EUROPEAN SECURITIES MARKET AUTHORITY
Mr Eduardo-Javier Moral Prieto, Policy Officer
201-203 Rue de Bercy
75012 Paris

XBRL Europe BPTF: BLOCK TAGGING in ESEF reports

Dear Mr Moral Prieto,

Since our meeting during the XBRL Europe days in Paris the 14th of June, the ESEF BPTF has worked on the block tagging issue.

We have prepared a document summarizing the guidance we believe should be included in the ESEF reporting manual for block tagging, including the reasoning for the technical issue. Our concern is to reduce the possible different interpretations of the rules that would be confusing for issuers.

The document has been prepared by the XBRL Europe ESEF Best Practice Task Force chaired by Roger Haddad, Pierre Hamon and David Bell. It has been also worked, reviewed and endorsed by Paul Warren and John Turner, respectively Technical director and CEO of XBRL International

We understand this document comes quite late in your process of drafting the manual, but we would appreciate that you consider it in finalizing the manual. We certainly will have other comments to share with you in the future.

Yours sincerely

Gilles Maguet

CEO
XBRL Europe

XBRL Europe AISBL c/o Accountancy Europe
Avenue d'Auderghem, 22-28/8
B-1040 Brussels – Belgium - www.xbrleurope.org

BLOCK TAGGING in ESEF reports

Authors: Members of the XBRL Europe ESEF Best Practice Task Force
Date: July 11, 2022

This document aims to comment the new requirement for block tagging of the disclosures in the ESEF reports.

This topic warrants some consideration for guidance.
The document also explains the differences between the two technical options and their effect on text blocks in ESEF reports.

Table of content

I.   Our understandings of the requirement and guidance needed

MANDATORY:

1. All disclosures in the financial statements that match the definition of an element in table 2 must be tagged.
2. *If there is no separate disclosure in the notes, tags of table 2 may remain unused.*
3. The granularity of tagging follows the general rule for mark-ups: issuers should select the element with the narrowest accounting meaning and/or scope. Multiple tagging of the same disclosure or tagging or portions of tables should be discouraged.
4. There is no obligation to create extension to tag notes that do not correspond to any of the elements in Table 2 of Annex II.
5. There is no obligation to anchor extension elements to mark-up their Notes.

ALLOWED: voluntary tagging:

1. Nevertheless, ESMA **encourages** issuers to create extension blocktags since this information is useful to end users
2. If multiple pieces of text corresponding to one block tag are disclosed in different sections of the Notes, issuers **should** tag such disclosures with one block tag by using the Inline XBRL constructs which allow the concatenation of text content within a document
3. The requirement for block tagging **should not** limit the discretion of issuers to mark up notes to IFRS consolidated financialstatements with a higher level of granularity". ESMA highlights in this regard that when tagging additional information, using either detail or block tags.

GUIDANCE NEEDED

It should be made clear in the ESEF reporting manual:

1. The meaning of: "Issuers shall not apply the markups only partially or selectively"?
2. The meaning of: "since this information is useful to end users"? The information is disclosed because it is useful in the human readable document. What is the usefulness of a disclosure block tagged with an extension in XBRL?
3. Guidance related to double or multiple tagging and/or embedded tagging. In which cases should/shall it be used?
   - How to mark up a statement in the notes that covers several disclosure requirements within the same sentence? When should a piece of information be marked up with multiple block tags?
   - How to deal with text block tags that refer to the same information (but in different ways)? To comply with the ESEF regulation, would issuers have to use all block tags, as all are mandatory?
   - How to deal with multiple core taxonomy elements that refer to the same piece of information, however at a different level of detail?
   - Information in tables: Issuers should not be required to mark up information present in tables with blocktags.

II.   Proposed technical guidance

We propose that the technical guidance be written as follows:

Use of the "escape" attribute for textblock tags:

For *ix:nonNumeric*text block tags (type derived from dtr:textBlockItemType):
- the escape attribute MUST be set to true, if the human readable content contains a "<" or "&" character to ensure that the resulting fact value is always valid for its data type.
- In all other cases, the escape attribute can either be set to "false" or "true".

The resulting XHTML fragment extracted to an XBRL instance and rendered in a viewer or browser is likely to look substantively different than when rendered from the original tagged iXBRL.

Issuers SHOULD ensure that the text content of the tag faithfully reproduces the original report:
- Words and numbers SHOULD be in the right order.
- Where there is space between words and numbers in the source, there SHOULD be at least some space retained in the text block (i.e., "intangible assets €3m" SHOULD NOT become "intangibleassets€3m").

 No matter how the escape attribute is set, the following applies:
- It may not be possible to replicate the look and feel of the Inline XBRL document within text rendered from extracted text block. It may not replicate the appearance of the tagged content in iXBRL. For example:
  o Table structure may be lost
  Multi-column text flow may be lost
  Styles (fonts etc.) may be different
  Line breaks may be different, etc.

Therefore, it MUST NOT be a requirement that issuers use semantic XHTML tags (e.g. <table> or <p>) or that the XHTML in an extracted XBRL instance that is re-rendered in a viewer or browser replicates the appearance of the original exactly.

ESMA will be observing and analysing the textblock tags of the first year and might change those rules for the following reporting periods.

III.   Requirement in RTS

RTS: Article 4 Marking up IFRS consolidated financial statements
2. Issuers shall, as a minimum, mark up the disclosures specified in Annex II where those disclosures are present in those IFRS consolidated financial statements.
ANNEX II
Mandatory markups
3. Issuers shall mark up all disclosures made in IFRS consolidated financial statements or made by cross-reference therein to other parts of the annual financial reports for financial years beginning on or after 1 January 2022 that correspond to the elements in Table 2 of this Annex.

ANNEX IV
Marking up and filing rules

3. When marking up disclosures, issuers shall use the core taxonomy element with the closest accounting meaning to the disclosure being marked up. Where there appears to be a choice of core taxonomy elements, issuers should select the element with the narrowest accounting meaning and/or scope.

9. Issuers shall ensure that the issuer's extension taxonomy elements marking up the IFRS consolidated financial statements' statement of financial position, statement of profit or loss and other comprehensive income, statement of changes in equity and statement of cash flows are anchored to one or more core taxonomy elements.

13. When marking up disclosures, issuers shall use non-numeric taxonomy elements in a way that it marks up all disclosures that match the definition of the respective element. Issuers shall not apply the markups only partially or selectively.

IV.   Guidance in ESFE reporting manual

Guidance 1.3.3 Tagging elements of Table 1 and 2 of Annex II **[last updated: July2021]**

The RTS on ESEF requires that issuers shall mark up all disclosures that correspond to the elements in Table 1 and Table 2 of Annex II if those disclosures are present in the issuer's financial statement. If those disclosures are not present in the issuer's financial statement, they should not be tagged.
Consequently, there is also no obligation to create an extension to tag the notes to the Financial Statements if an issuer's disclosure does not correspond to any of the elements in Table 2 of Annex II. Nevertheless, ESMA encourages issuers to create extension block tags since this information is useful to end users. As noted in Guidance 1.4.1, there is no obligation to anchor extensions in the Notes to the financial statements.
However, if multiple pieces of text corresponding to one block tag are disclosed in different sections of the Notes, issuers should tag such disclosures with one block tag by using the Inline XBRL constructs which allow the concatenation of text content within a document (see Guidance 2.5.5).
As highlighted by recital ten of the RTS on ESEF, "the requirement for block tagging should

not limit the discretion of issuers to mark up notes to IFRS consolidated financial statements with a higher level of granularity". ESMA highlights in this regard that when tagging additional information, using either detail or block tags, issuers need to ensure consistency across reporting periods to the maximum possible extent.

Guidance 1.4.1 Anchoring of extension elements to elements in the ESEF taxonomy that are wider in scope or meaning [last updated: July 2019]

Please note that the RTS on ESEF does not set an anchoring requirement for the Notes to the financial statements. Therefore, if issuers decide on a voluntary basis to create detailed tag extension elements to mark-up their Notes, there is no obligation to anchor such extension elements.

V. Technical considerations

Intended goal of block tagging

The goal is that block tagging should be able to designate "meaningful fragments" of well-formed XHTML that are extracted into XBRL for processing.

This means that for the extracted information:

- when displayed outside the context of the original document, these fragments resemble the original document as the semantic structure (but not necessarily style) is maintained.
- machines have a better opportunity to consume and interpret the content because the semantic structure is present.

The goal of block tagging should be clearly stated:

e.g., "publish machine readable information in the form of text or tables", "Allow for text mining of block tags with reasonable clean-up efforts"

What was iXBRL tagging intended for?

- Enable document navigation
  - o e.g., search for standard, iXBRL tags in an iXBRL viewer and just to relevant disclosure
- Improve automated analysis
  - o AI / text analysis is more effective on structured data
- Viewing data out of context
  - o e.g., analytic software putting policies from different companies in a table side-by-side

Retaining formatting and styling is only relevant to the last case, and even then, too much styling may hinder rather than help.

Technical options

textBlockItemType

IFRS text block tags use textBlockItemType from the Data Types Registry, which is derived from xmlNodesItemType which requires that: "the unescaped content is a valid sequence of XML text and well-formed XML nodes."

Valid fact value: "<p>Profit & Loss</p>"

Invalid fact value: "Profit & Loss" (& is a special character in XML/XHTML), so this must be represented as "Profit & Loss"

The "escaped" attribute

XML special characters (e.g., & and <) MUST be escaped for the resulting fact value to be valid XHTML. For the value to be reliably valid for dtr:textBlockItemType, tags must have escape="true"

iXBRL provides two possible modes when creating text tags (the "escaped" attribute):
- Retain HTML tags (escaped="true")
- Text only (escaped="false")

- Example of escaped="true"

This controls what information gets included in the extracted XBRL fact's value.

Example of a tagged accounting policy:

Accounting policy:
This is the first paragraph of my accounting policy
This is the second paragraph of my accounting policy

Text only representation:

Accounting Policy This is the first paragraph of my accounting policy. This is the second paragraph of my accounting policy.

Retained HTML representation:

```
<h3>Accounting Policy</h3>
<p>
This is the <b>first</b> paragraph of my accounting policy.
</p>
<p>
This is the <b>second</b> paragraph of my accounting policy.
</p>
```

Retain HTML tags (escaped="true")

In theory, retaining HTML tags allows formatting to be retained when displaying the fact's value.

In practice: Fragments of HTML are not self-contained:

- May rely on styling information (CSS) from elsewhere in the document — or a different document
- Appearance may be influenced by styling of a parent element e.g. An HTML tag could specify white text colour within a parent HTML element with a blue background. If the child HTML tag is extracted without the parent, you may get white-on-white (invisible) text.
- May not be valid HTML e.g., extracting a single row from a table – a row (<tr>) tag is not valid without an enclosing table (<table>)

Validity consideration

The requirement for values to be valid XHTML is not widely enforced by processors (but it should be)

Nonetheless, it is important that the datatype is used consistently, as otherwise consumers have to guess whether it is intended to be treated as XHTML or not.

If a text block does not contain either "&" or "<" in the text, it can be tagged with escape="false" and the result will be a valid (unformatted) XML/XHTML string.

If tagged with escape="true", the results are guaranteed to be valid XML (Results are not guaranteed to be valid XHTML as tags may not be nested correctly (e.g., a <tr> tag not without an enclosing <table> tag), but it will be valid, well-formed, and correctly escaped XML.)

Technical tag construction

The filing manual does not currently define whether escaped HTML tags should be included in text block tags.

In the absence of specific guidance, it is likely that preparers will refer to existing practice in the US, where text blocks contain (escaped) HTML fragments. However, there are several reasons why the approach taken there will not work well for ESEF filings:

1. SEC filing rules require all styling information to be included using inline style elements, rather than through the use of a stylesheet.
2. ESEF filings generally have much more complex HTML and styling than their US SEC counterparts. These mean that fragments of HTML extracted from an ESEF iXBRL document are unlikely to render correctly outside of the document from which they are drawn. It is worth noting that the text block tagging approach taken by the SEC predates the adoption of Inline XBRL. Inline XBRL means that XBRL data is directly linked to the HTML document, providing an easy way for users to view a disclosure as it was originally prepared. As such, the need to replicate the presentation of data within a text block tag is greatly reduced.

Example of SEC filings - Interactive data:

- SEC documents have very plain styling compared to ESEF reports
- HTML is much simpler (much closer to "meaningful", semantic HTML)
- Inline styles are required
- History (text tag usage pre-dates iXBRL)

Note that SEC would never accept a PDF conversion and that, in addition to block tagging, the issuers are required to detail tag their financial report.

Property, plant and equipment [abstract]	
Property, plant and equipment	Property, plant and equipment

D.3.1. Property, plant and equipment owned

Property, plant and equipment owned by Sanofi is comprised of the following items:

(€ million)	Land	Buildings	Machinery and equipment
Gross value at January 1, 2019	283	6,883	10,468
Acquisitions and other increases	—	10	50
Disposals and other decreases	(3)	(42)	(148)
Currency translation differences	6	80	64
Transfers[a]	(31)	351	619
Gross value at December 31, 2019	255	7,282	11,053
Changes in scope of consolidation	—	6	3
Acquisitions and other increases	—	16	40
Disposals and other decreases	(11)	(173)	(177)
Currency translation differences	(13)	(264)	(276)
Transfers[a]	5	(39)	484
Gross value at December 31, 2020	236	6,828	11,127
Changes in scope of consolidation	—	11	15
Acquisitions and other increases	—	10	51
Disposals and other decreases	(3)	(75)	(153)
Currency translation differences	6	169	155
Transfers[a]	1	227	453
Gross value at December 31, 2021	240	7,170	11,648
Accumulated depreciation & impairment at January 1, 2019	(19)	(3,796)	(7,230)
Depreciation expense	—	(357)	(586)
Impairment losses, net of reversals	(4)	(33)	(4)
Disposals and other decreases	2	54	140
Currency translation differences	—	(40)	(40)
Transfers[a]	10	107	60
Accumulated depreciation & impairment at December 31, 2019	(11)	(4,065)	(7,660)
Depreciation expense	—	(356)	(605)
Impairment losses, net of reversals	—	(24)	(12)
Disposals and other decreases	1	168	166
Currency translation differences	—	127	169
Transfers[a]	—	252	150
Accumulated depreciation & impairment at December 31, 2020	(10)	(3,898)	(7,792)
Depreciation expense	—	(306)	(592)
Impairment losses, net of reversals	—	(3)	(22)
Disposals and other decreases	—	74	149
Currency translation differences	—	(80)	(99)
Transfers[a]	1	23	16
Accumulated depreciation & impairment at December 31, 2021	(9)	(4,190)	(8,340)
Carrying amount at December 31, 2019	244	3,217	3,393
Carrying amount at December 31, 2020	226	2,930	3,335
Carrying amount at December 31, 2021	231	2,980	3,308

(a) This line mainly comprises property, plant and equipment in process brought into service during the period, but also includes the effect of the reclassification of assets to Assets held for sale or exchange assets on first-time application of IFRS 16.

The table below sets forth acquisitions and capitalized interest by operating segment for the years ended December 31, 2021, 2020 and 2019:

Example of SEC filings - XHTML viewer

Full disclosure text block or table text block

Expectations

- **Preparers**
  The preparers expect to keep their glossy reports. Block tagging should have no major impact on creation process and tools. The thousand companies that tagged PDF as of today, will not accept doing something else on a short notice.

- **Auditors**
  There is expectation by Auditors about this is that the "snippets" will look exactly like in the complete report, something which does simply not work when following the ESMA recommendations about styles. The professional body of chartered accountants in the Netherlands do not demand that a block tagged snippet should look exactly as in the complete report, the demand is that the information value must be the same. Auditors should bring in new requirements, with no legal basis

- **Data users**
  What is expected by analysts using text mining tools? Do we expect that text mining tools are adopted to iXBRL in future? As of today, those using text block tags (e.g., analysts), have not stated they need for semantical XHTML.

E BPTF - BLOCK TAGGING in ESEF reports   **11 July 2022**

Current state of tools and technology

The current state of tooling is such that this goal is currently unattainable for a very large part of the market. This is because there is little or no semantic structure available within the extracted XBRL content.

Thus, in the immediate and short-term, it must be understood that the content of block tags will most likely neither be structured nor meaningful.

So, we and the rest of the community (filers, auditors, regulators), must accept that:

- to be "human" meaningful, the block tag content must be read in context within the original document.
- extracted block tag content may have little resemblance, visually and structurally, compared to that of the original document.
- it is of no value (and detrimental) to try to impose any constraints on extracted block-tag content that has no semantic structure.

Possible approaches

1. Require "text-only" mode (escape="false")
   - Some effort required from vendors to ensure spaces are present between words after tags are stripped.
   - Text is available for automated analysis.
   - No formatting, so any non-trivial text block will be hard to read out-of-context.
   - No formatting possible -— removes opportunity for argument about what is "good enough".
2. Require "retain HTML tags" mode (escape="true")
   - Without additional effort, results may be little better than "text-only".
   - Results will never be perfect but leaves room for argument about what is "good enough".
   - "Meaningful" HTML could give results that are substantially better than "text-only» but would require a fundamental change of approach for many software products and preparers.
3. Let preparers choose
   - If this option is taken, the requirement for the output to be valid XHTML must be retained (i.e., "<" and "&" cannot be tagged with escape="false")

Preserving whitespace

- Some HTML elements are inline elements.
- Inline elements add mark-up within a line of text and do not introduce any space.
  e.g., whitespace => whitespace
- CSS styling can be used to add spacing to such tags:
  e.g. white<b style="padding-left: 30px">space => white   space
- A text-reader would still see this as single word "whitespace"
- When creating text blocks, software should ensure that spacing between words actually exists, and is not created only by CSS styling.

VI.   Considerations for future improvements

"Meaningful" HTML

- Many ESEF documents are generated by automated conversion from PDF documents.
- This process faithfully reproduces the *appearance* of the PDF but does not produce "meaningful" HTML.
- HTML provides semantic tags such as <p> for paragraphs, <h1> for headings, <table> for tables.
- Where these semantic tags are used, HTML can be successfully "re-styled" (i.e., displayed outside of its original context)
- We will refer to documents that do this as "meaningful" HTML.
- PDF converters do not produce these semantic tags.
-

"Meaningful" HTML is an improvement, but is not a complete or easy fix

- Custom fonts are still lost (no good solution to this)
- CSS Styles (e.g., font colour) must be inlined
    o This bloats the HTML document compared to using re-usable CSS classes
- Using "meaningful", "reflowable" HTML may create problems with the layout of the original document
    o e.g., If line breaks occur in different places, it may push text over a page or column boundary
- Switching to "meaningful" HTML would require a fundamental change of approach for many existing ESEF solutions — and *preparation processes*
    o Not possible to generate "meaningful" HTML from a presentation format such as PDF as the semantic information simply is not present. PDF only contains information about what the document should look like. It does not contain semantic document structure (paragraphs, headings, tables, etc.)

Perfect, "out-of-context" replication of formatting and styles is not realistic

- What does it even mean if a note spans a page or column?
- Preparers, auditors, and software vendors could expend a lot of effort worrying about what constitutes "good enough" — time that would be better spent on tag selection and other semantic issues.

Other considerations

We believe that further analysis of created ESEF filings will be required to provide rules and guidance for achieving same content. Therefore, filings should be analysed after 1-2 years of block tagging experiences.