

Executive Summary

RIXML member firms have identified the need to convene a working group to discuss how best to address the critical issue of **identifying third-party content contained within research reports** and other research content. The newly-formed **RIXML Third-Party Data Identification Working Group** will bring together the many constituencies involved (see below), with the goal of determining how best to identify third-party data that may require carveouts or other special handling when research is delivered via interactive platforms or fed into AI tools. This group will focus on **the business-side issues**; it will develop best practice guidance and will likely recommend specific functionality that would be helpful to add to the RIXML Research Standard to ensure that it meets the needs of RIXML member firms and the industry in general.

Context

- Until recently, investment research content was generally distributed as a static PDF or text file.
- The licensing arrangements that sellside firms have with third-party data providers generally allow tables, charts, etc. containing this data to be included in PDF/text-based research content, and for that content to be distributed via the traditional means – to research aggregators, direct feeds to buy-side firms, via email, and via their own websites.
- New distribution purposes and mechanisms have developed that will likely require research publishers to identify third-party data contained within research items, for example:
 - as some sellside firms become interested in allowing their research content to be ingested into some buy-side firms' and/or vendor firms' AI/NLP tools, all parties need a way to identify content that must be excluded or that is subject to different rules.
 - as new formats for delivering investment research - such as on a sellside firm's interactive research portal - become more widely used, firms need a way to identify the content that can/cannot be displayed.
- RIXML member firms have clearly indicated a need to develop this capability - nobody wants to be out of compliance with their contractual obligations.
- Improving transparency regarding the ownership of each component of a research item will benefit all the involved constituencies identified below.
- All of these constituencies are represented among RIXML member firms, and RIXML has a long history of bringing member firms and others together to discuss and develop solutions for key issues regarding research distribution and consumption.

Involved constituencies

This issue has implications for:

- **sellside firms** who wish to enter into arrangements with buy-side firms, vendors, or others to allow use of their content in AI tools and/or who wish to ensure that the content they provide on their own websites adheres to their contracts with third-party vendors.
- **buy-side firms** and other investment research consumers who wish to enter into agreements with sellside firms to incorporate sellside firm research into proprietary NLP or other AI-powered tools or for other non-traditional purposes beyond that which third-party data owners currently permit.
- the **third-party data providers** who own the proprietary data in question.
- **product vendors** who are partnering with sellside firms to incorporate sellside firm research into proprietary NLP or other AI-powered tools, or for other non-traditional purposes outside of what third-party data owners currently permit.
- **aggregation vendors** and others who need clarity on the allowed usage of the research items they receive (and the components contained within them).

Beyond identifying the ownership of third-party content, sellside firms may also want to indicate *their own* ownership of content that may become componentized, for a variety of reasons.

RIXML Third-Party Data Identification Working Group Key Tasks

Based on input provided at our all-member meetings, the key tasks we have identified include:

- identifying use cases, potentially including other instances where a different set of rules applies to sub-sections (components) of a research item.
- determining what ownership and licensing information needs to be captured for components within a research item as well as for the full research item.
- developing recommendations and guidance that will help both member firms and others to authorize use of their own content in a way that enables content that they do not own to be identified and excluded.

The group may identify additional tasks.

The initial focus of this working group will be to provide the framework necessary to allow sellside firms and other research providers with the tools needed to identify third-party data to the aggregators, buy-side firms, and vendors it is partnering with for the reasons stated in the Executive Summary. While the broader issue of AI/NLP tools obtaining access to sellside research without authorization is certainly important, it is out of scope for this project.

The output of this working group may involve developing guidance documents, enhancing the RIXML Research Standard, and/or other member-recommended solutions.

Considerations

In developing the mechanism for identifying third-party content contained within research reports and other investment research content, the working group will consider a wide range of questions, including:

- What information will need to be conveyed about the sub-sections (components) of research reports that contain third-party content?
- Will content *ownership* information need to be conveyed for these components, or just an indication that the publisher is *not* the owner?
- Will there be a need to provide information regarding how to contact the third-party data owner and/or request access?
- Will entitlement information need to be included (that is, will the publisher need to indicate that the content should be available to some consumers or for some purposes)?
- Will there ever be a time-based aspect (that is, will content ever be available after a certain amount of time)?

Leveraging the strengths of the RIXML Research Standard

Because the RIXML Research Standard already serves as the industry standard for providing the structured data needed to describe investment research content, it is likely that the working group will identify ways that it could be enhanced to address the needs identified by the working group. The upcoming version of the RIXML Research Standard (v3.0) has added the ability to identify components within a research item, so the framework already exists and can be extended as needed.

Next steps

We will convene a meeting to kick off this working group and identify key tasks. We encourage all RIXML member firms to participate.