

Study on Data Formats for text-based disclosures

Independent study to evaluate the different data formats options for the MiCA white papers and other similar text-based disclosures

Table of Contents

1	Executive Summary	5
2	Introduction	6
2.1	Background of the Study	6
2.2	Goal of the Study	7
2.3	Approach.....	7
3	Alternative Data Formats.....	9
3.1	Regulatory Market Trends	9
3.2	Overview of Data Formats.....	11
4	Evaluation of the Suitability of Data Formats	12
4.1	Introduction	12
4.1.1	Fitness-for-Purpose.....	12
4.1.2	Fitness-for-Future.....	13
4.2	Evaluation Scale	14
4.2.1	Fitness-for-Purpose.....	14
4.2.2	Fitness-for-Future.....	14
4.3	Evaluation	15
4.3.1	Fitness-for-Purpose.....	15
4.3.1.1	Machine-Readability	15
4.3.1.2	Human-Readability	17
4.3.2	Fitness-for-Future.....	17
4.3.2.1	Level of Adoption	17
4.3.2.2	Reusability	18
4.3.2.3	Governance	19
4.3.2.4	Non-Proprietary	20
4.3.2.5	Implementation Feasibility	21

4.3.3	Summary of the Results	22
4.4	Cost Considerations	23
4.4.1	Cost considerations for the issuers.....	23
4.4.2	Cost considerations for the regulators and NCAs	24
5	Conclusions.....	25
5.1	Recommendations	25
5.2	Further Insights of iXBRL	25
5.3	Relationship with other EU Data Initiatives	26
5.3.1	European Single Electronic Format (ESEF)	27
5.3.2	Corporate Sustainability Reporting Directive (CSRD)	27
5.3.3	Sustainable Finance Disclosures Regulation (SFDR).....	28
5.3.4	European Single Access Point (ESAP)	29
6	Annexes	31
6.1	Market Scan of Machine-Readable Data Formats	31
6.2	PDF and Machine-readability	35

1 Executive Summary

Reasons for publication

In line with ESMA practices for selecting formats for data requirements, an independent study was carried out to assess the appropriate machine-readable format for the MiCA white papers and other similar textual disclosures.

Content

The study conducted market research to identify relevant data formats and use cases within similar regulatory authorities and public service organizations. This initial phase led to the identification of six technical data formats deemed relevant for text-based disclosures. These data formats were assessed using an evaluation framework focused on two key aspects: fitness-for-purpose and fitness-for-future. This approach ensured the chosen format would meet the immediate needs of text-based disclosures under the MiCA regulation but also remain adaptable and future-proof for similar textual disclosures covered by ESAP.

Based on the results of the suitability evaluation, the independent study recommends adopting Inline XBRL (iXBRL) for the disclosure of MiCA white papers and for other similar text-based disclosures in scope of ESAP. iXBRL offers a seamless combination of machine- and human-readability, with the flexibility of handling unstructured and qualitative information. Furthermore, the format meets the requirements set forth by other EU data initiatives and presents an opportunity for standardization of data formats used in text-based disclosures. By adopting iXBRL, regulatory authorities and NCAs would ensure consistency and efficiency in their reporting processes.

Disclaimer

This document has been prepared for the European Securities and Markets Authority (ESMA) by Gartner Consulting. It reflects the views only of its authors, and the European Securities and Markets Authority is not liable for any consequence stemming from the reuse of this publication.

2 Introduction

2.1 Background of the Study

1. The European Securities and Markets Authority (ESMA) has been entrusted with the responsibility of defining the standard forms, formats, and templates of the white papers for different categories of crypto assets under the Markets in Crypto-assets Regulation (MiCA).
2. MiCA explicitly requires white papers to be made available in a machine-readable format for all three crypto assets classes: (a) Crypto assets other than asset-referenced tokens or e-money tokens (Article 6), (b) Asset-referenced tokens (ARTs) (Article 19), and (c) e-money tokens (EMTs) (Article 51).
3. Unlike the mandate under Article 21 of the Prospectus Regulation, which primarily focuses on metadata, MiCA mandates the standardization of the white papers themselves. This means that both the data necessary for the classification of the documents *and* the content of the white paper must be provided in a machine-readable format. Doing so is expected to ensure accessibility and ease of use for both investors (including retail investors) and National Competent Authorities (NCAs).
4. MiCA also clearly defines the information that issuers and Crypto Asset Services Providers (CASPs) must provide within the white papers. To ensure consistency and transparency on the information provided to investors and NCAs, no additional information beyond the specified requirements can be included by preparers of the white papers.
5. However, issuers and CASPs are allowed to provide additional information in different documents, for example as part of their “marketing communication”. This additional information must meet relevant requirements under MiCA and be consistent with the information provided in the crypto asset white paper (recital 24).
6. The primary objective of white papers under MiCA is to inform and protect retail investors (recitals 24 and 47). For this, the white papers must be easily usable and accessible to ensure that retail investors are able to make informed decisions. To this purpose and to avoid proliferation of non-official documents, even if not explicitly mentioned in Articles 6, 19 and 51, it is considered essential that the white papers are human-readable.
7. Public offers of crypto assets other than ARTs or EMTs that are addressed to fewer than 150 persons per Member State or are aimed exclusively at qualified investors or whose total consideration does not exceed EUR 1,000,000 over a certain period of 12 months are excluded from the obligation to prepare a white paper.

8. Furthermore, as part of the European Single Access Point (ESAP), ESMA is mandated to assess whether other, heavily text-based disclosures aimed at investors, including retail investors, should be mandated in a machine-readable format and if so, in which format. For many sectorial legislations falling under the scope of ESAP (see ESAP Omnibus Directive¹ and Regulation²), there is a specific mandate for ESMA to assess which information should be presented in a machine-readable format and to determine the most suitable machine-readable format for that purpose.
9. Hence, this study must also serve the purpose of assessing the existing machine-readable formats that may be considered to effectively develop this mandate.
10. The reference to ESAP is also relevant because the MiCA white papers will be submitted to ESAP starting from 2030. Therefore, all policy decisions concerning MiCA should be consistent with the policy decisions concerning ESAP, and vice versa, since MiCA is one of the sectorial legislations falling in scope of ESAP.

2.2 Goal of the Study

11. Based on the aforementioned context, ESMA has commissioned an independent study to assess the available machine-readable formats that could meet the data requirements set forth in MiCA and in ESAP.
12. The study was commissioned to an external consultant to support the following tasks:
 - Scan the market for possible format options,
 - Compare and assess the different data format options according to jointly agreed criteria³ (i.e., fitness for purpose, fitness for future, and high-level incremental cost for issuers, CASPs, and NCAs), and
 - Derive recommendations.

2.3 Approach

13. The independent assessment followed a phased approach where the following activities were performed:

¹ DIRECTIVE (EU) 2023/2864

² REGULATION (EU) 2023/2869

³ The criteria for the assessment have been agreed with the members of the Disclosure Working Group (DWG) operating under the Data Standing Committee of ESMA. <https://www.esma.europa.eu/about-esma/governance-structure/standing-committees>

- **Market trends analysis:** Research and analysis of the current market trends in the regulatory ecosystem across various jurisdictions. The aim was to identify data formats used in similar use cases as well as related general market trends.
- **Data formats overview:** Creation of a comprehensive longlist of relevant and similarly used data formats based on the market research results.
- **Definition of appropriate evaluation framework:** Definition and implementation of a structured procedure and evaluation criteria for assessing the suitability of the various data formats.
- **Comparative analysis:** Assessment and comparison of the various data formats based on the evaluation framework.
- **Conclusion:** Provide an independent and external view of the available machine-readable formats that could meet the disclosure requirements set out in MiCA and highlight whenever those requirements are similar in the context of ESAP. Based on the outcome of the analysis, the study will also provide clear recommendations.

3 Alternative Data Formats

3.1 Regulatory Market Trends

14. The study began with a market study of representative relevant regulatory authorities and other public service organizations. The aim of this analysis was to understand the extent to which different data formats are currently being used and what related market trends there are.

15. The market study spanned across multiple jurisdictions and identified relevant data formats and use cases within 20 representative organizations. The results are summarized in the following table⁴.

Entities	Data Formats
Australian Securities and Investments Commission (ASIC)	XBRL iXBRL PDF
Bank of Japan (BOJ)	XBRL iXBRL
Bank of Korea (BOK)	XBRL XML
Bank of Mauritius (BOM)	XBRL
Canadian Securities Administrators (CSA)	XBRL
Dutch Government	iXBRL
HM Revenue & Customs (HMRC)	XBRL iXBRL
Hong Kong Monetary Authority (HKMA)	iXBRL

⁴ Section 5.3 provides more details on the EU Data initiatives by ESMA, EBA and EIOPA (ESAs) pertaining to iXBRL

	XML CSV JSON
Israel Securities Authority (ISA)	iXBRL
Japan Financial Services Commission	XBRL iXBRL
Mauritius Financial Services Commission	XML CSV Docx
Ontario Security Commission (OSC)	PDF
Reserve Bank of India (RBI)	XBRL
Saudi Arabian Monetary Authority (SAMA)	XBRL iXBRL PDF
Securities and Exchange Commission (SEC)	XBRL iXBRL XML PDF
ESMA (EU)	XBRL
EBA (EU)	XBRL CSV (iXBRL)
EIOPA (EU)	XBRL

South Africa Companies and Intellectual Property Commission (CIPC)	XBRL
Switzerland Task Force on Climate-related Financial Disclosures (TCFD)	iXBRL
Taiwan Stock Exchange	XBRL iXBRL

3.2 Overview of Data Formats

16. The market study of representative organizations revealed a long list of seven relevant reporting data formats that form the commonly used basis for various types of regulatory reporting. Please note that the study focused exclusively on machine-readable data formats that are already used in related domains.
17. **XBRL (eXtensible Business Reporting Language):** XBRL⁵ is an open standard based on XML (Extensible Markup Language) commonly used for digital business reporting. It provides a language in which reporting terms can be authoritatively defined and including in a taxonomy, acting as a dictionary. XBRL allows unique tags to be assigned to reported facts, ensuring more accurate and efficient reporting. The standard is managed by XBRL International, a non-profit organization.
18. **iXBRL (inline eXtensible Business Reporting Language):** iXBRL is a structured language based on the HTML (HyperText Markup Language) standard. This format is an open standard and is recognized as a format that allows users to create documents that are both machine-readable and human-readable without need for special software (viewing can be done via any internet browser). iXBRL is based on the HTML standard and on XBRL tagging.⁶
19. **XML (Extensible Markup Language):** XML is a markup language designed to store, transmit, and reconstruct arbitrary data. This format was created to facilitate and support information exchange between computer systems.⁷

⁵ <https://www.xbrl.org/the-standard/what/an-introduction-to-xbrl/>

⁶ <https://www.sec.gov/structureddata/osd-inline-xbrl.html>

⁷ <https://www.w3.org/XML/>

20. **CSV (Comma-separated values):** CSV is a simple format, considered as a flat format, often used to represent arrays of numeric and textual values. It is a delimited format that has fields separated by a comma and rows separated by characters indicating a line break.⁸
21. **JSON (JavaScript Object Notation):** JSON is a data interchange format based on a subset of the JavaScript Programming Language Standard. The text format is language-independent but uses conventions that are familiar to programmers fluent in the C-family languages, making it easy for human to write and read and easy for machine to parse and generate.⁹
22. **Docx:** Docx is the format for Microsoft Word Documents used by Microsoft Office 2007 and later versions. The document evolved from plain binary to a combination of XML and binary files.¹⁰
23. **PDF (Portable Document Format):** PDF is an open standard maintained by the International Organization for Standardization (ISO), initially created by Adobe. This format is commonly recognized as the format for document presentation and exchange. This format is agnostic of software, hardware, or operating system.¹¹

4 Evaluation of the Suitability of Data Formats

4.1 Introduction

24. The suitability of the different data formats was assessed based on two main criteria: “Fitness-for-purpose” and “Fitness-for-future” of the respective data format for the MiCA white papers as well as other similar textual disclosures in the context of ESAP.
25. “Fitness-for-purpose” describes the ability of the data format to meet current regulatory requirements, and “Fitness-for-future” describes the ability to meet future requirements.

4.1.1 Fitness-for-Purpose

26. This criterion is intended to assess whether a particular data format is suitable for the specific use case of the MiCA white papers as well as other similar textual disclosures in

⁸ <https://www.loc.gov/preservation/digital/formats/fdd/fdd000323.shtml>

⁹ <https://www.json.org/json-en.html>

¹⁰ <https://docs.fileformat.com/word-processing/docx/>

¹¹ https://www.adobe.com/be_en/acrobat/about-adobe-pdf.html

the context of ESAP for which ESMA has a similar mandate. Please note that the MiCA white papers will be in the scope of ESAP starting from 2030 (“phase 3” of ESAP).

27. MiCA explicitly requires that white papers for all three crypto asset categories (ARTs, EMTs, and other crypto assets) must be machine-readable. In addition, the white papers serve to inform and protect investors. Therefore, it is also important (even if this is not required by MiCA) that the white papers be human-readable.
28. Essential aspects of “fitness-for-purpose” are therefore the degree of machine- and human-readability of the data formats, which are defined in more detail as follows:
 - Machine-readability: the Open Data Directive (EU 2019/1024) defines machine-readability as enabling data identification, recognition, and extraction (this is the definition of machine-readability to which the ESAP L1 text makes reference).
 - Human-readability: This requirement must allow readers, and particularly retail investors, to read the document without the need to use special software tools.

4.1.2 Fitness-for-Future

29. The second main criterion assessed the future viability of the selected data formats, i.e., the ability of the specific data format to meet ESMA's future requirements.
30. This main criterion was divided into five sub-criteria:
 - Level of Adoption: Level of adoption in other similar regulatory framework in Europe and other jurisdictions.
 - Reusability: Ability to accommodate diverse data structures, adapt to varying data requirements, and support customization and extensibility.
 - Governance: Involvement of industry participants and regulators in the maintenance of the standard.
 - Non-Proprietary: Ability to benefit from the solution without any IP restriction and free of charge.
 - Implementation Feasibility: Ability to be implemented at a reasonable incremental cost for both NCAs and the market.

4.2 Evaluation Scale

4.2.1 Fitness-for-Purpose

	High / (Yes)	Medium	Low / (No)
Machine-Readability	The data format meets the criteria set forth by the Open Data Directive.	N/A	The data format does not meet the criteria set forth by the Open Data Directive.
Human-Readability	The data format is easily understandable and accessible without requiring specialized knowledge or expertise. The information presented is clear, well-structured, and user-friendly, enabling individuals to comprehend and interpret the data effortlessly.	The data format requires a moderate level of effort and expertise to comprehend. Individuals with specialized knowledge should be able to understand the information with reasonable ease. It may require third-party software to facilitate its human-readability.	The data format is challenging to understand and interpret without significant specialized knowledge or expertise. It requires a proprietary third-party software to facilitate its human-readability.

4.2.2 Fitness-for-Future

	High / (Yes)	Medium	Low / (No)
Level of Adoption	The data format is widely adopted in similar regulatory frameworks across Europe and other jurisdictions.	The data format is moderately adopted in similar regulatory frameworks across Europe and other jurisdictions.	The data format has limited or no adoption in similar regulatory frameworks across Europe and other jurisdictions.
Reusability	The data format has a high degree of reusability and can easily accommodate	The data format has some degree of reusability but may require additional	The data format has a limited degree of reusability and may not adapt to diverse data

	various data structures and requirements.	efforts to accommodate certain data structures or requirements.	structures or requirements.
Governance	The industry and regulators actively participate in the maintenance and development of the data format, providing clear collaboration and regular maintenance/update.	The industry and regulators are moderately involved in the maintenance and development of the data format, lacking clear collaboration and regular maintenance/update.	The data format lacks industry and regulatory involvement in its maintenance and development.
Non-Proprietary	The data format is non-proprietary.	N/A	The data format is proprietary.
Implementation Feasibility	The data format can be implemented at a reasonable cost for both NCAs and market participants.	The implementation may require some additional costs and resources but remain feasible for most NCAs and market participants.	The implementation is associated with high costs and significant challenges, making it difficult for both NCAs and market participants.

4.3 Evaluation

4.3.1 Fitness-for-Purpose

4.3.1.1 Machine-Readability

Data Format	Rating	Comments
CSV	No	While CSV is a simple tabular format that can easily be parsed by a machine, it does not feature data recognition and identification capabilities out-of-the-box. Hence, in the context of our study, we do not consider CSV as

		machine-readable as per the Open Data Directive definition.
iXBRL	Yes	iXBRL is an extension of the XBRL format (see below), sharing similar characteristics and capabilities with the enhanced capability of being human-readable.
JSON	Yes	JSON is widely used in web APIs, allowing for the identification, recognition and extraction of data. These characteristics make JSON a machine-readable format according to the Open Data Directive.
PDF	No	PDF is primarily designed for human-readability. Extracting data can prove complex. For example, if a PDF document contains data tables, this information is not machine-readable as a computer would have difficulty accessing the tabular information - although they are very readable by humans. The corresponding tables in a format such as a spreadsheet would be machine-readable.
XBRL	Yes	XBRL is a global standard for exchanging business information based on structured data language, making it natively machine-readable.
XML	Yes	XML is a markup language defining sets of rules for encoding document, allowing for the identification, recognition, and extraction of data. These characteristics make XML a machine-readable format according to the Open Data Directive.
XLS	No	XLS is the default format for Microsoft Excel. Currently, XLS does not have built-in capabilities for data identification and recognition, thus failing to meet the Open Data Directive definition for machine-readability

4.3.1.2 Human-Readability

Data Format	Rating	Comments
CSV	Medium	CSV is a simple tabular format that can easily be understood by humans, although its readability is often enhanced by a text editor or a third-party software.
iXBRL	High	iXBRL is designed to be human-readable as it embeds XBRL data within an HTML document, which can simply be opened and viewed with a standard web browser and does not therefore require additional third-party software.
JSON	Low	While JSON uses human-readable text to store and transmit data, doing so requires a third-party software (JSON viewers and formatters).
PDF	High	Designed specifically for human-readability, PDF is the most used human-readable format across all industries.
XBRL	Low	XBRL is primarily designed for machine-readability, and while it is text-based, it is not suitable for humans to read without additional software and requires a third-party software such as an XBRL viewer.
XML	Low	XML cannot easily be understood and interpreted by someone without the specialized knowledge. It also requires third-party software to allow experts to comfortably read it.
XLS	Low	XLS does not meet the human-readable criteria as it always requires a proprietary software.

4.3.2 Fitness-for-Future

4.3.2.1 Level of Adoption

Data Format	Rating	Comment
-------------	--------	---------

CSV	High	CSV format is still widely used for financial reporting/disclosures for simple organizations with smaller data sets or in countries where XBRL has not yet been introduced.
iXBRL	High	iXBRL has seen recent widespread adoption for disclosures requiring human-readability (e.g., financial statements, tax disclosures, ESG reporting)
JSON	Low	The market analysis has revealed very limited examples of JSON being used by regulators and then for data access via APIs.
PDF	High	As of the time of this study, PDF is the most widespread and common format for disclosures which do not require machine-readability.
XBRL	High	XBRL has been used for the past decades by regulatory bodies seeking to formalize and streamline the disclosure of financial statements or tax disclosures.
XML	High	XML format is still widely used for disclosure of financial data sets and forms, as an alternative to CSV format.
XLS	Low	The market trend assessment found that relevant regulators tend to favor the XML format instead of the basic format for Excel (XLS).

4.3.2.2 Reusability

Data Format	Rating	Comments
CSV	High	CSV is a simple, widely supported format that accommodates diverse data structures and is easy to customize. It may lack certain features for complex structures, but it is generally reusable for various data types.
iXBRL	High	iXBRL allows the creation of reusable taxonomies developed by regulators, capturing

		the meaning of it and the relationships between terms.
JSON	High	JSON is highly reusable and flexible format for data interchange. It accommodates diverse data structures, is easy to adapt, and supports extensibility.
PDF	Low	PDF is not designed for data interchange but only for document presentation. Extracting data from this format can prove to be challenging and it may not support diverse data structures or customization as easily as other formats studied here.
XBRL	High	XBRL allows the creation of reusable taxonomies developed by regulators, capturing the meaning and the relationships between terms.
XML	High	XML is highly reusable and extensible. It allows the definition of custom data structures through user-defined tags, making it adaptable to varying data requirements.
XLS	Low	While Excel spreadsheets are widely used, they may not be as adaptable as other format studied here. They are most suitable for tabular data but may have limitations for other more complex types of data.

4.3.2.3 Governance

Data Format	Rating	Comments
CSV	Low	CSV is a rather informal format which lacks a formal governance structure, either at industry or regulatory level.
iXBRL	High	It is managed by XBRL International, which is a global non-for-profit organization. Major regulators have a proven track of implementing

		and maintaining the iXBRL standard. (e.g., SEC, JFSA, BOJ, HMRC)
JSON	High	JSON is an open standard and supported by the ISO20022 standard.
PDF	Low	PDF is governed by Adobe, which means that its updates and maintenance are mainly under their control. Involvement of external bodies is limited.
XBRL	High	Major regulators have a proven track of implementing and maintaining the XBRL standard. (e.g., SEC, JFSA, BOJ, HMRC). It is also managed by XBRL International.
XML	High	XML is an open standard benefiting from the governance of the World Wide Web Consortium (W3C) and is often used in conjunction with specific industry standards. Additionally, XML is supported by the ISO20022 standard.
XLS	Low	XLS is a proprietary format, and its governance is thus determined by Microsoft. Involvement of external industry participants and regulators is very limited.

4.3.2.4 Non-Proprietary

Data Format	Rating	Comments
CSV	High	CSV is a text file format using commas to separate values and can thus be used with a variety of software, ranging from paying and open source.
iXBRL	High	The open XBRL specifications are freely licensed to anyone seeking to use the standard developed and maintained by XBRL International.
JSON	High	JSON is a non-proprietary format used for data interchanging, derived from JavaScript but not limited to it.

PDF	High	While originally a proprietary format, PDF has been released as an open standard in 2008.
XBRL	High	The open XBRL specifications are freely licensed to anyone seeking to use the standard developed and maintained by XBRL International.
XML	High	XML is an open standard developed by the World Wide Web Consortium and not limited to one single software.
XLS	Low	The XLS format is still considered proprietary, even if since 2006 it has been covered by Microsoft's Open Specification Promise, promising not to assert any Microsoft Necessary Claims against the individuals.

4.3.2.5 Implementation Feasibility

Data Format	Rating	Comments
CSV	High	Implementing CSV is typically straight-forward and can be done at low cost. Most programming languages and applications have built-in support for CSV.
iXBRL	Medium	Converting files into the iXBRL formats either requires the use of a dedicated software or the use of external third-party services.
JSON	Medium	Most programming languages have built-in support for JSON, making it relatively accessible and available. However, implementing this data formats requires coding expertise and efforts, driving up the cost.
PDF	High	Converting documents to the PDF format has a very limited, near inexistant, impact on costs.
XBRL	Medium	Converting files into the XBRL formats either requires the use of a dedicated software or the use of external third-party services.

XML	Medium	XML is a well-established standard and can be implemented with a wide range of tools and libraries. However, designing an effective XML schema and implementing complex data structure can require robust expertise and efforts, driving the cost up.
XLS	Medium	The use of XLS may involve moderate costs as it requires an Excel license or similar software and handling large complex data structures might require additional resources and/or expertise.

4.3.3 Summary of the Results

31. The results of the suitability assessment for the seven data formats are summarized in the following table:

	CSV	iXBRL	JSON	PDF	XBRL	XML	XLS
Fitness-for-Purpose							
Machine-Readability	No	Yes	Yes	No	Yes	Yes	No
Human-Readability	Medium	High	Low	High	Low	Low	Low
Fitness-for-Future							
Level of Adoption	High	High	Low	High	High	High	Low
Reusability	High	High	High	Low	High	High	Low
Governance	Low	High	High	Low	High	High	Low
Non-Proprietary	Yes	Yes	Yes	Yes	Yes	Yes	No
Implementation Feasibility	High	Medium	Medium	High	Medium	Medium	Medium

32. Based on the outcome of the evaluation, we observe that iXBRL seems to be the most suitable data format for the disclosure of white papers under the MiCA regulation.

33. iXBRL appears to be the only data format that provides a seamless combination of machine and human readability, as required in the context of MiCA. The data format is an open-standard and offers a high degree of flexibility through its reusable taxonomies. iXBRL also benefits from an increasing adoption rate in similar jurisdictions and use cases, confirming and solidifying its position as most suitable data formats for white paper disclosure.

34. iXBRL remains an emerging technology used in similar contexts, which limits the availability of data on its implementation related costs. Given the data scarcity and the time

constraints of this study, the following subchapter aims to identify cost proxies that can be utilized to gain insights on cost considerations associated with iXBRL.

4.4 Cost Considerations

35. This subchapter identifies cost proxies to evaluate the high-level incremental costs associated with using iXBRL, which will be assessed separately for issuers and for regulators and NCAs.

4.4.1 Cost considerations for the issuers

36. To date, there is a lack of information about the cost of producing white papers iXBRL format since such a requirement does not exist. Therefore, the study examined two different possible scenarios to assess the potential incremental costs for issuers and CASPs.

37. The first scenario assumes the conversion of a PDF, Word, or Excel document directly into the iXBRL format through a third-party provider.

38. Current estimation for iXBRL tagging or conversion costs vary significantly across different sources. HMRC iXBRL filing¹² typically ranges from 199€ to 299€ while ESEF reporting services¹³ tend to range from 900€ to 1350€, depending on the size of the file and the expected time it should take to convert the file.

39. The second scenario considers the use of an online form, or an Excel spreadsheet provided by ESMA, to report the required white paper information. The received information is then assumed to be converted by ESMA (in case of the online platform) or by the issuer (in case of an Excel spreadsheet tool) into the iXBRL format. This approach would not incur any additional costs for issuers as they would simply copy and paste the relevant white paper content into the appropriate field in the online form or Excel spreadsheet. However, the cost impact on ESMA to develop the platform would need to be considered.

40. In summary, adopting iXBRL as a reporting data format involves costs that may vary depending on the implementation approach. Prices from service providers in the UK show that conversion services are already available today in a price range of €199 to €1350, depending on the specific reporting context. This should be a strong indicator of the potential price range of similar services for white papers. It is also expected that as market

¹² <https://www.fintags.co.uk/ixbrl-tagging-cost>; <https://www.pdf2ixbrl.co.uk/pdf2ixbrl/prices.aspx>

¹³ <https://www.fintags.eu/esef-reporting-services-pricing>

adoption increases, competition between service providers should increase and therefore prices of conversion services will tend to decrease.

4.4.2 Cost considerations for the regulators and NCAs

41. The cost considerations for the regulators and NCAs was investigated by mean of interviews with two comparable regulatory organizations who already have experience with implementing iXBRL for text-based disclosures.
42. To this day, none of the interviewed regulatory organizations were able to quantify the costs associated with implementing iXBRL. The main reason being that those organizations used internal resources to develop and implement the data formats, and assessing the costs would require an analysis which has not been possible for them to undertake in the given timeframe.
43. Nonetheless, given the information collected in these interviews, it is possible to form working assumptions on the cost considerations for the regulators and NCAs:
 - The interviewed organizations shared the view that incremental costs decrease for iXBRL as its adoption progresses. As more disclosures are mandated in iXBRL, the data format becomes an industry standard and the incremental cost, both for the preparers and the regulator, decreases over time.
 - Text-based disclosures, such as MiCA white papers, require fewer and simpler validation rules compared to IFRS financial reporting disclosures.
44. Moreover, the NCAs are already developing or expecting to develop their technical capabilities to handle iXBRL due to the ESEF reporting format. In the various discussion leading to this report, relevant stakeholders have recognized ESEF reporting as being more complex than MiCA. Considering these factors, the additional cost of handling MiCA white papers in an iXBRL format can be considered as very limited.
45. As iXBRL adoption grows in Europe and similar jurisdictions, it will be crucial to closely monitor cost-benefit analyses performed by experienced organization. However, with the current level of information available in the market, the aforementioned elements will be the only reference points for the cost consideration for regulator and NCAs in this study.

5 Conclusions

5.1 Recommendations

46. Gartner recommends adopting iXBRL as technical data format for the disclosure of white papers under the MiCA regulation and other similar text-based disclosures that will be covered by ESAP.
47. The suitability evaluation results allow us to identify iXBRL as the most suitable data format based on the assessment performed. iXBRL is an open and non-proprietary technical data format offering a seamless combination of machine- and human-readability, with the flexibility of handling unstructured and qualitative information.
48. The market research also allowed us to observe a positive trend in the adoption of iXBRL in other related regulatory authorities and public service organization as a standard for heavy narrative text-based disclosures. The increased adoption in iXBRL is reflected in the increased support offered by software companies and conversion service providers in Europe and other jurisdictions.
49. At the time of the study, and based on the available market information, the cost considerations for the issuers are deemed reasonable, with expected conversion services costs ranging from 199€ to 1350€, depending on the nature and size of the disclosures. As iXBRL's adoption rate increases, it is expected that the market offering will continue expanding and lead costs for issuers to decrease over time.
50. It can be argued that deviating from the standardization trend of using iXBRL for text-based disclosures could result in higher costs for the issuers in the medium to long term.
51. Finally, iXBRL meets the requirements set forth by other EU data initiatives and could be reused for other similar text-based disclosures covered by ESAP. The standardization efforts are essential in reducing the incremental costs for the regulator and/or NCAs.
52. The following subchapter highlights the additional advantages of adopting iXBRL for the white paper disclosure under the MiCA regulation.

5.2 Further Insights of iXBRL

53. Further analysis of iXBRL reveals additional advantages as those covered in the previous subchapter.

54. As a standard functionality, iXBRL offers compatibility with standard browsers, allowing for an easy display and printing without resorting to proprietary software. Furthermore, our observation of the market trends in the United Kingdom allows us to assume that with the increasing amount of software companies supporting iXBRL, the need for extensive software development efforts and the associated risks will be significantly reduced.
55. iXBRL is considered as an accessible format for developers given that it is based on HTML, which allows the data format to support graphics, various formats, and languages providing further flexibility to present information. This feature could be interesting for white papers, which may contain graphs and images to complement the narrative and unstructured data, even if it is not foreseen yet in the MiCA regulation.
56. Those characteristics would facilitate visualization of information, facilitating consistency and correctness checks. This capability reduces efforts for both the preparers and consumers, ensuring cost-effective, useful, and reliable results from iXBRL reporting.
57. Finally, in the context of the MiCA white papers, iXBRL would require few and simple validation rules. Various stakeholders interviewed in the context of the study confirmed that ESEF required more complex validation rules compared to what would be needed under the MiCA regulation.

5.3 Relationship with other EU Data Initiatives

58. The project team also investigated other related EU data initiatives to see how iXBRL is treated in their context. Such initiatives investigated are:
- European Single Electronic Format (ESEF)
 - Corporate Sustainability Reporting Directive (CSRD)
 - Sustainable Finance Disclosures Regulation (SFDR)
 - EBA Pillar 3 Reporting
 - European Single Access Point (ESAP)
59. Overall, the results of the analysis showed that the use of iXBRL for MiCA white papers and other similar textual disclosures was largely consistent with all of these initiatives.
60. The following pages provide some more details of the relationship between the different EU data initiatives and iXBRL.

5.3.1 European Single Electronic Format (ESEF)

Source: <https://www.esma.europa.eu/press-news/esma-news/esma-publishes-2022-esef-xbrl-taxonomy-files-and-esef-conformance-suite>

Background

- The European Single Electronic Format (ESEF) is the electronic reporting format in which issuers whose securities are admitted to trading on EU regulated markets must prepare their annual financial reports to facilitate accessibility, analysis and comparability of annual financial reports.
- The ESEF means providing information in a machine-readable and human-readable format - iXBRL.

Data Format

- The ESEF Regulation requires that all issuers with securities listed on an EU regulated market prepare their annual financial reports in xHTML and mark-up the IFRS consolidated financial statements contained therein using XBRL tags and the iXBRL technology.
- The 2022 ESEF taxonomy is based on the 2022 IFRS Taxonomy, prepared and updated annually by the IFRS Foundation.
- The IFRS Foundation also publishes a versioning document explaining the changes compared to the previous taxonomy. The ESEF taxonomy package includes labels in all EU languages.

5.3.2 Corporate Sustainability Reporting Directive (CSRD)

Source: https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en

Background

- EU law requires all large companies and all listed companies (except listed micro-enterprises) to disclose information on what they see as the risks and opportunities arising from social and environmental issues, and on the impact of their activities on people and the environment.
- This helps investors, civil society organisations, consumers and other stakeholders to evaluate the sustainability performance of companies, as part of the European green deal.

- Companies subject to the CSRD will have to report according to European Sustainability Reporting Standards (ESRS). The standards were developed by the EFRAG, previously known as the European Financial Reporting Advisory Group, an independent body bringing together various different stakeholders.

Data Format

- The Corporate Sustainability Reporting Directive (CSRD) determines that sustainability information will have to be reported digitally in accordance with the European Single Electronic Format (ESEF).
- The ESEF means providing information in a machine-readable and human-readable format – iXBRL.

5.3.3 Sustainable Finance Disclosures Regulation (SFDR)

Source: https://www.esma.europa.eu/sites/default/files/2023-12/JC_2023_55_-_Final_Report_SFDR_Delegated_Regulation_amending_RTS.pdf

Background

- The EU has put in place a transparency framework, the Sustainable Finance Disclosure Regulation (SFDR).
- By setting out how financial market participants have to disclose sustainability information, it helps those investors who seek to put their money into companies and projects supporting sustainability objectives to make informed choices. The SFDR is also designed to allow investors to properly assess how sustainability risks are integrated in the investment decision process. In this way, the SFDR contributes to one of the EU's big political objectives: attracting private funding to help Europe make the shift to a net-zero economy.

Data Format

- The ESAs have taken into account the responses from stakeholders, of whom a substantial majority expressed support for machine-readable SFDR disclosures.
- Many stakeholders specifically recommended iXBRL to meet the format requirements, ensuring that disclosures cater to both human and automated processing needs.
- Recognizing the importance of aligning with the forthcoming European Single Access Point (ESAP) Regulation, the ESAs propose that the disclosures made available on

ESAP should be presented in a format that is both human-readable and machine-readable, specifically iXBRL – a move which will affect all disclosures covered by SFDR. This approach aligns with sustainability reporting requirements under the Corporate Sustainability Reporting Directive (CSRD), fostering consistency and ease of interpretation.*

**Please note that this proposal is still under the European Commission’s consideration.*

EBA Pillar 3 Reporting

Background

- Following the recent updates to the regulatory frameworks for credit institutions and investment firms, and the publication in 2018 of the European Commission’s action plan on sustainable finance, the EBA is implementing a new policy strategy on institutions’ Pillar 3 disclosures that seeks to increase efficiency of institutions’ disclosures and reinforce market discipline by developing a comprehensive framework with consistent and comparable disclosures.
- In addition, the EBA aims to promote transparency on Environmental, Social or Governance (ESG) risks, encouraging institutions to strengthen their management of these risks and promoting awareness of their key role in the transition to a green economy.

Data Format

- EBA is proposing XBRL-CSV as the Pillar 3 reporting data format and not iXBRL due to time constraints.
- EBA is investigating the use of iXBRL for qualitative (unstructured) data fields - this question was asked by participants in the public hearing on January 23, 2024.

5.3.4 European Single Access Point (ESAP)

Background

- The European single access point (ESAP) will provide centralized electronic access to information relevant to capital markets and financial services (including sustainability-related information) that must be made public under EU legislation, but also to other

types of information relevant to capital markets and financial services that companies voluntarily decide to publish.

- The ESAP is the number one Action of the 2020 Capital Markets Union (CMU) action plan.

Data Format

- Some data will be available in machine-readable format, helping users search and compare data more easily,
- All information will be available in a data-extractable format, enabling comparability, under open terms of use.
- The ESAs will have powers to develop machine-readable formats on a case-by-case basis.
- iXBRL is expected to be one of the format allowed for machine-readable data on the basis of the JC Consultation on the ITSs on Article 5 and Article 7 of ESAP¹⁴.

¹⁴ [JC_2023_78_CP_on ITS on ESAP tasks of collection bodies and ESAP functionalities.pdf \(europa.eu\)](#)

6 Annexes

6.1 Market Scan of Machine-Readable Data Formats

Entities	Data Formats	Comments
Australian Securities and Investments Commission (ASIC)	XBRL iXBRL PDF	<p>The ASIC has set forth a financial reporting taxonomy which allows digital financial reports to be provided in XBRL or iXBRL format.</p> <p>Starting from 2015, listed companies are required to submit the XBRL file and PDF file simultaneously to the Ministry of Commerce and Industry (MCI) annually.</p>
Bank of Japan (BOJ)	XBRL iXBRL	<p>The Japan Exchange Group (JPX) uses iXBRL to view XBRL data on the Company Announcements Service screen in XHTML using a web browser without dedicated software.</p> <p>Bank of Japan is using XBRL for its financial data reporting. (2006)</p>
Bank of Korea (BOK)	XBRL XML	<p>South Korea has established new requirements in 2023 such as requirements to include the notes to the accounts in XBRL formats, digital reporting in XBRL for financial firms, and the inclusion of financial statements from large unlisted companies in XBRL.</p> <p>In the context of ISO 20022, KOB is using XML as a basis for its message standards format, including payment and settlement, securities, and foreign exchange. The use of the markup language facilitates the expansion of the information contained in the message.</p>
Bank of Mauritius (BOM)	XBRL	<p>The bank of Mauritius has implemented a XBRL-based e-filing system for compliance reporting of all the entities under its jurisdiction, including banks, non-banking deposit-taking institutions,</p>

		money changers, cash lenders, and cash dealers.
Canadian Securities Administrators (CSA)	XBRL	The CSA has established a program to allow issuers to voluntarily file their financial statements in XBRL format on the Canadian electronic filing system, SEDAR (System for Electronic Document Analysis and Retrieval)
Dutch Government	iXBRL	The Dutch government and institutions are using iXBRL for business reports with a financial component, also referred to as the Standard Business Reporting (SBR).
HM Revenue & Customs (HMRC)	XBRL iXBRL	UK businesses, including non-resident companies, are required to file a Company Tax Return to HMRC, in the XBRL data format and in iXBRL, the standard form of presentation of business reports tagged in XBRL.
Hong Kong Monetary Authority (HKMA)	iXBRL XML CSV JSON	<p>The Inland Revenue Department (IRD) is using iXBRL for uploading financial statements and tax computation.</p> <p>The HKMA also uses XML and CSV for trade information submission to suit Trade Repository Members of different levels of technological sophistication.</p> <p>Finally, the HKMA uses JSON as an access to data through APIs.</p>
Israel Securities Authority (ISA)	iXBRL	ISA has proposed the introduction of mandatory reporting digital reporting in iXBRL from 2024. In 2023, the iXBRL format was already one of the options available for publishing financial reports in PDF format.
Japan Financial Services Commission	XBRL iXBRL	The Financial Services Agency (FSA) uses EDINET (Electronic Disclosure for Investors' NETwork) system. All filers are, in principle, mandated to submit in XBRL format the financial statements included in their Annual Securities Reports, Semi-annual Securities Reports, Quarterly Securities Reports and Securities

		Registration Statements for fiscal years starting in or after April 2008.
Mauritius Financial Services Commission	XML CSV Docx	All reporting issuers have to file their annual reports/audited financial statements and quarterly financial statements through the FSC Online Data Capture System Platform (ODSC), using either XML, CSV, or Docx formats.
Ontario Security Commission (OSC)	PDF	The OSC uses a PDF format for Capital and financial requirements (e.g., Form 13-502F4 for firms registered under the Securities Act)
Reserve Bank of India (RBI)	XBRL	The Institute of Chartered Accountants of India (ICAI) has been tasked with forming a national jurisdiction of the standards for financial reporting, and the Reserve Bank of India is responsible for implementing the XBRL standard for banks' reporting.
Saudi Arabian Monetary Authority (SAMA)	XBRL iXBRL PDF	Starting from 2015, listed companies are required to submit the XBRL file and PDF file simultaneously to the Ministry of Commerce and Industry (MCI) annually.
Securities and Exchange Commission (SEC)	XBRL iXBRL XML PDF	<p>The SEC uses several formats for reporting:</p> <ul style="list-style-type: none"> - XBRL is used for Financial Statement Data Sets, Financial Statement and Notes Data Sets, Mutual Fund Prospectus Risk/Return Summary Data Sets. - iXBRL is used to file cover pages, financial statement information, and certain non-IPO registration statements for Operating Companies. Open-end funds can also use this to tag the risk/return summaries in their filing. - The SEC requires the following forms to be filed in the XML format: Form SF-1/SF-3, Form ABS-EE, Form 10-D, Schedules 13D & 13G, Form N-CEN, and Form 13F. - PDF is used for Alternative Trading System ("ATS") List, and other lists based on text only. <p>On March 6, 2024, the SEC adopted rules to mandate climate-related disclosures for issuers</p>

		and has chosen iXBRL as the required format. The SEC has mandated iXBRL tagging for the rule, noting that it will enable automated extraction and analysis of the information, allowing investors and other market participants to efficiently identify responsive disclosure, as well as to perform large-scale aggregation, comparison, filtering, and other analyses of this information across registrants. The SEC received comments that continuing with the iXBRL would ease compliance as investors and filers are already familiar with the language.
ESMA (EU)	iXBRL	ESMA requires the use of iXBRL for the European Single Electronic Format (ESEF). ESEF is the electronic reporting format in which issuers whose securities are admitted to trading on EU regulated markets must prepare their annual financial reports to facilitate accessibility, analysis and comparability of annual financial reports.
EBA (EU)	XBRL CSV (iXBRL)	EBA is proposing XBRL-CSV as the Pillar 3 reporting data format. iXBRL is being investigated for qualitative data fields. (Information shared during the public hearing on January 23, 2024)
EIOPA (EU)	XBRL	Following a decision from the Board of Supervisors on Collection of Information, EIOPA mandated the use of XBRL as the standard for reporting data submission between EIOPA and NCAs.
South Africa Companies and Intellectual Property Commission (CIPC)	XBRL	The services of an external consulting firm were enlisted to develop the taxonomy and its comprehensive architecture to support the implementation of the program for digitizing financial reporting through the open data exchange standard XBRL.
Switzerland Task Force on Climate-	iXBRL	While no decisions have been taken yet at the time of the study, the Swiss TCFD has highlighted the need for a machine-readable climate disclosure through an ordinance that

related Financial Disclosures (TCFD)		should come into force in January 2024. For this purpose, the task force has mentioned the use of iXBRL but more discussions will be required before a final decision is taken.
Taiwan Stock Exchange	XBRL iXBRL	To keep pace with the international trend and promote the internationalization of financial information, Taiwan’s capital market adopted the XBRL format for financial reporting since the second quarter of 2010. Over the recent years, further development to move towards iXBRL have been made.

6.2 PDF and Machine-readability

61. Following a request from ESMA, the study also investigated the machine-readability of PDFs. The following subchapter provide elements in that regard.

62. As stated in the evaluation, the standard PDF format cannot be considered as machine-readable. However, our limited market scan allowed us to identify a constraint form of Adobe PDF version 1.7, used in the context of ISO-19005-3, which meets the machine-readability criteria set forth by the Open Data Directive (EU 2019/1024).

63. The Library of Congress defines PDF/A3 has “a constrained form of Adobe PDF version 1.7 (as defined in ISO 32000-1) intended to be suitable for archiving of page-oriented documents for which PDF is already being used in practice. PDF/A-3 adds a single and highly significant feature to its predecessor PDF/A-2 (ISO 19005-2) specification, to permit the embedding within a PDF/A file a file, or files, in any other format, not just other PDF/A files (as permitted in PDF/A-2).”¹⁵

64. The limited market scan allowed us to identify existing use cases for PDF/A3:

- **XML Schema imbedded in PDF files**¹⁶: German ZUGFeRD standard for electronic invoices, in which a visible, printable PDF document has a machine-processable version of the invoice based on an XML schema embedded in the PDF file.

¹⁵ <https://www.loc.gov/preservation/digital/formats/fdd/fdd000360.shtml#:~:text=PDF%2FA%2D3%20is%20a,already%20being%20used%20in%20practice.>

¹⁶ https://fnfe-mpe.org/factor-x/factor-x_en/

- **V3 Framework**¹⁷: V3 uses an XML document as the master format from which plain text, HTML, and PDF versions are derived. The PDF is a PDF/A-3u document with the XML master embedded. The first RFC published in the new format was RFC 8650, published in November 2019.
- **Manufacturing**¹⁸: A PDF/A-3 file can present an interactive 3D model in the document and related files can be embedded, creating what is sometimes called a "technical data package" or TDP. Changes introduced with PDF/A-4e, a PDF/A-4 profile intended as a successor to the PDF/E-1 standard, are likely to increase adoption of PDF/A in this domain.
- **Hybrid Archiving**¹⁹: If every revision of a document is accompanied by storing a PDF/A-3 file with the source word-processing file embedded, the document is archive-ready whenever editing stops.
- **Embedding richer metadata in "native" discipline-specific or application-specific format**²⁰: Although PDF/A already requires and supports XMP, many metadata schemes have rich XML representations that are not easily converted to RDF, on which XMP is based. However, rich metadata in a well-known application-specific format could be embedded and associated with the document as a whole - in addition to XMP metadata using built-in schema.
- **Ministry of Transport of Quebec**²¹: In April 2018 Dietrich von Seggern described use cases for PDF/A-3 where the Ministry of Transport of Quebec to archive engineering documents as PDF/A-3 with embedded CAD files and presented the use case of what he terms a "digital dossier," using PDF/A-3 as a container for project documentation.

65. The agreed scope of the study focused on the analysis of standard data formats. Given the information provided above, we cannot consider PDF/A3 as a standard format and thus cannot include it in the list of data formats to be evaluated.

¹⁷ <https://www.loc.gov/preservation/digital/formats/fdd/fdd000360.shtml>

¹⁸ <https://www.loc.gov/preservation/digital/formats/fdd/fdd000360.shtml>

¹⁹ <https://www.loc.gov/preservation/digital/formats/fdd/fdd000360.shtml>

²⁰ <https://www.loc.gov/preservation/digital/formats/fdd/fdd000360.shtml>

²¹ <https://www.loc.gov/preservation/digital/formats/fdd/fdd000360.shtml>