experts;
- **Strengthen ESMA's participation in international groups and governance bodies** relevant for data-related standards and technologies, while increasing knowledge and practice sharing;
- Extend the use of **data and information sharing with other European authorities and institutions** such as the ECB, the ESRB, other ESA, ACER and other authorities whose remit intersects with ours;²²
- **Intensify collaboration with the ECB on common data initiatives**, particularly on SupTech, RegTech and technological innovation, including in the area of crypto assets²³.

3.2.5 Objective E – Efficient data policy: promote efficiency, transparency, and cooperation in data policy, and reduce reporting burdens

Since its establishment, ESMA has focused on the developments related to supervisory data standardisation and has driven the evolution of reporting standards from divergent national reporting regimes to standardised across the EU reporting regimes based on common standards and messaging. Nevertheless, as evidenced by the fitness check of supervisory reporting requirements in EU financial services legislation²⁴, there is a need for further harmonisation of reporting regimes, reducing of the compliance cost by the reporting entities, increased cooperation between the authorities and increased use of modern technologies to support the reporting process.

Therefore, based on the experience gathered so far, ESMA intends to actively contribute to / be the driving force behind the implementation of the **EC supervisory data strategy**, by fulfilling the mandates envisaged in the strategy and influencing the way the strategy is implemented through ambitious proposals in the specific contributions required by ESMA.

²⁰ E.g., the Digital Europe Programme used for the ESAP development, or the Technical Support Instrument Programme (TSI) used for the project on the use of AI in market abuse.

²¹ Examples include common projects with a lead authority (e.g., DORA, MICA, ESAP) and joint procurement procedures as is already done in several areas such as for IT solutions and commercial data providers.

²² E.g., ESMA already uses ECB registers for its report on Trends, Risks and Vulnerabilities and plans to extend its use of ECB data with RIAD data.

²³ E.g., to study the underpinning technology and assess its reusability in other applications in the context of MICA or DLT Pilot.

²⁴ [Results of the fitness check of supervisory reporting requirements in EU financial services legislation \(europa.eu\)](#).

The key focus will be to **reduce the compliance burden** for reporting entities by reducing duplicative and inconsistent requirements, optimisation of reporting flows, ensuring effective and efficient data sharing between authorities, and exploiting emerging technologies.

ESMA also intends to increase **communication, transparency, and feedback** on data, including closer cooperation with market participants and other authorities, in particular the other ESAs, ECB and others to find ways to increase efficiency.

The specific actions envisaged by ESMA will include:

- Progressively, by taking the opportunity of the review of reporting regimes, **increase standardisation and efficiency of reporting**, by ensuring the alignment of standards, streamlining the reporting arrangements and mechanisms to collect information²⁵, promoting data sharing and avoiding duplicative reporting obligations²⁶, as well as increase the usability of public information by **ensuring machine-readability of the disclosed data**;
- Explore the possibility of the use of new technologies to ensure **cost-efficient reporting** solutions and the opportunities to **reduce reporting burden**, in particular, study further the possibility to implement the concept of the machine-readable and executable reporting²⁷;
- Develop a **data dictionary for securities markets**²⁸, as envisaged by the European Commission Strategy on supervisory data in EU financial services to contribute to the consistency across reporting regimes.

3.2.6 Objective F – Systematic data use: systematically use data for evidence-based policy development, supervision and risk assessment

Historically, technology and data have always served as enablers to achieve ESMA's mission. In its new strategy, ESMA has the ambition of taking a more systematic data-driven approach to regulation, supervision and risk assessment to deliver across its strategic priorities: to foster effective markets and financial stability, strengthen the supervision of EU financial markets, and enhance the protection of retail investors.

To achieve this, ESMA aims to **reinforce its ability to rely on and make use of data** to support its regulatory and supervisory work, help monitor supervisory convergence, detect financial stability threats, and address investor protection concerns. All of that through the **promotion of evidence-based ways of working**, both internally and in cooperation with NCAs.

²⁵ E.g., pull and push approaches.

²⁶ E.g., in the AIFMD and UCITS reporting frameworks where the legislative proposals envisage that ESMA should study integrated supervisory data collection.

²⁷ E.g., in the AIFMD and UCITS reporting regimes that are currently under review.

²⁸ This dictionary will contain a description of the content and format of all data collected under various reporting frameworks in a structured, comprehensive, consistent and unambiguous manner, using terms anchored in legislation to establish a clear link between collected data items and the relevant legislative requirements.

The planned actions are:

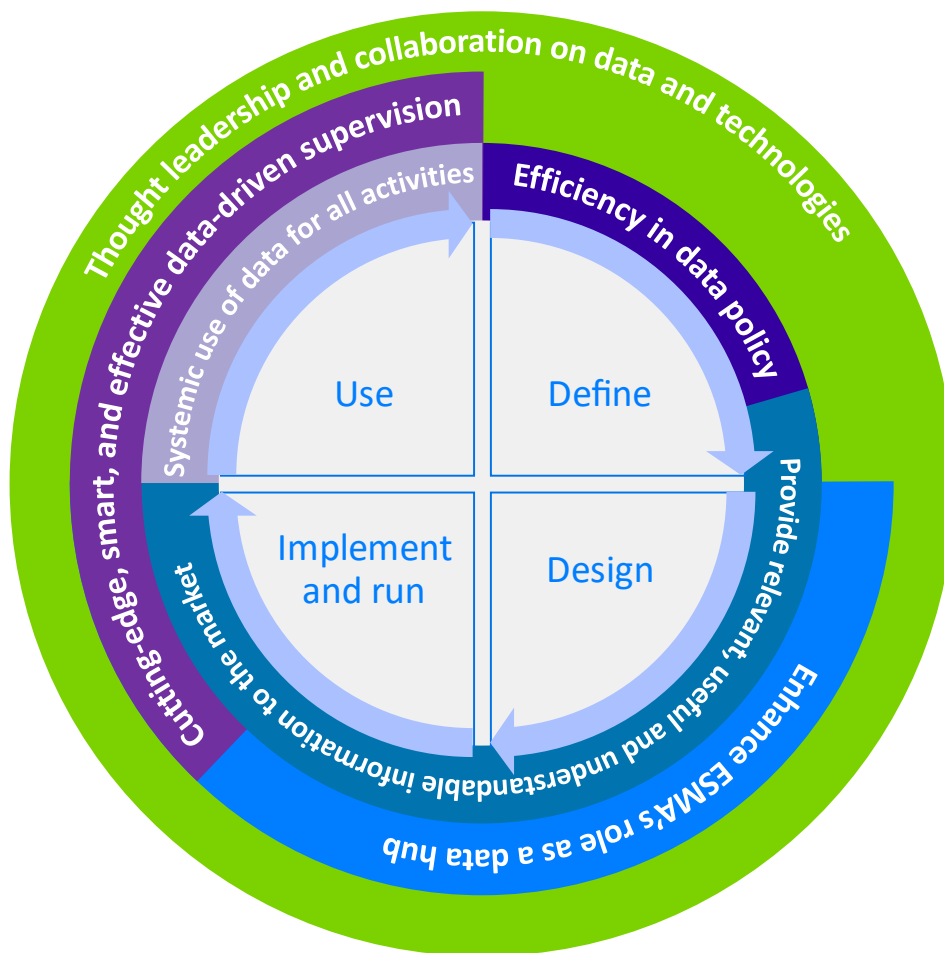
- Promote a **data-driven approach**, in particular:
 - Ensure ESMA **deliverables**²⁹ are **supported by data** findings and evidence wherever possible;
 - Develop an **ESMA framework** defining the data governance supporting ESMA's **internal activities**;
 - Develop **common methodologies** in collaboration with NCAs and other stakeholders to enable **consistent and comparable work on data** amongst them and ESMA;
- Build a central access point for **ready-to-use standardised data and analyses** to ensure data interoperability, connect the dots between different datasets, and avoid siloed work:
 - Complete the **migration of datasets** to the ESMA Data Platform to leverage the synergies from data ingestion to dashboard production across all datasets;
 - Create a **unified catalogue** of existing datasets, analytics, and processes, thereby promoting data literacy at ESMA;
 - Develop the **ESMA metadata dictionary**³⁰ in lockstep with the **ESMA Data Platform** to ensure more controls across datasets and centralise relevant information (data sources, data models, glossary of terms, etc.).

²⁹ RTS/ITS, guidelines, Q&As, peer reviews, opinions, supervisory outputs, risk assessment reports, etc.

³⁰ In this context, an inventory of data assets at ESMA through the discovery, description, and organisation of datasets.

4 Target capabilities

The actions listed in the previous section will impact ESMA activities across the entire data lifecycle, from the definition of data reporting requirements to the use of data by way of the design and implementation of data reporting systems. Given their scale and ambition, achieving all data strategic objectives within the next five years calls for a profound revamp of ESMA’s approach towards the different activities the ESMA data lifecycle consists of, and how they do and should interact with each other.



ESMA’s first response was to establish a new Data Intelligence and Technology department (DIT) that enables the consolidation of the full set of competences required to best leverage data assets, talents, and benefit from economies of scale and synergies, while supporting internal and external stakeholders throughout the data lifecycle. In this regard, **DIT will serve as the enabler of the Data Strategy**, working in close cooperation with other ESMA departments internally, and NCAs and other stakeholders externally.

Drawing from the results of the capability assessment conducted at the end of 2021³¹, DIT will focus its efforts on the **development, restructuring, or enhancement of the following capabilities:**

- Organisation and governance: Data functions should be designed to respond to stakeholder needs, avoid siloed work, and enable effective cross-department as well as cross-organisation coordination and collaboration (in particular with NCAs) at each step of the data lifecycle. It is under that premise that ESMA aims to leverage industry best practices and frameworks to **overhaul its existing data governance, establishing and clearly delineating roles, responsibilities, and processes**, so that it becomes more efficient and nimbler as an organisation³², while ensuring transparency and clear communication with its stakeholders. This new data governance model will be one of the first undertakings in the implementation of the Data Strategy.
- Processes and ways of working: For ESMA to become truly efficient in the coordination and execution of the many activities taking place throughout the data lifecycle, it will be of utmost importance to converge towards more streamlined processes, advance towards more standardisation and interoperability of data assets, while developing common ways of working. In that respect, important endeavours for ESMA will include designing **a common dataset framework** for defining prioritisation and measuring data quality, and providing **standardised data and analytics** to its stakeholders, both being particularly relevant in the context of our engagement with NCAs³³.
- Technology and tools: The technology and tools deployed at ESMA must be fit for purpose in that they should enable an efficient use of data by its stakeholders. To ensure this, as first step it is necessary to identify the key focus points to make ESMA best suited to help achieve the objectives laid out in the Data Strategy and, at a second stage, **prioritise and implement the required changes**. Amongst these changes, some such as the big data platform migration³⁴ are already under way, some are planned enhancements (e.g., a metadata dictionary, a unified catalogue of existing datasets, analytics, and processes), whereas some remain to be further specified (e.g., risk-assessment tools to identify risks and initiate supervisory actions);
- Talents and skills: In a fast-evolving technological environment, the palette of skills needed in data functions requires constant (re)-calibration as some practices, tools and (data) standards become outdated or supplanted by others. ESMA should maximise the use of existing resources and **demonstrate a strong capacity to not only attract, but also develop and retain data talents** in order to thrive in the medium- to long-term. More broadly speaking, and outside of the mere data functions, **data literacy**

³¹ See section [Current ESMA Data Situation](#).

³² By improving the identification of business needs and how to address them, ultimately facilitating the decision-making process.

³³ E.g., the many data quality engagement frameworks in place for major reporting regimes.

³⁴ To make data, information, and analytical tools facilitating the use of modern technologies available to ESMA stakeholders, enable collaborative analytics, and allow for “connecting the dots” between datasets.

also needs to be better embedded in the organisation culture to allow ESMA to realise the maximum value out of data for its stakeholders.

With regards to the above, ESMA has defined in the next section a **roadmap with clear milestones**. Going forward, this roadmap will be **communicated to stakeholders and updated regularly**. In addition, ESMA will monitor its fulfilment of the different objectives in the strategy and will develop appropriate metrics to assess the **progress** on the implementation of the listed points (e.g., by performing broad-based or targeted assessments and comparing their results with the results from previous ones). This will ensure that it remains on track to achieve the desired target state within the planned timeframe.

5 Implementation plan

5.1 Roadmap

This section outlines the key actions, milestones and outputs that will be delivered over the period of the next 5 years, in order to achieve the strategic objectives. Some of the actions are already planned or ongoing projects, therefore they are more specific and more information is available, whereas the other actions, in particular those planned for later in the strategy implementation period, remain high-level and will need to be more elaborated on before the implementation.

5.1.1 Objective A – Enhanced data hub

- **2023:** Develop the project to share data with NCAs, by onboarding them to the ESMA data platform
- **2024-2028:** Progressively extend data sharing with NCAs by implementing further use cases (e.g., set up a common repository of analytical codes with NCAs) and datasets (e.g., those related to the supervision of ESG disclosures, commercial datasets³⁵, or master data), on the basis of the identified needs and priorities
- **2024-2025:** Develop orderbook data sharing platform, on the basis of the new legislative proposal included in the listing package
- **2025-2027:** Develop integrated reporting system under AIFMD and UCITS, including arrangements to facilitate data exchanges between ESMA and other authorities (ECB, other ESAs, etc.)³⁶
- **2025-2026:** Pending the outcome of the feasibility study (see objective E), implement, in collaboration with other ESAs, the system to support ICT-related incident reporting and/or exchange under DORA

5.1.2 Objective B – Access to data of public interest

- **2023-2027:** Implement the ESAP system, in phases as envisaged in the regulation
- **2024-2028:** Potentially through ESAP, implement ways of sharing data and analytics with the market (e.g., via interactive dashboards), extending gradually the number of datasets covered

³⁵ Depending on contractual restrictions, considering the possibility of joint procurements, etc.

³⁶ The development of the system will be preceded by the study on integrated supervisory collection and the development of technical standards on reporting (see section 5.1.5), the implementation timeline will depend on the legislative deadlines.

- **2024:** Engage with the EC to discuss the possibility to publish (on an anonymised basis) ESMA data through the EU Digital Finance Platform Data Hub

5.1.3 Objective C – Data-driven supervision

- **2023-2024:** Develop the proof-of-concept for the detection of potential market abuse cases using AI techniques. Subsequently, depending on the outcome of the project, implement common AI-based tools for detection of market abuse cases
- **2023:** Develop a tool for detection of anomalies on CRA data using machine learning techniques
- **2023-2028:** Starting with the MiFIR reporting framework, progressively revise data quality frameworks for the key reporting regimes to ensure more outcome-focused, data-driven and risk-based approach
- **2023-2024:** Develop advanced monitoring dashboards, with early warning indicators, alerts, triggers and scoring to support an efficient prioritisation of supervisory and convergence work

5.1.4 Objective D – Thought leadership

- **2023-2028:** Propose proof-of-concepts / projects on the use of modern technologies related to, e.g., anomaly detection on different datasets (EMIR, SFTR, etc.) or the detection of greenwashing practices, and consider using alternative funding sources such as the TSI programme or similar
- **2023:** Join ISO governance bodies relevant for data standards in financial markets
- **2023-2024:** Engage with relevant stakeholders on the development of joint supervisory tools for crypto-assets services provides under MiCA
- **2023-2024:** Leverage on existing and future common IT projects with ESAs to explore ways to ensure an efficient exchange of data. In particular, provide access to TRACE data to EIOPA
- **2023-2024:** Establish a forum with ECB to discuss common data and technological initiatives, including data exchanges
- **2024-2028:** Reinforce ESMA engagement with external stakeholders, including different categories of market participants and academics, on data and data-related technologies, e.g., by organising workshops, undertaking joint projects and collecting feedbacks

5.1.5 Objective E – Efficient data policy

- **2023-2024:** In collaboration with the other ESAs, perform the study on further centralisation of ICT-related incident reporting under DORA
- **2023-2026:** Perform the study on the integrated reporting system for AIFMD/UCITS, in cooperation with the ECB and the other ESAs, and subsequently develop the RTS/ITS on reporting
- **2023-2024:** Develop a data dictionary for securities markets, as a contribution to the Strategy on supervisory data in EU financial services
- **2024:** Study the possibility to implement machine-readable and executable reporting

5.1.6 Objective F – Systematic data use

- **2023-2026:** Complete migration of all suitable datasets to the ESMA data platform
- **2023:** Develop ESMA data governance framework
- **2023-2024:** Develop a unified catalogue of datasets, analytics and processes, including services and activities that ESMA performs for or on behalf of its stakeholders, and finalise the development of the ESMA metadata dictionary
- **2024:** Establish a data literacy program across the organisation to raise awareness amongst ESMA stakeholders on what data is available at ESMA and how it can be or is already used to support their work
- **2024-2028:** Use RegTech and SupTech solutions, to consolidate and analyse multiple sources of data, including unstructured data, to make the use of data more effective and efficient

The implementation of the ESMA Data Strategy will be monitored continuously to ensure that it remains aligned with ESMA priorities and, if needed, it will be revised accordingly. Annual Work Programmes will ensure yearly reassessment of the prioritised projects.

5.2 Resources

The objectives and implementing actions of this strategy are ambitious and will have a substantial impact on resources of ESMA and NCAs. It is evident, that the objectives cannot be fully achieved under the existing resources constraints. Therefore, it becomes critical to explore new ways of funding and running projects. To that end, ESMA intends to:

- Explore the opportunities of new funding sources: Traditionally, all work performed by ESMA has been funded either by joint contribution by NCAs and the EU budget or by

specific NCAs contributions for delegated tasks. However, some of the projects highlighted in this strategy use alternative funding sources. Firstly, the Digital Europe Programme used for the implementation of the ESAP. Secondly, the Technical Support Instrument Programme (TSI) is used for the proof-of-concept project for detection of potential market abuse cases using AI techniques. This is the first instance where the TSI Programme is used to support a multi-country project coordinated by an ESA where the objective of the project is to enhance data-related processes. ESMA intends to explore the possibility to use further the above-mentioned or similar support programmes for specific projects included in this strategy. Finally, delegated projects delivered in the past allowed to mutualise the development cost and make it cheaper for the individual participating authorities. Further delegated projects to develop common tools instead of multiple national systems could be considered in the future.

- Explore new approaches to organising data and technology related projects: ESMA also intends to explore possible alternative approaches to coordination of joint projects with NCAs. The above-mentioned proof-of-concept project for detection of potential market abuse cases using AI techniques is already a good example of a project that will be developed within a group of 16 NCAs, with facilitation by ESMA, and where it is expected that the outcomes of the project will be shared with all NCAs. In collaboration with NCAs, ESMA will explore the possibilities of similar projects done by groups of NCAs, with participation or facilitation by ESMA, with the view to share the deliverables and experience with the entire supervisory community. Finally, joint procurements for tools and data could also be considered.
- Increase collaboration with the other ESAs: ESMA intends to increase the collaboration with EIOPA and EBA on data and IT-related projects. The examples of such joint cross-sector projects include the implementation of ESAP that is led by ESMA but will support data and information disclosures across the entire financial sector, the work on data collection system for the purpose of oversight under DORA that is led by EBA or the system Fitness and Propriety Information envisaged by the ESAs funding regulations which is led by EIOPA. In order to achieve economies of scales, more joint projects with other ESAs will be considered in the future. The joint work with the other ESAs is already envisaged and will continue to further consider joint procurement and contracts with suppliers of products and services related to data or technology.

The detailed planning of the work on the implementation of the strategic objectives will be included in the **annual IT and Data Work Programmes**. These documents will include further details on the projects to be launched in a given year, including their funding source(s) and allocated budget, supporting the annual planning that will help to prioritise the work in order to maximise the delivered value to the stakeholders given the resources constraints. It is also worth noting that some of the initiatives are already being started or ongoing and the funding

for those projects³⁷ has been already secured, whereas the funding of the future projects will be defined on a case-by-case basis in due time.

³⁷ Or, in some cases, for the initial phases of the projects.

6 Communication plan

In its communication activities related to data, ESMA intends to continue to:

- position itself as the leading public authority in financial market data reporting policy making, in the EU as well as at the international level;
- increase awareness about ESMA's work on regulatory and supervisory data;
- promote the use of data and the value added of the reporting regimes.

ESMA intends to undertake the following activities:

- publish the Data Strategy;
- increase the number of publications / articles on the work related to data and data related technologies, e.g., specific use cases, data related projects;
- consider new ways to engage with and get feedback from stakeholders, e.g., interactive webinars on data-related work;
- progressively standardise the communication and documentation on data reporting regimes.

7 Annexes

7.1 Annex I – Acronyms and definitions

AIFMD	Alternative Investment Fund Managers Directive
CCPs	Central Counterparties
Corporate search engine	Insight engine combining search with AI to deliver actionable insights derived from the full spectrum of content and data sourced within and external to the organization
Data Catalogue & Data Dictionary	A data catalogue and data dictionary maintain an inventory of data assets through the discovery, description, and organisation of datasets
Data literacy	The ability to read, write and communicate data in context, with an understanding of the data sources and constructs, analytical methods and techniques applied. It is the ability to describe the use-case application and resulting users' value or outcome
DORA	Digital Operational Resilience Act
DLT	Distributed Ledger Technology
DRSPs	Data Reporting Service Providers
ESAs	European Supervisory Authorities
ESAP	European Single Access Point
ESFS	European System of Financial Supervision
ESMA Data Platform	A set of data processing tools implemented by ESMA based on cutting-edge tools, such as Big Data, Cloud Computing, etc., allowing to ingest, store, process and analyse all data available to ESMA.
Experimental environment	Environment, also known as sandbox, that can be used to build prototypes and create new data outcomes in an iterative way.
KPI	Key Performance Indicator
Machine-readable data	Data in structured so that software applications can easily identify, recognise and extract specific data, including individual statements of fact, and their internal structure
Metadata management	Elements that allow creating active metadata by: 1. Collecting and performing graph analysis on traditional design-based metadata and runtime metadata; 2. Using results as input for ML/ AI algorithms and 3. Capturing the results in a knowledge graph of the data fabric that is ready for action
MDM / RDM	Master Data Management and Reference Data Management. These components manage the master data (= the consistent and uniform set of identifiers and extended attributes that describe the core entities of the enterprise) and the reference data in a consistent, quality assured way
MiCA	Regulation on Markets in Crypto-Assets
NCA	National Competent Authority
NLP	Natural Language Processing
RegTech	Regulatory technology

SupTech

Supervisory Technology

UCITS

Undertakings for the Collective Investment in Transferable Securities
