Taxonomy Task Team meeting summary

Date of meeting	August 12, 2025
Topic	Tags related to publishing content and framing the context
Background information	See the Tags related to publishing content and framing the context document on the website
Upcoming meetings	August 26: People, groups, and organizations Additional meetings to be scheduled

Many of the tags we discussed at this meeting are often invisible to the end user. For example, they may be automatically tagged behind the scenes based on the template a user chooses. However, these tags are extremely important in helping aggregators organize and display the content, and in helping end users find the research they need, so we appreciate members' input regarding these lists.

Below is a summary of the input we received before and during this meeting, along with answers to some questions that came up during the meeting requiring additional research. As with all meetings in this series, we are still looking for more input, so feel free to send your thoughts regarding the below.

Summary

PublishingActionType

This list is fairly straightforward. In many cases, the only PublishingAction a research report would receive would be *Published*; the remaining terms are available to capture any pre-publication statuses a firm needs internally, or for any updates or changes that occur post-publication. Since this list was created when most research content was published in print, we are particularly interested in assessing whether other research formats – interactive content, audio/video files like podcasts, etc. – require additional terms.

For firms that use RIXML "baked in" to their entire workflow (rather than adding tagging as a final step before distribution), we would like input regarding whether additional terms covering pre-publication stages would be helpful. We currently only provide pending, published, and released. If your firm has had to create workarounds to capture the stages of your publishing process – particularly pre-publishing – please let us know.

One participant asked for clarification regarding the difference between *Recalled* and *Deleted*. According to our 2.5 documentation, *Recalled* is used if the file was published as a mistake or before intended; *Deleted* is when it was published as intended, but the publisher has decided that it should later be removed.

Another question came up regarding the difference between *Published* and *Released: Published* indicates that the product is being published for the first time; *Released* is used to indicate the date/time that the product is being released. This would be used when the publication and release times are not the same.

The separation between *RevisedMaterialChange* and *RevisedNoMaterialChange* was well received, as allowing publishers and aggregators to distinguish between items where a spelling error has been corrected vs. a more important change can help determine whether a report should be displayed based on the original timestamp or the new timestamp, etc.

The lists involved with indicating what "kind" of report is being described

This topic made up a significant portion of the meeting. As described in the pre-meeting documentation, there are some tags that are used to indicate what "kind" of report something is – not the content of the report, but what format it is, what discipline was used, etc. These tags work together, so we looked at them together so that participants can consider the different kinds of content they publish and think through what terms they would pick in each list.

We presented an updated version of the chart at this meeting, adding the lists for at *EventVenueType* and *InteractionMode* because they have some overlap with the other lists in the chart. We will take this conversation up again when we discuss events and interactions, but including them in this meeting resulted in several helpful insights. The updated chart is:

ResearchProduct	Form	Event	Interaction	Discipline	
principalMedium	form	EventVenueType	InteractionMode	disciplineType	researchApproach
Written Audio Video Interactive Mixed	Comment Report Model Chart ** Compilation Presentation Event Primer Directory Glossary Overview EventInvitation PostEventSummary Podcast * Interview * PanelDiscussion * ExhibitExplainer * EventRecording * Marketing * TrainingMaterial * PublisherDefined *	ConferenceCall OneOnOneMeeting GroupMeeting Webcast WrittenRelease InternetAudio InternetVideo MediaAppearance Transcript Brief Podcast Interview Panel	Call DataFeed Email IM Inperson Model Product Service Video VoiceMail PublisherDefined	Investment Strategy Economics	Fundamental Quantitative Technical
* indicates new term ** indicates updated terminology for existing term (*Chart* was formerly *Charts*)					

Our goal is to create a set of tags that works for all our standards without unnecessary or unintended overlap, so we will be looking at these lists again when we discuss events and interactions. Taking a cohesive look at these different lists, and all of the types of research, events, and interactions that use them, will allow us to identify both gaps in the current lists and commonalities that may allow us to adjust or combine lists to create a more-intuitive end result.

Participants brought up a wide range of examples of research content their firm or their clients produce, allowing us to brainstorm about the terms needed to describe different kinds of research (that is, not simply single-company reports).

To summarize the discussion:

In the *principalMedium* attribute:

- CLARIFICATION: there can only be one principalMedium per research item, because it is an attribute of the top-level tag. It is an optional tag; however, we would like to ensure that the list includes terms that reflect the types of research content being published.
- Adding the term *Data* was recommended, for models and other research content that is mainly numbers-based vs. written analysis.

In the *form* attribute:

- Adding the terms *Transcript* and also *TranscriptSummary* were recommended.
- Adding the term *ValuationModel* was suggested. We will need to discuss this again to determine whether this value should be added, or if a valuation model would have form=*Model* with some mechanism to indicate what kind of model. Regardless, we will ensure that valuation models can be accurately tagged.
- CLARIFICATION: a participant wanted to know the difference between a *Comment* and a *Report*: a comment is meant to be a short, generally time-sensitive update – an analyst's take on news regarding a company, etc.; a report is meant to be a more in-depth piece of content.

In the *EventVenueType* and *InteractionMode* lists:

- *InvestorDay* and *Fieldtrip* were suggested. In reviewing the Interactions standard, these are both currently available in the InteractionPurpose tag set; we encourage firms that would like to tag for this type of content to review the Interactions Standard documentation (https://rixml.org/standards/interactions-standard) to learn more; we will discuss how InteractionMode and InteractionPurpose interact in more depth at a future meeting.
- *VirtualConference* was suggested. In reviewing the Interactions standard, that can be accomplished using InteractionMode=Video and InteractionPurpose=Conference; we will need to determine whether using these tags allows virtual conferences to be accurately identified.
- *FiresideChat* was suggested. This is not currently available as a value in any of the Events or Interactions tag sets. We will include this in the discussion document for events and interactions, to make sure we determine how best to identify this fairly common type of interaction. (We discussed that the term "fireside chat" is distinct from a panel discussion – less formal. As such, that may make it more valuable, so it would be helpful for the interactions standard to have a way to make this distinction. Note that neither *fireside chat* nor *panel discussion* are available terms in any tags in the interactions standard; however, *Panel* is an option in the EventVenueType list.
- The following additional interaction modes were suggested: APIs, research portal, screen share, chatbot. We will include these in the discussion document for events and interactions. **Question**: should any of these be available for research content? We will revisit this discussion when we discuss events and interactions and will assess how any overlapping terms should be handled.
- We need to discuss the types of non-company content that is published – macroeconomic analysis, thematic research, etc. – and make sure that we provide clear guidance on what tags should be used for various scenarios. We may also need to add terms in various lists to facilitate accurate tagging for non-company research, particularly for thematic research and content created using various AI-powered methodologies.
- For all research, we want to minimize having multiple valid ways to describe similar content; rather, either the tags themselves or the documentation should provide a single best practice whenever possible.

Background information regarding *Discipline*:

Sal provided some helpful information about the history of the *Discipline* tag set and its intended usage:

- the *Discipline* tag set was added in an early version of the research standard when the terms *Economics* and *Strategy* were proposed for the *Asset Class* list; while they are topics that are frequently written about, they are definitely not asset classes. Thus, the Discipline.disciplineType attribute was created to allow this content to be identified; the term "Investment" was added to mean "everything else" – including company research, sector overviews, etc.
    - QUESTION: in the interest of clarity, would it be beneficial to rename "Investment" to "Other" to increase likelihood that firms use this value in a consistent manner?
    - QUESTION: where would thematic research fall – primers, thought pieces, etc.?
- Thematic and cross-sector research has become more common, so we need to ensure that we provide sufficient guidance for tagging such content and that the terms available in our enumeration lists capture the types of content being produced.
- Quant and data driven reports are also more common, particularly with the introduction of machine learning model outputs – sentiment shifts, signal reports, etc.). We need to examine our enumeration lists to determine whether added granularity or specificity is needed to accurately describe this content.
- We discussed the fact that disciplineType and researchApproach are not meant to act in a vacuum, and that generally speaking, they would be used in conjunction with tag sets like asset class, subject, etc. That is, disciplineType and researchApproach aren't meant to convey the topic being discussed, simply the methodology used. We will improve our documentation to clarify this, as it is often a point of confusion.
- The researchApproach attribute is confusing to many people; we should update the documentation to stress that it is an optional tag, and to clarify when and how it is to be used. (The v2.5 documentation indicates that the attribute is required when the disciplineType is *Strategy* or *Investment*.)
- Some potential additions were suggested for either the disciplineApproach or researchApproach lists, some representing terms that would be particularly helpful in describing content that isn't company related; others representing new methods of creating research. These terms are: sentiment analysis, machine learning, ESG, forensic, event, macro, thematic, corporate access. Some of these terms may be better suited to the Subject enumeration list, so we will include these terms when we review that list; others may already be covered in other ways; others would be valuable additions to disciplineApproach or researchApproach lists. Regardless, we will ensure that content of all these types can be consistently tagged.

Focus

The purpose of the *Focus* tag set is to allow a publisher to indicate what the report is mainly about – that is, is it a company report or a sector overview? Based on the value chosen, the expectation is that the publisher will indicate which company (or which sector, etc.) is the primary focus of the report. While multiples are allowed, the recommendation is to select a single focus whenever possible.

The focus tag set is required, so it is essential that the list we provide allows for all types of research content to be accurately tagged. In the current release candidate for RIXML v3, we have updated the *Focus* enumeration list to include all of the top-level tag sets in the Research standard. As such, we discussed the following:

- Can a specialty be the focus? The Specialty tag set currently does appear as a top level tag set in the Research standard; however, the original intent for Specialty was to allow a publisher to indicate that its firm specializes in a particular area; it wasn't intended to describe the content in a research report, but instead to describe the author to indicate why you might want to give their research extra weight. The values in the Specialty list and the Subject list have gotten a bit entangled over the years. We introduced this question at this meeting but will discuss it in depth at the upcoming meeting focused on the subject and specialty lists.

- Can a report be "unfocused"? Since the Focus tag is required, the list needs to allow for research content that isn't accurately described as having any of the other focus values. For example, a report that has short updates on a number of topics and companies. The Focus enumeration list has the term "Unfocused" for any report that doesn't fall into any of the other areas; however, one participant mentioned that his firm uses the term "Generic" to describe this type of content. Would this term be more intuitive? Would it be confusing to replace the term "Unfocused" with "Generic" for the v3.0 release? We would like your input.

AI implications

- One participant observed that there is increasing interest in exploring the possibility of having AI tools provide the first pass at tagging a research item, with a human checking it for accuracy. Other participants agreed that this will almost certainly become more common.
- The Research v3.0 standard is adding the capability to indicate whether a tag was added using AI, but we need to continue to discuss the issue to ensure that the RIXML standards keep up with industry trends.

Remaining tag sets

There were 2 enumeration lists we did not have time to discuss: *LengthUnit* and *ResourceType*. It is unlikely that these require any changes; however, we would appreciate it if participants could review the v3.0 Data Dictionary to confirm. In particular:

- *LengthUnit* is an optional value that allows publishers to indicate whether the value in the *Length* tag represents a time value or a number of pages. Are there any other measures of length that would be helpful for the content your firm publishes?
- The only current option in the *ResourceType* list is *XBRL*. Does your firm publish research content with any type of XML data embedded in it other than XBRL data?