RIXML Research Standard v3.0

June 16, 2022

RIXML Quarterly Meeting

Goals

✿ Cleanup

- Better integrate enhancements made in the v2.x series that were implemented with backward-compatibility in mind

- Remove elements of the schema not used by adopters

🧩 Enhance

- Respond to the evolving publishing space by better covering topics such as linkbacks, audio/video content, and episodic content

- Respond to emerging topics such as ESG, blockchain, cryptocurrency, and cannabis

👤 Explain

- Improve how we use object diagrams in our documentation

- Unify terminology definitions across schema and documentation

- Capture and incorporate more recent best practices

Structure

Administrative Track
- Develop timeline
- Assemble teams
- Convene meetings
- Report on progress
- Coordinate feedback
- Update website (NEW)
- Plan and execute marketing strategy

Documentation Track
- Ensure definitions exist for all tags and tag sets
- Ensure definitions are stored within design tool
- Revise & enhance best practice guidance, implementation guidance, etc.
- Create cohesive documentation that contains definitions, object diagrams, and implementation and best practice content

Development Track

Toolbox
- Determine required tools
- Schema design tool
- Object diagramming tool
- Documentation extraction tool
- Project management tool

Cleanup
- Evaluate current standards
- Identify tags to be retired
- Identify past design compromises made to maintain backward-compatibility and simplify
- Identify and adjust areas that industry has outgrown
- Review & update enumeration lists

Enhancements
- Review backlog of change requests
- Incorporate findings on Topics
- Gather input from RIXML community

Topics
- Research
- Interactions
- ESG
- Linkbacks
- Entitlements
- Sector/Industry Classification
- Interactive Content
- Audio/Video Content
- Episodic Content
- Unstructured Data & AI
- Entitlements & Subscriptions
- Blockchain & Cryptocurrency
- Cannabis

Schema
- Develop updated set of schema files
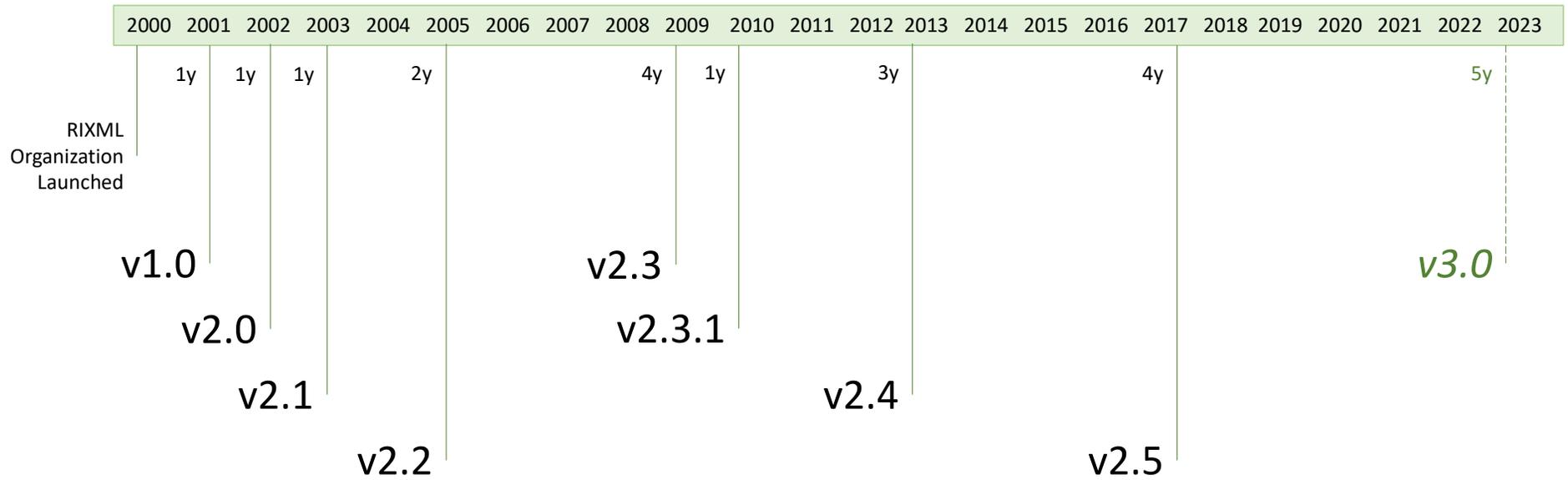- Test and validate

Progress

Plan

Phase 1: Planning and outreach	Phase 2: Scoping	Phase 3: Development	Phase 4: Testing and validating	Phase 5: Public release
Gathering requests; encouraging participation	Discussing requests & defining scope; determining major milestones and target deadlines	Incorporating changes into Standards Suite and updating documentation	Community testing and feedback	New version of Standards Suite becomes available

Actual

40%	30%			

Target Dates

July 29	August 31	November 18	February 28	March 31

History

2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023

1y 1y 1y 2y 4y 1y 3y 4y 5y

RIXML
Organization
Launched

v1.0

v2.0

v2.1

v2.2

v2.3

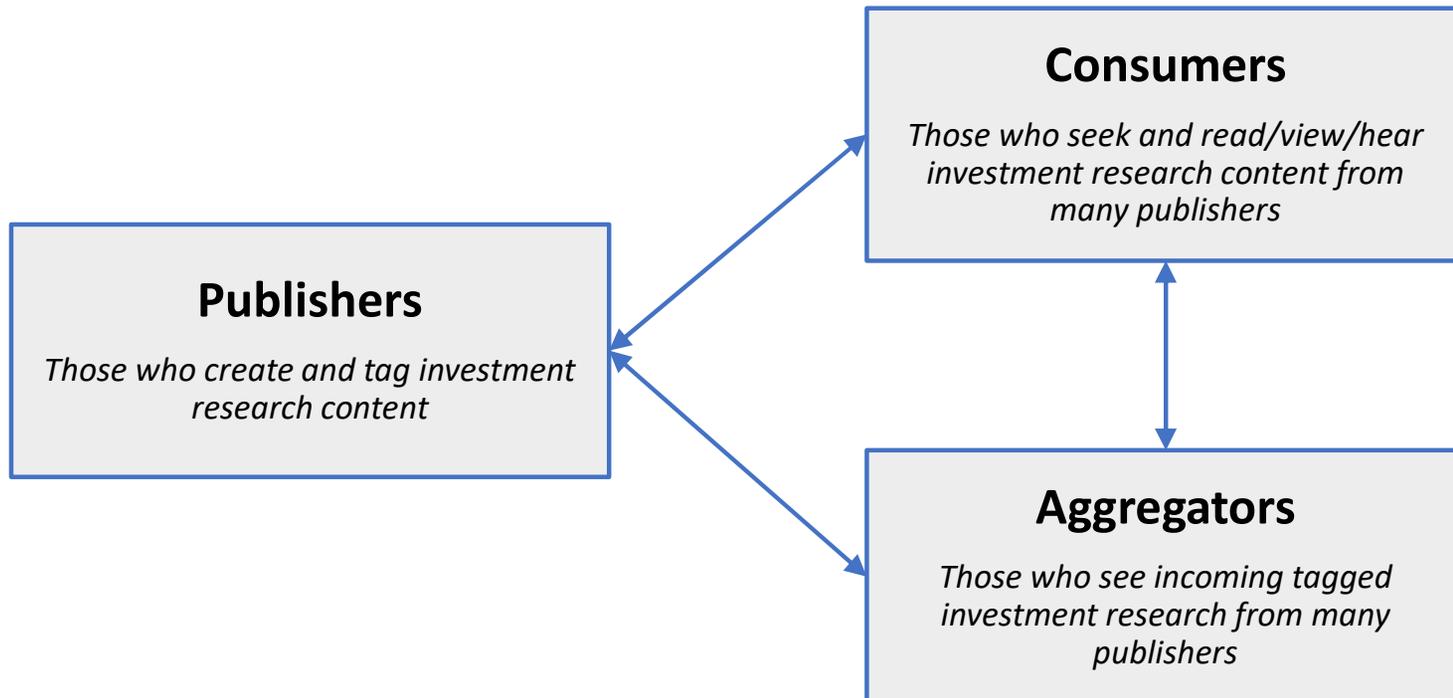v2.3.1

v2.4

v2.5

v3.0

RIXML Research Standard
Production Releases

How has the research landscape changed?

- Since we released v2.5 of the Research Standard in 2017
- Since we launched the organization in 2000

Voices

Publishers

Those who create and tag investment research content

Consumers

Those who seek and read/view/hear investment research content from many publishers

Aggregators

Those who see incoming tagged investment research from many publishers

Publishers

Which RIXML tags from the v2.x series do you **never** supply in your tagging? Or very rarely supply?

What **publisher-defined** tag values are you using most frequently that could be added to v3.0?

What information are your production processes capturing and storing outside RIXML that you would like to see **added** to v3.0?

Who at your firm should RIXML consult during the **outreach** phase the v3.0 cycle?

Consumers

What research content do you find most difficult to **search** for? What kinds of search terms are usually ineffective?

What are some other sources of **friction** in getting to the research content you need when you need it?

How would you like to see search **results**? What useful information could be added to the typical *"list of headlines"* presentation of results?

Who at your firm should RIXML consult during the **outreach** phase the v3.0 cycle?

Aggregators

Which RIXML tags from the v2.x series are **never** supplied by publishers? Or very rarely supplied?

What information missing in v2.x would you like see **added** in v3.0?

Which RIXML tags from the v2.x series are **not used** on your client-facing platform(s)?

Who at your firm should RIXML consult during the **outreach** phase the v3.0 cycle?

Everyone

How will research **content** evolve? What will be the next formats, the next categories?

What new **topics** will research cover? What comes after ESG, blockchain, cryptocurrency, etc.?

How will research **consumption** evolve? What comes after PDF, browser, and mobile apps?

What other tagging **standards** are out there, established or emerging, that RIXML should consider?

Example Ideas

Audio & Video Content

**Thinking about audio & video content both embedded within a report and standalone**

- Use-cases for embedded within a report vs. standalone, ad-hoc vs. episodic

- Calls for component-level tagging with context (authors, companies, etc.)

- Calls for indexing to specific time markers

- Perhaps new ProductCategoryEnums
  - *Podcast*
  - *Interview*
  - *PanelDiscussion*
  - *EventRecording*
  - *Marketing*
  - Any others?

Episodic Content

Research published on a regular schedule, such as a Periodical or Podcast, with sequential episodes

- Capturing episode identifiers, such as absolute episode number and/or episode within series
    - Should we expand on volume/number combination to also support year/week, year/month, year/quarter?
- Other descriptive data about the series
    - Longer-form description of the purpose of the series
    - Series Owner (Team or Individual)
    - Some context tags like Sector, Industry, AssetClass, Region, Country
- Expected search keywords for aiding discovery
- Inventory Check Schema for listing full catalog of episodes

- RIXML already supports
    - Product Series
    - Series Name
    - Series ID
    - Periodicity
    - Volume
    - Number
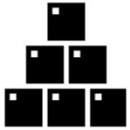
Component-Level Tagging

When components (charts, videos, etc.) are embedded within a specific report, should the RIXML include…

- Component-Level Tags
    - Title, Sub-Title, Source Line
    - Index within report (exhibit number, section number, time index)
    - Some kind of direct address or link?
- Component-Level Context Tags
    - Authors
    - Companies, Industries
    - Asset Classes
- Component Catalog
    - List of major sections
    - List of charts, tables, audios, videos
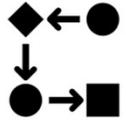
Related Content

Capturing how individual content items (reports, etc.) relate to others, such as derivations and follow-ups

- Company Report relates to latest Company Model
- Company Report relates to latest Industry Report
- Company Report relates to original Company Initiation Report
- Report in Japanese is a translation of same Report in English
- Economics Report references external Economic Dataset
- Podcast Video references original Annual Outlook report
- Could look like a kind of bibliography?
- Other cases?

- RIXML already supports
  - Based On
  - Part Of
  - References
  - Replaces
  - Requires
  - Translation Of
  - Related To

- By report IDs only

Resource-Level Context Tagging

Ability to mate specific payload files to local tags for author, company, industry, asset class, etc.

- Some publishers make more use of the multiple-payload support in RIXML

- No way to mark one as the Company ABC Model and the other as the Company XYZ Model

- No way to indicate if different payloads are associated with different authors

- A big cross-asset compendium report could have many payloads that call for context tagging

- How much should we support?
  - Author
  - Company
  - Sector/Industry
  - Asset Class
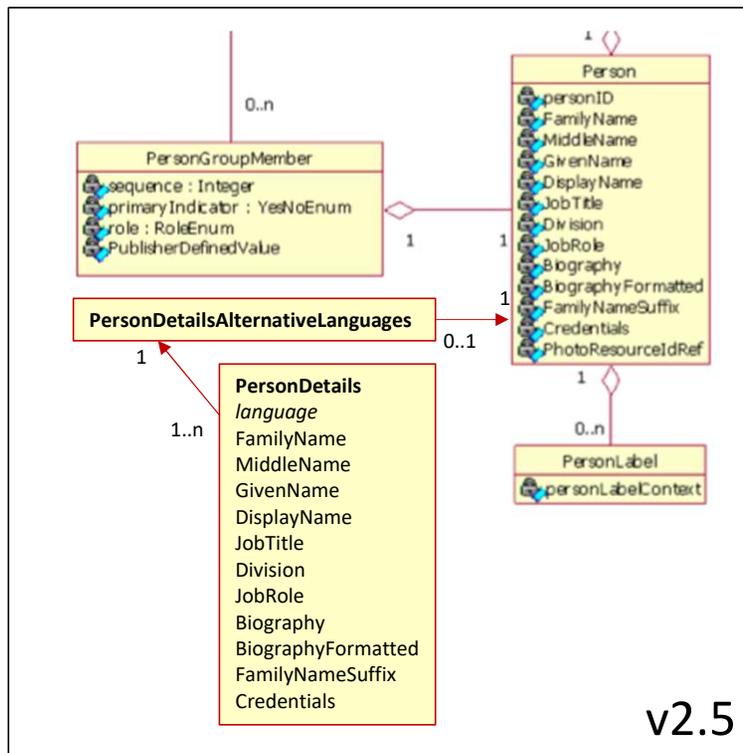
- How about also…
  - Translations
  - Alternative formats

Tag Names in Multiple Languages

Simplify our v2.x design for supporting tag names in multiple languages

- Use proper *language* attribute within each relevant element combined with a more natural cardinality to establish peer relationship between language variants of the same element data
- This is simpler and better organized than the *AlternativeLanguages* construct built into v2.5

- Covers language support for
  - Person name
  - Person Group (Team) name
  - Title
  - Sub-Title
  - Synopsis
  - Abstract

Go from this...   To this

v2.5

Person
- personID
- FamilyName
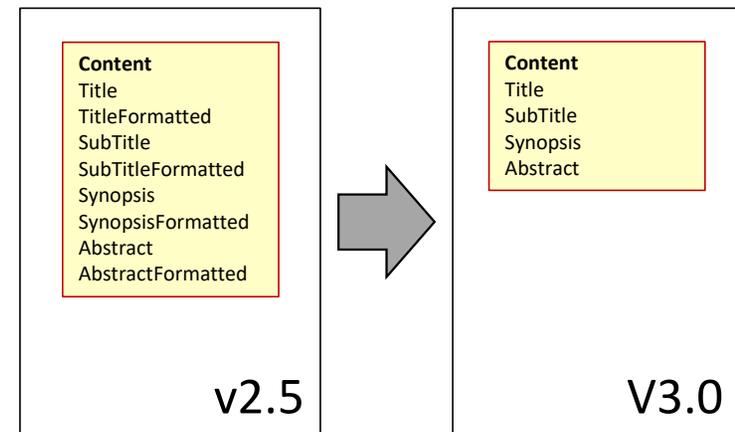- MiddleName
- GivenName
- DisplayName
- JobTitle
- Division
- JobRole
- Biography
- BiographyFormatted
- FamilyNameSuffix
- Credentials
- PhotoResourceIdRef

PersonGroupMember
- sequence : Integer
- primaryIndicator : YesNoEnum
- role : RoleEnum
- PublisherDefinedValue

PersonDetailsAlternativeLanguages

PersonDetails
- *language*
- FamilyName
- MiddleName
- GivenName
- DisplayName
- JobTitle
- Division
- JobRole
- Biography
- BiographyFormatted
- FamilyNameSuffix
- Credentials

PersonLabel
- personLabelContext

V3.0

Person
- personID
- *language*
- FamilyName
- MiddleName
- GivenName
- DisplayName
- JobTitle
- Division
- JobRole
- Biography
- BiographyFormatted
- FamilyNameSuffix
- Credentials
- PhotoResourceIDRef

PersonGroupMember
- sequence
- primaryIndicator
- Role
- PublisherDefinedValue

PersonLabel
- personLabelContext

✿ Unformatted Tags vs. Formatted

Simplify our v2.x design for supporting both unformatted and formatted tag names

- We currently support both unformatted and formatted variants for these elements below the Content element
  - Title
  - Sub-Title
  - Synopsis
  - Abstract

- A simpler and better organized alternative would be to have one set of these four sub-elements with a *formatted* attribute, rather than having a duplicate set of sub-elements for each variant

Content
Title
TitleFormatted
SubTitle
SubTitleFormatted
Synopsis
SynopsisFormatted
Abstract
AbstractFormatted

v2.5

Content
Title
SubTitle
Synopsis
Abstract

V3.0

✿ Remove Elements Not Adopted

Simplify our schema by removing parts of the design not adopted by the marketplace

- *Root element*
  - The Research element at the top of the schema was meant to include multiple Product elements (corresponding to reports)

- *Organization Expertise*
  - Does it make sense to keep this somewhat complex structure under the Organization element or perhaps spin it off into a Side-Car Schema?

- *Resource.Price*
  - Including pricing data in the report tagging file is probably not the best approach since it is subject to change

- *Certain Entitlement Types*
  - Is anyone using Country, Region, or Sector entitlements?

- *Events*
  - Does it make sense to keep this somewhat complex structure under the Context element or perhaps spin it off into a Side-Car Schema?

Asset Class Hierarchy

Simplify our schema by restructuring Asset Class, Asset Type, and Security Type into a real hierarchy

- We already have a documented hierarchy that structures the three Asset Class levels but do not implement them that way in the schema

- We present them as peers with no interdependencies

- Only in the documentation do we recommend awareness of the hierarchy

- Should we formalize this better in v3.0 so that validation of an instance document against the schema requires it?

- RIXML currently supports
  - Asset Class
  - Asset Type
  - Security Type

- But as independent peers