Practical Tips for Setting up Taxonomy Governance

Taxonomy Boot Camp London
12 October 2022

Heather Hedden
Data and Knowledge Engineer,
Semantic Web Company

About the Speaker

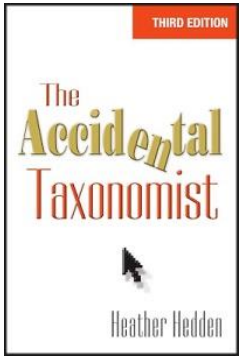
Heather Hedden
Data and Knowledge Engineer
Semantic Web Company

Over 25 years of experience in developing and managing taxonomies, metadata, and other knowledge organization systems for various organizations and applications.

Prior taxonomy consultant and staff taxonomist.

Instructor of self-paced online taxonomy courses.

Author of the book *The Accidental Taxonomist, 3rd edition (2022)*.



About Semantic Web Company and PoolParty

SWC is developer / vendor of **PoolParty Semantic Suite**

Most complete and secure **Semantic AI platform** on the global market

W3C standards compliant

ISO 27001:2013 certified (since 2019)

First release in 2009

Current version **8.1**

On-premises or **cloud-based**

Over **170** **customers world-wide**

Semantic AI: Fusion of graphs, NLP, and machine learning

Gartner named SWC a Visionary in their **Magic Quadrant** for Metadata Management Systems 2019 and 2020

KMWorld listed PoolParty as **Trend-Setting Product** 2015 - 2022 and SWC in the **AI 50** list of companies in 2020 and 2022

Forrester listed SWC as sample vendor in their **report** on *The Document-Oriented Text Analytics Platforms Landscape 2022*

Outline

- Introduction - *Why* taxonomy governance
- Planning governance - *When* is governance done
- Scope of governance - *What* is governed
- Governance roles and responsibilities - *Who* governs
- Policies and procedures - *How* is governance done
- Other governance sources
  - Taxonomy management software
  - Taxonomy-related standards
- Governance documentation types
- Governance plan example

Why Taxonomy Governance

Why Governance?
Taxonomies need to grow and change over time.

To ensure that the process for taxonomy maintenance defines in a repeatable and controlled way how changes in the taxonomy are approved and implemented

- To be prepared for change
- To mitigate risk
- To involve the proper stakeholders

Why Taxonomy Governance

Taxonomies need to grow and change over time.

Getty images

Taxonomy governance covers policies and procedures for maintaining the taxonomy through updates and all kinds of changes.

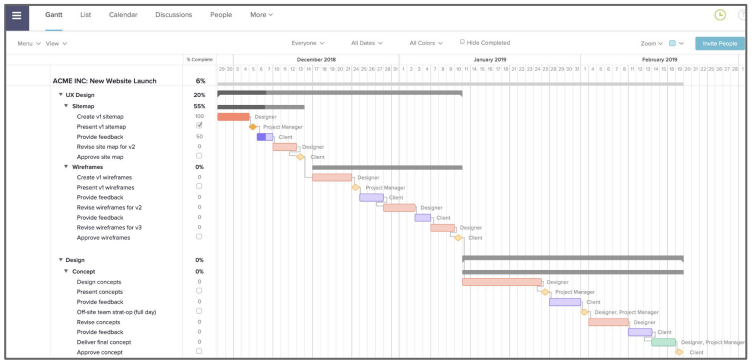
Why Taxonomy Governance

Change Triggers include:

- New concepts in the world (*COVID-19*, *Hybrid work*, *Digital nomads*)
- Changed label trends, for example to reduce bias (*Waitress* to *Server*; *Blind* to *Blindness*)
- New content sets, bringing up new concepts
- New requirements, users, needs, markets, etc.
- User feedback (taggers or searchers) suggesting improvements
- New taxonomy management, tagging, or search systems call for taxonomy restructure or support enhanced features

Planning Governance - When

- Governance should not come at the end of a taxonomy project.
- Taxonomy governance process starts with the start of creating the taxonomy
- As issues come and get resolved, they get documented.
- Documented decisions become the basis of governance policy.

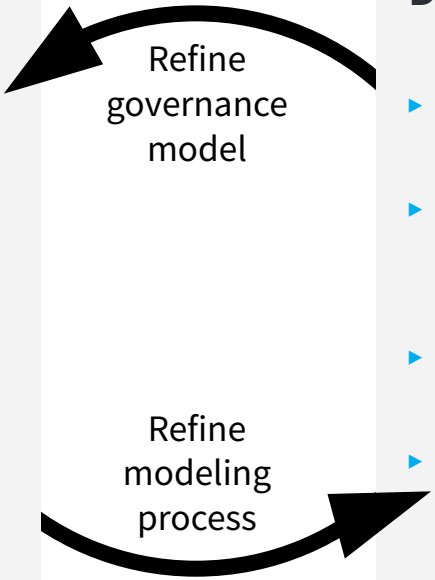


Planning Governance

Governance is an iterative process.

Develop governance model

- Define guiding principles for process governance.
- Determine members and management of a governance council.
- Define roles and responsibilities for managing the taxonomy system.
- Implement workflows for taxonomy development and maintenance.
- Define and refine update scenarios and cycles.



Develop knowledge model

- Analyze domain by reviewing representative content.
- Obtain stakeholder input, such as by interviews and cardsorting.
- Reuse existing controlled vocabularies.
- Build and extend the taxonomy structure in iterations:
  - Corpus analyses
  - User testing
  - User feedback

Planning Governance

In planning governance, especially the procedures and roles, consider the goals and factors important to the organization.

Example governance principles (Semantic Web Company):

- A lightweight and agile process without significant overhead
- Asynchronous communication when possible for coordination within the workgroup
- Avoiding introducing new regular meetings, both at implementation and strategic level, so ad hoc interactions are to be followed as much as possible
- Only involving people whose time is scarce when they are truly needed (ad hoc composition of the governance council)
- Harnessing existing infrastructure elements - e.g. the Jira support desk for processing requests

Scope of Governance - What

Taxonomy/thesaurus governance comprises:

- Documented tasks, roles, responsibilities, and decision-making processes for taxonomy maintenance and for major revisions
- Taxonomy descriptive documentation (purpose, type, scope, users, indexing method, history/sources)
- Taxonomy editorial policy/guidelines for maintenance
- Possibly descriptive metadata element rules and governance
- Indexing or tagging policy/guidelines
- Instructional/how-to documents (system-specific)

Scope of Governance

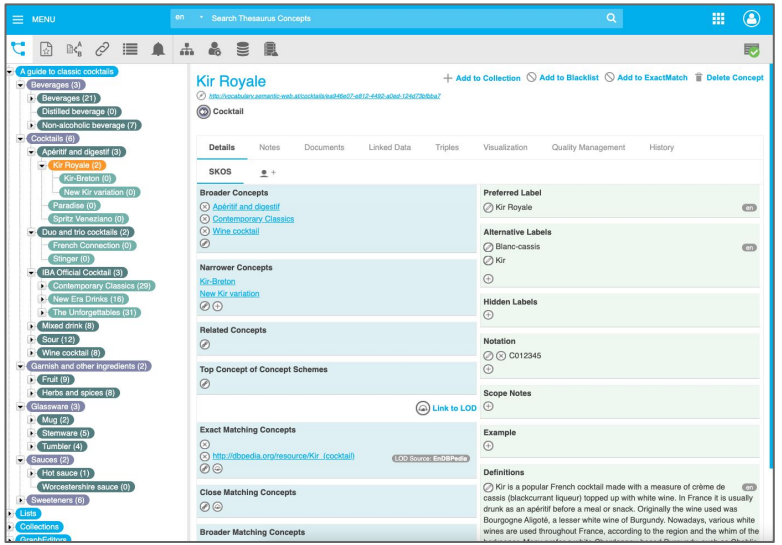
At a minimum, document policy and procedures for:

Routine changes:

- Modifying concepts
  - Changing preferred labels
  - Adding/deleting alternative labels
  - Editing notes and other attributes
  - Moving concepts/changing broader concept
- Adding new concepts
- Deleting or merging concepts

Occasional, major changes:

- Adding concept schemes or classes
- Modifying concept schemes or classes
- Deleting concept schemes or classes



Governance Roles and Responsibilities - Who

Why define governance roles?

- Stakeholders in a taxonomy tend to be cross-departmental
- Different organizations manage taxonomies differently
- Taxonomy management may be a part-time role and a shared responsibility
- The taxonomy may grow to include a greater scope and additional people and departments

Governance Roles and Responsibilities - Who

Define processes with roles, documenting

- Who can initiate/request changes in the taxonomy
- Who can implement approved changes
- The responsibilities of the involved roles

Who should be involved in governance

- Taxonomist(s)/taxonomy editor(s)
- Representatives of taxonomy users (taggers/indexers, content managers,)
- Representatives of business units, products, services
- Representatives of IT/software development
- Executive or directorship leadership
- Possibly legal department

Governance Roles and Responsibilities

Identify levels of changes:

- Major changes - new and changed concepts schemes
- Medium changes - multiple concepts, branches, or concepts with many narrower
- Minor changes - individual concepts, deeper in the taxonomy

Identify different roles and groups responsible for approving each of different level of changes:

- Senior/managing taxonomist
- All taxonomy editors
- Business stakeholders
- Subject matter experts
- IT systems administrators/developers (Identify different systems impacted by changes.)

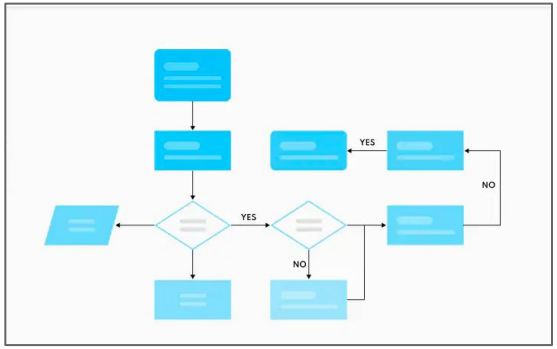
Determine to what extent taxonomy management is centralized or decentralized when there are multiple taxonomies.

Determine change communication plan and method(s).

Governance Policies and Procedures - How

Policies and procedures should be create for

- Editing and expanding the taxonomy
  - Maintenance changes
  - Major changes
- Gaining approval for changes to the taxonomy
- Communicating changes to the taxonomy
- Tagging with the taxonomy



Governance Policies and Procedures

For maintenance policy and procedures, need to determine:

- What kinds of changes are routine, which are not
- What information needed to determine the changes
- What group should maintain the taxonomy
- What role indexers play in suggesting changes
- The processes for proposing and resolving changes
  - Comment-handling, appeals, issue logs, announcements, update schedules, etc.

For maintenance, review:

- Newly added content sources
- Search logs
- Sections of the controlled vocabulary covering high-change topics
- Tagging statistics (human or auto) to find:
  - high-use terms needing further differentiation
  - low/no-use terms that should be merged
  - identifiable indexing errors
-

Governance Policies and Procedures

Questions to be addressed:

- How can taggers, content managers, and other uses request additions and changes?
- How are changes from those who are not to be approved?
- What is the process for reviewing suggested changes?
- How are changes evaluated?
- How are requests prioritized?
- Will implemented suggestions be tracked for further review later?
- How are change decisions documented, if at all?

Other Governance Sources

“We’ll use dedicated taxonomy management software, and that will provide governance”

Governance through Taxonomy Management Software

Many features are supported.

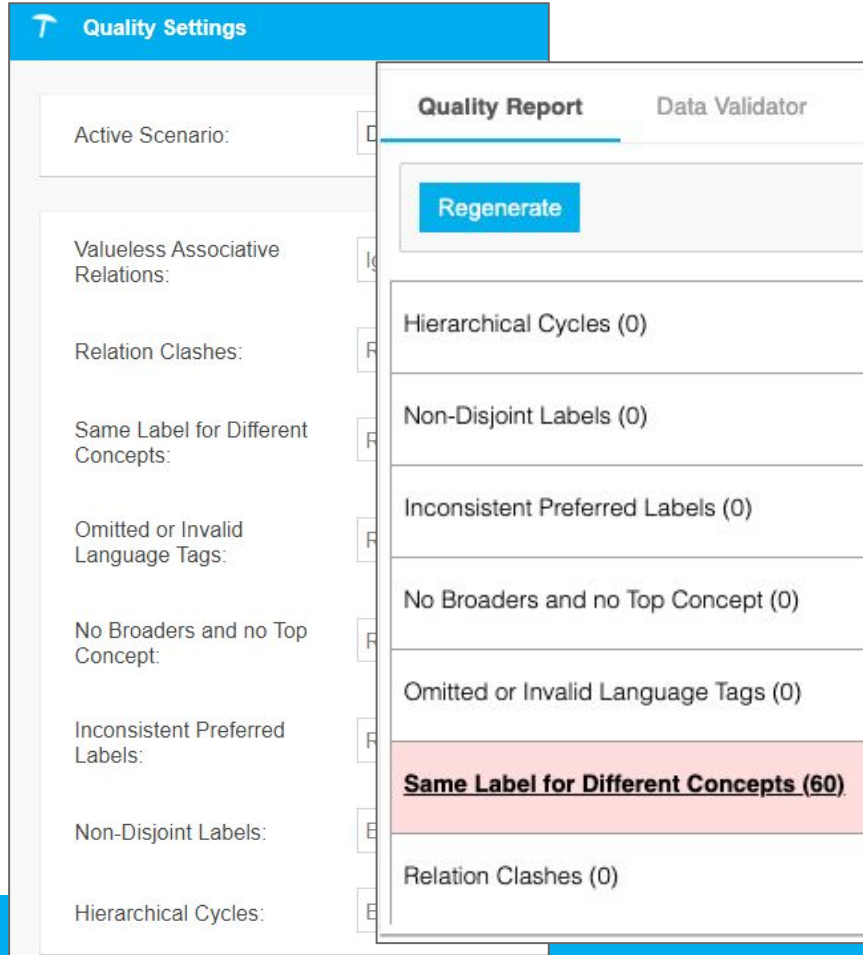
But allows choices and configurations, which require decisions and policy.

- Interoperability standards and data validation - does not ensure taxonomy quality
- Quality standards - can be configured
- Workflow management - can be used as desired

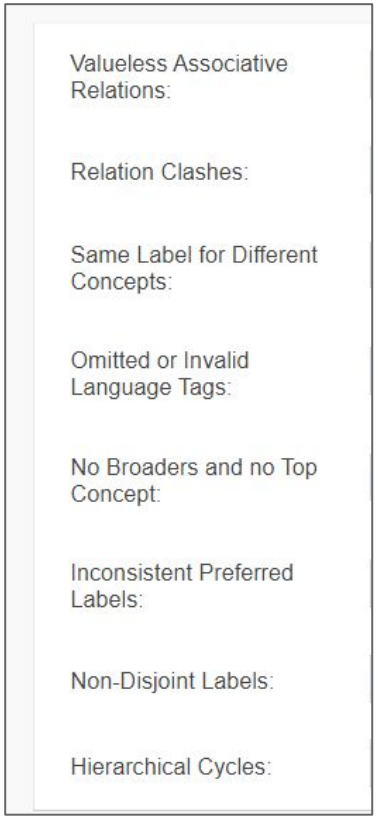
Governance in Taxonomy Management Software

Governance through PoolParty **Quality Settings** (Advanced Menu):

- Customizable settings for:
  - Ignore
  - Report (when running a quality report)
  - Enforce (will prohibit the action/creation)
- Options of:
  Default, Checks Disabled, Custom
- Certain violations arise only from an import file.



Governance in Taxonomy Management Software

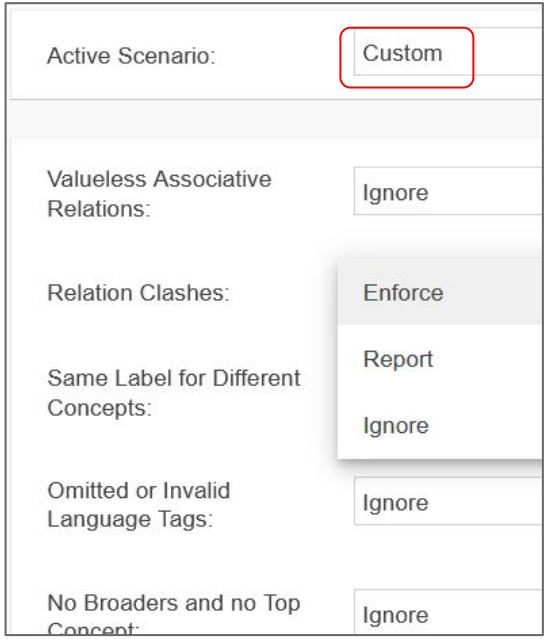


- Related concept (RT) relationships between sibling concept with the same broader concept
- Broader/narrower and related concept relationships between the same pair of concepts.
- Two (or more) concepts have the same label of any type (preferred, alternative, or hidden).
- **Language tags that have been set incorrectly or not set at all.**
- **Concepts that do not have a broader concept and are also not a top concept. (Ideally, only top concepts lack broader concepts.)**
- **Concepts that have more then one skos:prefLabel per language.**
- **Concepts that have the same label two times as skos:pref/alt/hiddenlabel in the same language.**
- **Concepts that are related to each other as both broader and narrower, either directly or indirectly through additional hierarchical levels of concepts.**

Governance in Taxonomy Management Software

Quality Settings Recommendation

- Used Custom quality settings.
- Determine which setting type (*Ignore*, *Report*, or *Enforce*) shall be set for each scenario.
- Document the choice, providing the reason

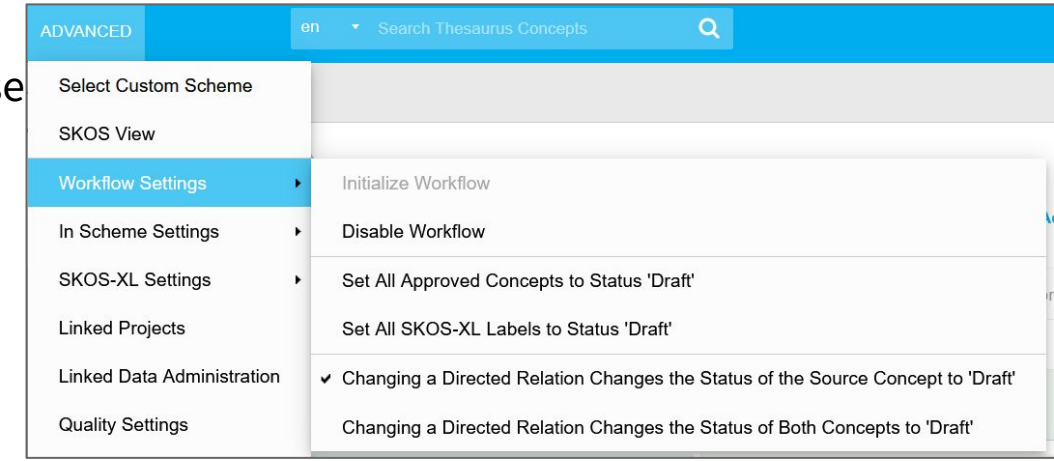


Governance in Taxonomy Management Software

Governance through User **Workflow Management**

Decisions regarding workflow that form governance policy:

- Whether or not to use Workflow Settings
- Who can self-approve and who needs approval
- When (with what kinds of changes) can taxonomy editors self-approve and when do they require approval
- Who should grant approval for those users or situations requiring it
- What notes or documentation are required when requesting and granting approval



Governance in Taxonomy Management Software

Decisions regarding workflow that form governance policy:

- Whether or not to use Workflow Settings
- Who can self-approve and who needs approval
- When (with what kinds of changes) can taxonomy editors self-approve and when do they require approval
- Who should grant approval for those users or situations requiring it
- What notes or documentation are required when requesting and granting approval

Possible option:

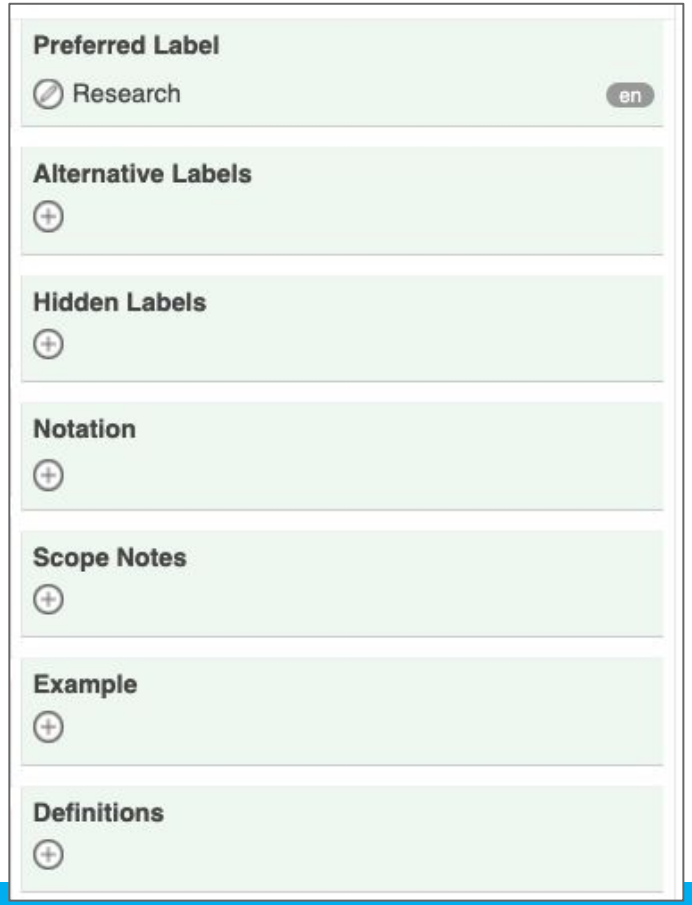
- Temporarily turn on workflow settings during an editing project of a junior taxonomist.
- Other experienced taxonomists can self-approve their changes during this time.

Governance in Taxonomy Management Software

Decisions regarding concept details that form governance policy

How shall you use **SKOS elements**?

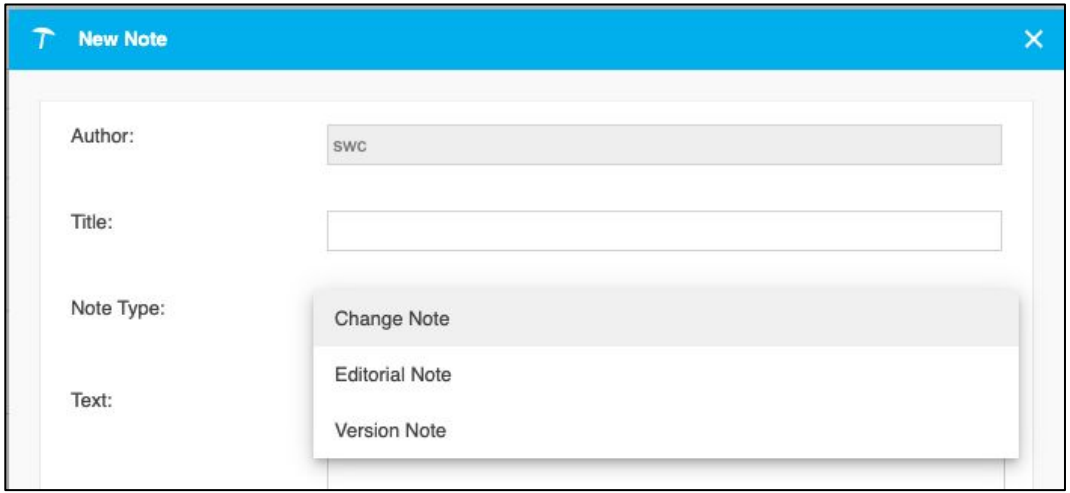
- Use of SKOS fields for: Hidden Labels, Notation, Scope Notes, Example, and Definitions
  - As *never use*, *always use*, or *sometimes use*
  - What criteria for sometimes use
  - What editorial form of entry to follow
- Language tag and usage policies



Governance in Taxonomy Management Software

Decisions regarding **Notes policies**

- Whether, when, how and what types of notes to use: Change, Editorial, Version

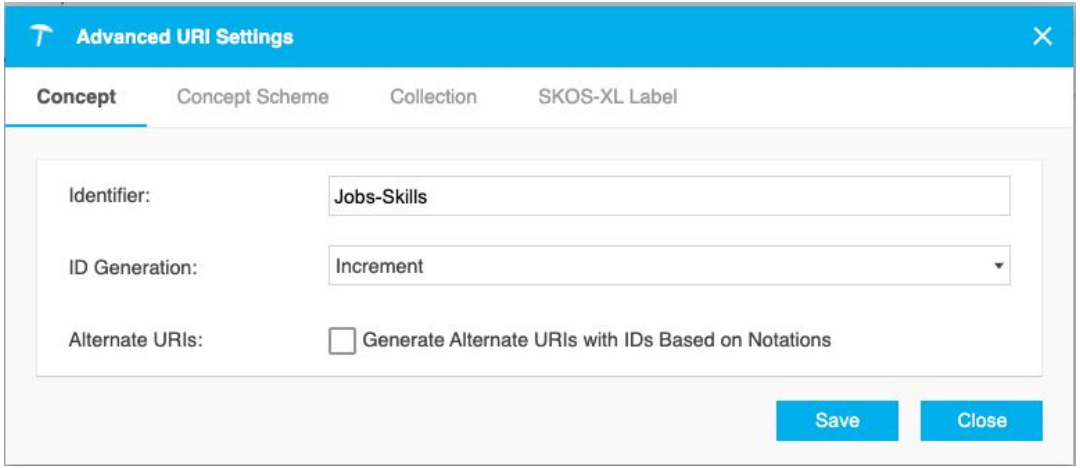


Governance in Taxonomy Management Software

In SKOS vocabularies, **URI naming and incrementation policy**

For URIs of concepts, concept schemes, collections, and SKOS-XL labels
Options:

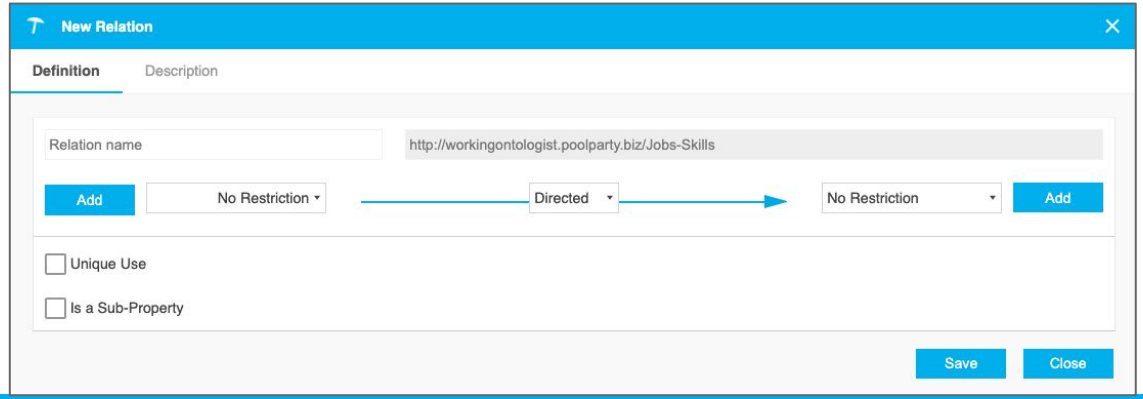
- UUID
- From Preferred Label
- Increment
- Manual



Ontology Governance

OWL ontologies inherently support governance with “restrictions” and other rules

- Class disjointness - *option when creating a Class*
  - Two Classes are disjoint if they cannot share a member concept
- Class, Relation, and Attribute descriptions and domains and ranges - *option when creating Relations or Attributes*
  - Make as restrictive as needed for business purposes
- Relation and Attribute unique use and sub-property
- Attribute type: No Restriction, URI, Literal, Integer, Long, Float
  - Enforce types, leave as No restriction, or all as Literal?
- Attribute unique or multiple use for a give concept



Other Governance Sources

“We’ll follow the standards, and that will provide governance”

Taxonomy-related standards

- ISO 25964 and ANSI/NISO Z.39.19
  - For ***thesauri***, not for all ***taxonomies***, which are more flexible.
  - Need to determine how closely you want to follow the thesaurus standards

- SKOS, RDF, RDF-S, OWL (W3C guidelines)
  - For interoperability and machine-readability only
  - Do not ensure a taxonomy custom is designed for its users and organizational needs; do not address processes

Governance Documentation Types

1. General taxonomy documentation/description

- Taxonomy purpose(s) and intended use(s)
- Taxonomies owners and users
- Basic type (hierarchical taxonomy, faceted taxonomy, thesaurus, a combination)
- History of the origins and development of the taxonomy
- Taxonomy size and expected growth rate
- Taxonomy languages
- Scope of concepts and content
- Method for tagging

Governance Documentation Types

2. Taxonomy editorial policy

- Concept label editorial style (case, abbreviations, acronyms, special characters, diacritics, punctuation, plurals, spelling)
- Extent of compound/pre-coordinated concepts and specificity of concepts
- Types of concepts included and types of terms not
- Authoritative sources for concepts and labels
- Guidelines for including alternative labels
- Relationship policy (hierarchy depth, polyhierarchy, associative relationships, relationships between topics and named entities, orphans)
- Use of custom classes and attributes for concepts
- Notes policy

Governance Documentation Types

3. Tagging policy

For each taxonomy, concept scheme, term list, facet, or metadata element

Policies for tagging include:

- Required or not required for each content item
- Required for certain kinds of content
- Tagging with broader concept in addition to specific concepts
- Single only or multiple concepts from the same concept scheme or facet for each content item
    - Example: Audience - Only one or multiple permitted?

Guidelines for manual tagging or manual review

- Depth of tagging: approximate range of number of concepts per content item

Governance Plan and Role Example (SWC)

Taxonomy Working group roles and responsibilities (Semantic Web Company)

Role	Responsibilities	Example
Taxonomist	Making day-to-day well-defined changes (specified by the domain experts) to taxonomies and concepts. Supporting domain experts in organising a new part of the taxonomy. Ensuring structural consistency and that good practices are followed. All changes have a Draft status until the Taxonomy manager approves.	Supporting HR developing a new taxonomy for employee skills.
Taxonomy manager	Approving (complex) taxonomy changes. Understanding the taxonomy (strategically as well). Changes made by a taxonomy manager can be review by another taxonomy manager (if needed)	Approving a new concept, for which, the domain expert did not provide its location in the taxonomy.
Taxonomy system administrator	Managing of the taxonomy management system from within the system.	Access management task, exporting taxonomies, workflow management, creation of new servers...

Governance Plan and Role Example (SWC)

Taxonomy Governance Council Roles

Role	Description
Change proposer	Any person who feels the need to introduce a change in the taxonomy can propose it. The proposal should be discussed within their department The request will be done via a form (to be defined), which will make use of the PPT Suggest service
Taxonomy sponsor	Stakeholder within the business, who recognises, supports and sponsors the development of the taxonomy. In our case, the primary sponsor is [name], COO

The council will oversee at strategic and business level the development of the taxonomy, including endorsing new developments.
The composition of the council will be fluid, adding new members as specific areas of knowledge are needed.
The core of the council will be [named individuals/job titles]

Governance Plan and Role Example (SWC)

1. An employee notices that the taxonomy needs to be changed.
2. The employee discusses the need within their department.
3. If the department agrees that it is needed, the n employee requests the change via the provided form (to be defined).
4. The taxonomy workgroup receives notification of the change request.
5. The taxonomy workgroup discusses new tasks as they arrive via Slack, and distributes them for execution within the workgroup. For complex changes, a taxonomy manager is assigned for reviewing the implementation.
6. The workgroup members execute the changes on their own, or for more complex ones they involve a domain expert, possibly the same person who requested the change.
7. Complex changes are reviewed by the designated taxonomy manager.

The reviewer (or the executor, it if no review was needed) notifies the employee who requested the change.

Conclusions

Setting up taxonomy governance

- Start addressing governance questions from the start of the project
- Document taxonomy description, policies, procedures, roles
- Consider all the stakeholders broadly
- Follow the plan; if it does not work, adjust it

Resources

- “Taxonomy Governance Best Practices” (2017), Zach Wahl, Enterprise Knowledge White Paper https://enterprise-knowledge.com/taxonomy-governance-best-practices/
- “Taxonomy Governance” presentation (Taxonomy Boot Camp conference, 2021), John Horodyski
- “Knowledge Graph Governance” (2020) Andreas Blumauer and Helmut Nagy, excerpt from *The Knowledge Graph Cookbook,* pp 67-68.
- “Taxonomy Governance” (2022) Cynthia Knowles, chapter in *Taxonomies: Practical Approaches to Developing and Managing Vocabularies for Digital Information*, edited by Helen Lippell.

- *The Accidental Taxonomist*, 3rd edition (2022) Heather Hedden https://books.infotoday.com/books/Accidental-Taxonomist.shtml

Questions/Contact

Heather Hedden
Data and Knowledge Engineer
Semantic Web Company Inc.
One Boston Place, Suite 2600
Boston, MA 02108
USA
+1 857-400-0183
heather.hedden@semantic-web.com
www.linkedin.com/in/hedden
http://accidental-taxonomist.blogspot.com

Semantic Web Company www.semantic-web.com

PoolParty software www.poolparty.biz