CHEMICAL ENGINEERING TRANSACTIONS
VOL. 107, 2023
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it
Guest Editors: Petar S. Varbanov, Bohong Wang, Petro Kapustenko

ISBN 979-12-81206-07-6; **ISSN** 2283-9216
DOI: 10.3303/CET23107097

Overview of XBRL Taxonomy Usage for Structured Sustainability Reporting in European Filings

Alex Suta*, Orsolya Pintes, Péter Molnár, Bence Lukács, Loránd Kedves, Árpád Tóth

Széchenyi István University, Vehicle Industry Research Center, H-9026 Győr, Egyetem Sq. 1, IS-201, Hungary
suta.alex@ga.sze.hu

The increasing requirement for businesses to disclose sustainability information digitally has prompted significant changes in the content and format of Environmental, Social, and Governance (ESG) disclosures. However, companies mandated to adapt to these changes face technological and information challenges regarding 'what' and 'how' to report. For European filers, the Corporate Sustainability Reporting Directive (CSRD) and its requirements, the European Sustainability Reporting Standards (ESRS), along with the International Financial Reporting Standards (IFRS S1 and S2), propose the use of the eXtensible Business Reporting Language (XBRL) as the anticipated technical solution for the digital data structure. The objective of this paper is to provide a methodological framework for effectively navigating the complex and interrelated concepts relevant to stakeholders. Rather than relying on cumbersome textual guides, this framework leverages an examination of existing taxonomies to offer readers insights into the essential glossary of disclosures and metrics considered crucial by official regulatory sources. Furthermore, the research discusses the emphasis on qualitative and narrative disclosures in ESG reporting and their feasibility of comparable results. Employing this methodology facilitates the implementation of corporate case studies and enables the analysis of mass amounts of future annual reports for comprehensive sustainability performance measurement.

1. Introduction

In recent years, the growing emphasis on sustainability and responsible business practices has led to increasing pressure on companies to disclose comprehensive environmental, social, and governance (ESG) information. This demand for transparency and accountability, especially in the case of large companies, stems from a global shift towards more sustainable and socially responsible business models and innovative practices (Ritala et al., 2018). As a result, companies are expected to communicate externally via not only traditional financial-, but also detailed sustainability reports in various styles and formats to address stakeholder concerns and demonstrate their commitment to sustainable practices (Hossain et al., 2022). This information enables the understanding of the positive and negative sustainability-related impacts of companies and how they affect their development, performance, and position (EFRAG, 2022). Different regions have developed their own reporting requirements, adding to the complexity of multinational companies. Taxonomies were developed to streamline the reporting content and technical format, of which Jastrzębska (2023) thoroughly analyzed the acceleration of European standard setters, publishing the first set of European Sustainability Reporting Standards (ESRS).

For the purposes of this study, we focus on the European region, which is governed by the Corporate Sustainability Reporting Directive (CSRD) and, for the majority of large companies, the International Financial Reporting Standards (IFRS). It is worth noting that due to mainly internal managerial purposes, companies could use additional reporting standards, such as the Global Reporting Initiative (GRI) and the Sustainability Accounting Standards Board (SASB), which are used alongside regional requirements to meet the needs of different stakeholders (Pizzi et al., 2022).

At the same time, the transition from traditional paper-based reporting to digital corporate reporting has been a transformative process, offering significant benefits to companies and stakeholders alike. Enhanced opportunities are being introduced to standardize and rationalize sustainability disclosures with the introduction

Paper Received: 17 April 2023; Revised: 19 June 2023; Accepted: 23 October 2023
Please cite this article as: Suta A., Pintes O., Molnár P., Lukács B., Kedves L., Tóth Á., 2023, Overview of XBRL Taxonomy Usage for Structured Sustainability Reporting in European Filings, Chemical Engineering Transactions, 107, 577-582 DOI:10.3303/CET23107097

of digital taxonomies. The technological medium of current digital reporting has consolidated in the eXtensible Business Reporting Language (XBRL), which has become the superior standard globally for streamlining data exchange and improving data accuracy in both financial and sustainability disclosures. XBRL is an extension of the Extensible Mark-up Language (XML), developed over 20 y ago to possess a tagging feature on otherwise moderately structured business information (Mousa and Ozili, 2023). XBRL provides a clear and stable document structure, a set of default file extensions, and a standardized method to describe and refer to the taxonomies (XBRL International Inc., 2023). However, along with its benefits, this transition has presented companies with a prolonged and relatively slow adoption phase and also technical challenges, particularly in navigating the complexities of sustainability reporting, such as determining 'what' data to report (e.g., the effect of sustainable aspects on financial outcomes and vice versa) and 'how' (i.e., method of technology integration in internal strategy and operational frameworks) to communicate it effectively (Mora-Rodriguez et al., 2017).

There has been significant research effort on the observation of such practices to generate insights into the ESG performance of individual companies and provide an objective comparative analysis with the use of XBRL reports (Faccia et al., 2021). On the other hand, when digital corporate reports are not available, researchers have tried to consolidate the report disclosure content by introducing standardization measures, often supported by *ex-post* methodologies. These methods define a system of sustainability disclosures by retrieving key concepts directly from sustainability reports already published. This often includes disclosures to be evaluated using pre-defined taxonomies, for instance, by manual coding to define quality measures by the content and format of integrated reports (Lueg, 2022). Usually, unique text mining dictionaries were created based on keyword retrieval from reports and matching to sustainability concepts, e.g., official categories of 'Governance', 'Strategy', 'Risk Management', and 'Metrics and Targets' (Moreno and Caminero, 2022), or keywords implying different sustainability concepts (Tóth and Suta, 2021). Multiple taxonomies could be used in the creation of dictionaries, e.g. Jiang et al. (2023) used GRI standards and the United Nations' Sustainable Development Goals (SDGs) to define word combinations (N-grams) to implement on annual reports. Alternatively, case studies could be used to develop taxonomies to analyze specific processes from sustainability aspects (Tahir and Darton, 2010). While taxonomy development could be based on Community regulations and directives (CSRD and its predecessor NFRD) or global scope (TCFD) (Andersson and Arvidsson, 2022), heuristics are often incorporated, which ultimately limits standard outcomes.

In addition, as similar research often uses text documents as data sources, both manual and hybrid data processing methods can result in a heavy workload, thus limiting the sample size. In the drive towards comprehensive sustainability reporting, there is a research gap for a standardized methodology to effectively manage the intricacies of ESG disclosure. XBRL, coupled with recent EU regulatory efforts towards digital ESG taxonomies, addresses this critical gap to enable companies to present coherent and comparable sustainability information to stakeholders and ensure meaningful analysis and decision-making.

This research aims to contribute to the existing knowledge by proposing a methodological framework that uses digital taxonomies, in particular the ESRS and the International Financial Reporting Standards (IFRS S1, S2), to streamline ESG reporting. We employ an *ex-ante* approach, by analyzing the technological background and reporting taxonomies that guide companies through the official publication of their ESG disclosures. In addition, the study explores the challenges and feasibility of incorporating qualitative and narrative disclosures into the digital reporting landscape. For practical knowledge, one case study is carried out by the observation of the integrated annual statement of Aviva Co., a major UK-listed oil-industry company, which is considered among the first report preparers to publish their sustainability information in an XBRL-tagged format (XBRL International Inc., 2022). In doing so, the research aims to facilitate the implementation of corporate case studies and enable stakeholders to comprehensively measure the sustainability performance of different European companies.

While previous research has focused on the development of solutions for processing XBRL files, there is a lack of descriptive research on compiling and analyzing digital sustainability reports. This is due to a lack of clarity in taxonomy data and file structure for both preparers and consumers. From a technical point of view, the possible method of ESG disclosures was examined. Methodologies for ESG disclosures do not follow the structured format facilitated by the XBRL framework (Faccia et al., 2021).

2. Material and Methods

Rather than creating our own taxonomy, the authors used a current regulatory reporting prototype taxonomy, the ESRS Proof-of-Concept (EFRAG, 2022), which outlines reporting requirements in the desired data structure. It no longer comprises simple recommendations but is now required by law for a large group of. While the content of the taxonomy is not yet finalized, understanding the implementation dates, scope, and XBRL data structure can provide a high level of understanding of sustainability reporting obligations, including checklists for content, possible links between data points, and technical rules of disclosure, illustrated in Figure 1.

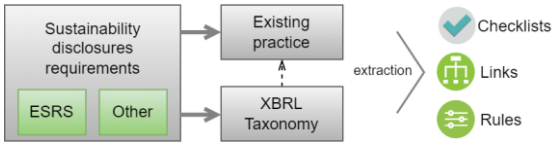

Figure 1: Outline of research design

2.1 Effective Dates of Adoption

The European Commission's CSRD, which came into effect on January 5, 2023, has introduced a mandatory sustainability reporting scheme for large companies (250+ employees) and listed SMEs in Europe. The implementation of the CSRD follows a phased approach, commencing with NFRD reporters in 2024, other CSRD stakeholders in 2025, and listed SMEs in 2026. To ensure standardized and comprehensive reporting, the CSRD–in the matter of defined reporting disclosures–aligns with the ESRS, covering Environmental (E), Social (S), and Governance (G) aspects. Concurrently, the International Sustainability Standards Board (ISSB) finalized the IFRS S2 Standard on June 26, 2023, facilitating global interoperability. Embracing digital transformation, sustainability reporting will transition to a structured digital format with the European Single Electronic Format (ESEF) requirements, starting from the first publication in 2025. These progressive measures collectively drive transparency, comparability, and sustainable practices across European enterprises.

2.2 Source and Processing of Data

An Officially Appointed Mechanism (OAM) is the term used throughout Europe to describe national databases for regulated financial information. OAMs are located in each member state of the European Union and serve as a mechanism for companies to file their financial reports with their regulator. To date, there is no requirement to produce sustainability reports in an XBRL-tagged format, but it is expected that companies will have to provide OAMs with filing requirements similar to those for financial statements.
In order to do this, the XBRL ecosystem has seen the creation of several tools by business (for-profit and not-for-profit), legal (government and regulatory), and academic entities. In general, we can assume that:

- For-profit tools are more reliable but focus on business opportunities by providing tools and services to publishers (report-creating companies) and regulators (who process those reports). They are closed products, not for general research.
- Non-profit, community solutions emerge from generalizing and sharing local developments. That leads to a large variety of overlapping solutions with serious dependency on external components and still a significant amount of code. Integrating or improving such solutions is time-consuming, with limited reliability.

To help clarify the landscape for using the framework, XBRL International maintains and continually updates a list of XBRL Certified Software™ that conforms to the XBRL specifications in three functional areas: Validating Processors (11); Report Creation Software (33), Review and Consumption Software (17); in addition, other non-certified tools (33) can be found online.

2.3 XBRL Data Structure for Sustainability Reporting

To carry out standardized content analysis on an XBRL Taxonomy, a processor is required. Processing can be done by using pre-built software or by own development if familiar with the technical details. This research involved writing a processing code in the Java environment, which aggregates the relationships between data points in standard taxonomies and reports. Simplified Excel spreadsheets were exported for further inspection of results, which were also made available publicly (Suta and Kedves, 2023). This code organizes the disclosures required by companies into a list that can be interpreted by anyone and shows the specificity of these disclosures for sustainability information. XBRL has a more organized and structured data format compared to traditional spreadsheet-based approaches, which are not reduced to accommodating rows and columns. This data structuring can be done in two ways. One way is by organizing data into tree structures, which is a top-down, hierarchy-oriented approach typically used for taxonomies and values structured according to accounting rules. This approach provides automatic validation for user-supplied numbers. As the standard has evolved, the need to subdivide existing data has arisen, leading to the inclusion of *dimension items*. Additionally, the *hypercube* data format introduced a higher dimensional approach to organizing data points, whereby new data, which has a cloud-like organization, are predefined and can be linked with the existing data.

3. Results and Discussion

3.1 Technical Analysis of the ESRS Taxonomy

In the XBRL framework, the intricacies of the taxonomies used for sustainability reporting facilitate accurate and standardized data exchange. A key aspect is the communication between XML Schema Definition (XSD) and linkbase files (including the label linkbase file), which form the backbone of the XBRL framework.
In this context, linkbase files act as a key component by providing additional information to the elements defined in XSD. Specifically, they augment the elements with essential contextual details that enable a more complete understanding of the reported data. For example, the 'label' attribute in the linkbase files supplements the concepts with explanatory text, presentation formats, and other relevant details. These labels play a critical role in conveying information to both XBRL-aware readers and stakeholders, promoting transparency and facilitating effective data interpretation.
In addition, linkbase files establish relationships through 'arcs' that link concepts and their associated labels. This structure enhances the navigability of the taxonomy and enables a coherent representation of the complex framework of the different meanings of the same core concepts. This is done by linking tags that serve as primary concepts (and represent values) to tags that represent certain labels. Figure 2 presents a total of 72 tags that were extracted from the ESRS Taxonomy, categorized by ‘substitution groups’ (in parentheses), and the number of each data type in the respective groups in bullet points. The number of data types and items is expected to increase as the ESRS Taxonomy becomes more prepared despite XBRL's fixed data structure.

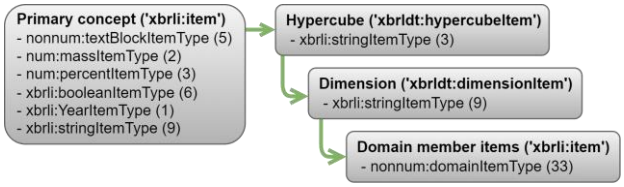

Figure 2: Linking primary concepts, hypercubes, dimensions, and domain member items in the definition linkbase of the taxonomy, based on (Burger and Roos, 2009)

Within the taxonomy, different ‘substitution groups’ are used to capture sustainability information content via being linked through ‘arcs’. For example, the ‘xbrli:item’ group of tags is commonly used and can be assigned a value indicating both quantitative and qualitative (textual) data reported. Part of the group serves as primary concepts, but most as domain items, which allows the inclusion of specific items from domain value sets, often representing hierarchical relationships. This is particularly relevant for the reporting of emissions, where a scope 2 emission element may contain several member items from an emission-specific domain value set.

3.2 Possible Structure of Data Points for Tagging Sustainability Information

Depending on their type, the 72 members shown in Figure 2 can have different link combinations. To effectively organize and categorize data, XBRL taxonomies use dimensions (‘xbrldt:dimensionItem’) and hypercubes (‘xbrldt:hypercubeItem’). Dimensions categorize or classify data items based on specific attributes or characteristics. Dimensions can be understood as 'tags' that are applied to primary concepts to provide further context and granularity to the reported information that creates cross-tabulation in the end report. For instance, let's take a primary concept such as 'Mass of GHG Emissions', which represents the total greenhouse gas emissions reported by a company. Dimensions can be used to categorize this primary concept based on diverse attributes or characteristics. One possible dimension would be 'Emission Scope,' where the company specifies whether the emissions fall under Scope 1, Scope 2, or Scope 3. Another conceivable dimension is 'Geographic Region,' which enables the company to indicate the region or country where the emissions were generated.
On the other hand, hypercubes are a uniquely structured representation of multiple dimensions that facilitate the analysis of multidimensional data. Hypercubes are particularly useful for presenting complex and interrelated information that requires effective categorization of data across multiple dimensions. For instance, in the case of "Mass of GHG Emissions," a hypercube can be created to offer a multidimensional perspective on emissions data. Such a hypercube would be based on dimensions such as "Emission Scope" and "Geographic Region". Through the use of a hypercube, companies can report emissions data categorized by different emission scopes, such as Scope 1, Scope 2, and Scope 3, and further subcategorized by specific geographic regions, such as countries or continents. Thus, the key difference between the two item type is that dimensions provide a single aspect of categorization, while hypercubes combine multiple dimensions to provide a more detailed and comprehensive multidimensional view of data.

Domain items in XBRL play a crucial role in connecting dimensions and hypercubes to primary concepts in the taxonomy. They provide a way to specify the allowable values for dimensions and hypercubes, effectively defining the domain or range of possible categorizations. When reporting data in iXBRL format, domain items are typically used as column headers in tables, allowing stakeholders to understand the breakdown of information based on various dimensions or categories. Each data point type has a set of possible tags that, when combined in various ways, can represent different sustainability content for disclosure. Figure 3 shows an illustration of the possible information that can be derived from the ESRS Taxonomy.

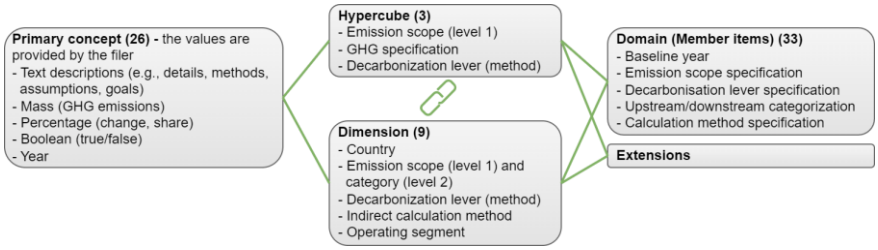

Figure 3: Indicative example of the sustainability information content of ESRS PoC Taxonomy tags

A notable feature of XBRL taxonomies used for sustainability reporting is the inclusion of external item lists, such as those sourced from Eurofilings, as value sets for dimension items. This feature increases the versatility and adaptability of the taxonomy, enabling seamless integration of relevant external data sources to enrich reporting. It is possible to distinguish tags by their source, which is contained in the tag name (which consists of a 'namespace', and a 'taxonomy concept', e.g., 'esrs'), usually the taxonomy of the accounting standard or the organization issuing it. Custom tags, commonly referred to as extensions are also available for reporting entities, but should not be overused as tags are expected to cover all aspects of necessary information. The structure of XBRL reports may contain additional information, such as for statistical or sustainability purposes. Sustainability information could be diversified by regional factors and subclasses, such as company divisions, product classes, or types of business activities. The data presented in firms' annual reports is challenging to locate and compile, but the structured format of XBRL makes it manageable. The standardized foundation of ESG disclosures also has the potential for wider application, such as creating more consistent ESG ratings calculations.

3.3 Data Point Analysis of Aviva Plc.

To gain insight into reporting practices, Aviva Plc's 2021 and 2022 statements were reviewed. Aviva Plc has utilized a forward-looking integrated accounting methodology that includes complete financial and sustainability disclosures for the year, following the IFRS regulations. Moreover, the company has voluntarily tagged this information according to the XBRL standard. They have utilized the taxonomy 'direp' (meaning Director's Report) proposed by the UK's Financial Reporting Council (FRC). This serves as an excellent example for observing data point relationships, following the examples provided in the previous section.
According to its 2021 report, the company disclosed 1,081 numeric and 160 non-numeric items. Among those, 207 numeric (19.2 %) and 41 non-numeric (25.6 %) items were sustainability-related. For both the preceding year (2020) and the business year (2021), each value was appropriately disclosed. The most commonly utilized dimension item was the reporting region, representing 57.0 % of numeric values, but none of textual. This dimension consisted of three distinct member items, including the regional classification of (1) UK, (2) Offshore, and (3) Total. Due to this classification and the data representing two years, only 10 unique numeric indicators were reported: Total Carbon Offsets, Emissions (Direct, Indirect, Other Indirect, Gross, Net), Energy Consumption, Intensity Ratio, and Pay Ratio. The tags could appear multiple times. As a result, they can be mentioned several times in the reports. Regarding textual data, 22 distinct tags were disclosed, with only a few instances of duplication (an average of 1.9 occurrences). The length of tagged text varied: the shortest text block was only 5 characters, and the longest was over 5,400 characters with an average length of 723.6 characters, which could be further evaluated using a fog index. The longest text blocks were tagged with the phrase 'Disclosure Metrics Used to Assess Climate-related Risks Opportunities In Line With Strategy Risk Management Process'.

4. Conclusion

XBRL taxonomies play a critical role in rationalizing sustainability reporting by providing a standardized and structured framework for data representation. The inclusion of dimensions, hypercubes, and external element lists helps to improve data organization and comparability. In addition, the communication between the

standardized files (XSD and linkbases), with the addition of explanatory labels (data points) and hierarchical relationships, ensures clarity and precision in the communication of sustainability information. By leveraging these advanced features, XBRL taxonomies enable organizations to effectively report and analyze comprehensive sustainability performance data for informed decision-making and stakeholder engagement. This research provided a mostly technical overview of the evolving European XBRL Taxonomy for sustainability reporting, which is essential due to upcoming adoption deadlines. Although the ESRS Taxonomy is not yet finalized, it presents an opportunity to comprehend persistent data formats and structures. Voluntary sustainability disclosures of companies could be examined further, similar to the instance provided with the data point analysis of Aviva Plc. Currently, our research is concentrated on developing and publishing an open-source data processing pipeline. This effort aims to improve the mass analysis of digital reports, including both financial and sustainability-related statements.

Acknowledgment

The research was supported by the European Union within the framework of the National Laboratory for Artificial Intelligence (RRF-2.3.1-21-2022-00004).

References

Andersson F.N., Arvidsson S., 2022, Understanding, mapping and reporting of climate-related risks among listed firms in Sweden. Climate Policy, DOI: 10.1080/14693062.2022.2116383.

Burger J.M., Roos M.R., 2009, Potential of XBRL Dimensions for International Trade Statistics. <https://www.cbs.nl/-/media/imported/documents/2009/13/2009-25-x10-pub.pdf?la=nl-nl>, accessed 16.10.2023.

EFRAG, 2022, [draft] European Sustainability Reporting Standards ESRS 1—General requirements. <efrag.org/lab6>, accessed 25.09.2023.

Faccia A., Manni F., Capitanio F., 2021, Mandatory ESG reporting and XBRL taxonomies combination: ESG ratings and income statement, a sustainable value-added disclosure. Sustainability, 13(16), 8876.

Hossain M., Islam M.T., Momin M.A., Nahar S., Alam M.S., 2022, Understanding communication of sustainability reporting: Application of symbolic convergence theory (SCT), In: Martin K., Shilton K., Smith J. (Eds.) Business and the Ethical Implications of Technology, 257-280, Springer Nature, Cham, Switzerland.

Jastrzębska E., 2023, The reporting of climate-related information by CSR leaders in Poland: a climate-related disclosures index. Economics and Environment, 84(1), 96-115.

Jiang L., Gu Y., Dai J., 2023, Environmental, Social, and Governance Taxonomy Simplification: A Hybrid Text Mining Approach. Journal of Emerging Technologies in Accounting, 20(1), 305-325.

Lueg R., 2022, Constructs for assessing integrated reports-Testing the predictive validity of a taxonomy for organization size, industry, and performance. Sustainability, 14(12), 7206.

Mora-Rodriguez M., Atemezing G.A., Preist C., 2017, Adopting semantic technologies for effective corporate transparency. In: The Semantic Web: 14th International Conference, ESWC 2017, Portorož, Slovenia, May 28–June 1, 2017, Proceedings, Part I(14), 655-670.

Moreno A.I., Caminero T., 2022, Application of text mining to the analysis of climate-related disclosures. International Review of Financial Analysis, 83, 102307.

Mousa R., Ozili P.K., 2022, A Futuristic View of Using XBRL Technology in Non-Financial Sustainability Reporting: The Case of the FDIC. Journal of Risk and Financial Management, 16(1), 1.

Pizzi S., Principale S., De Nuccio E., 2022, Material sustainability information and reporting standards. Exploring the differences between GRI and SASB. Meditari Accountancy Research, DOI: 10.1108/MEDAR-11-2021-1486.

Ritala P., Huotari P., Bocken N., Albareda L., Puumalainen K., 2018, Sustainable business model adoption among S&P 500 firms: A longitudinal content analysis study. Journal of Cleaner Production, 170, 216-226.

Suta A., Kedves L., 2023, Supplementary Material to "Overview of XBRL Taxonomy Usage for Structured Sustainability Reporting in European Filings", Zenodo.org, DOI: 10.5281/zenodo.8202556.

Tahir A.C., Darton R.C., 2010, Sustainability indicators: using the Process Analysis Method to select indicators for assessing production operations. Chemical Engineering Transactions, 21, 7-12.

Tóth Á., Suta A., 2021, Global sustainability reporting in the automotive industry via the eXtensible business reporting language. Chemical Engineering Transactions, 88, 1087-1092.

XBRL International Inc., 2022, UK's first Inline XBRL ESG report offers exciting new ways to analyse data. <xbrl.org/news/uks-first-inline-xbrl-esg-report-offers-exciting-new-ways-to-analyse-data>, accessed 25.09.2023.

XBRL International Inc., 2023, About XBRL. <filings.xbrl.org/about.html>, accessed 25.09.2023.