



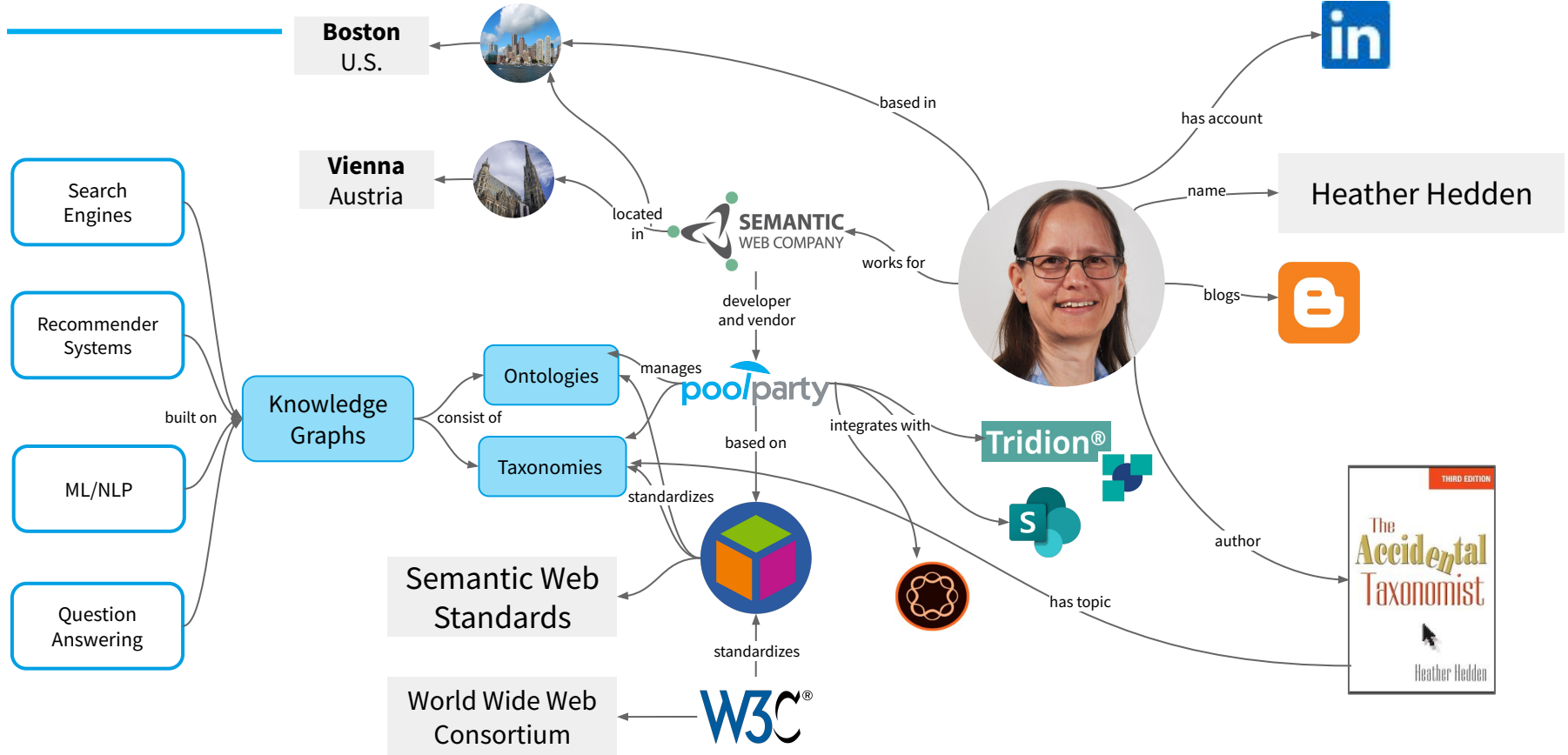
# Taxonomies in SharePoint: Challenges and Opportunities

Bite-Sized Taxonomy Boot Camp London  
11 October 2023

The background of the slide is a light blue-grey color. It features a top-down view of a person's hands using a tablet. The tablet screen shows a colorful circular chart with segments in green, yellow, red, and blue. To the right of the tablet, there is a complex network diagram consisting of white lines connecting various circular nodes. Some nodes contain icons, such as a bar chart, a gear, and a document. A large globe-like structure is also visible on the right side of the network diagram.

**Heather Hedden**  
Knowledge Manager  
Semantic Web Company

# Click the Graph—get in contact with us!



- ▶ Introduction
  - ▷ What is SharePoint and what is it for
  - ▷ Search issues
  - ▷ Taxonomy solution and benefits
- ▶ Taxonomy use and management in SharePoint
  - ▷ SharePoint capabilities and weaknesses
- ▶ Improving taxonomy management for SharePoint
  - ▷ Challenges & solutions of managed metadata/taxonomies
  - ▷ Taxonomy management system integration

A solid blue rectangular box containing white text. A thin white vertical line is positioned to the left of the text.

**Introduction:**  
**What is SharePoint**  
**Search issues & taxonomy benefits**

## What is SharePoint

- ▶ Since 2001, software from Microsoft for shared document management and collaboration
- ▶ Major past versions: 2003, 2010, 2013, 2016, 2019
- ▶ On-premise or online (included in Microsoft 365 for business)

## What is SharePoint used for?

- ▶ Intranet platform
- ▶ Content and document management - uploaded files
- ▶ Collaboration, including workflow, in team sites



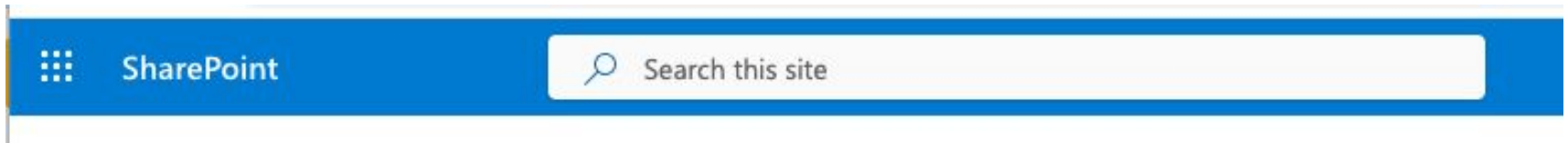
## SharePoint general user types

- ▶ **Visitors** - readers: need to search for content, such as the Intranet
- ▶ **Members** - contributors: can create pages and upload and tag content
- ▶ **Owners** - administrators: can create sites and edit metadata/taxonomies

## SharePoint's search and metadata features

- ▶ Default search engine
- ▶ Search refinements/filters
- ▶ Column type metadata, for sorting and filtering tables of documents
- ▶ “Content type” metadata, since 2003
- ▶ Hierarchical taxonomy management “Term Store,” since 2010

Default search in SharePoint (or in general) is useful, but insufficient.



- ▶ What are the issues?



## User Objectives

- ▶ Finding the right information **quickly**
- ▶ Finding **sufficient** information
- ▶ **Discovering** related relevant information not known to search for
- ▶ Being confident that the information is **complete**
- ▶ Being able to **expand and limit** the search
- ▶ Having results that can be **trusted** and are **explainable**

## User Expectations

- ▶ Getting search results like web search results



# Search Issues

---

## Organization Objectives

- ▶ Saved time through efficient search
- ▶ Knowledge discovery
- ▶ Cost savings through content reuse
- ▶ Better decisions from complete information
- ▶ Improved competitiveness through better information
- ▶ Improved employee and customer satisfaction



## Internal Search User Pain Points

- ▶ Retrieving too many inappropriate search results
  - ▷ Wasting time reading through irrelevant information
- ▶ Not confident that that *all* relevant information was retrieved
  - ▷ Not finding content believed to exist
  - ▷ Not knowing *if* the desired information exists
- ▶ Repeating searches in different, unconnected content repositories



# Solution: Controlled Metadata & Taxonomies

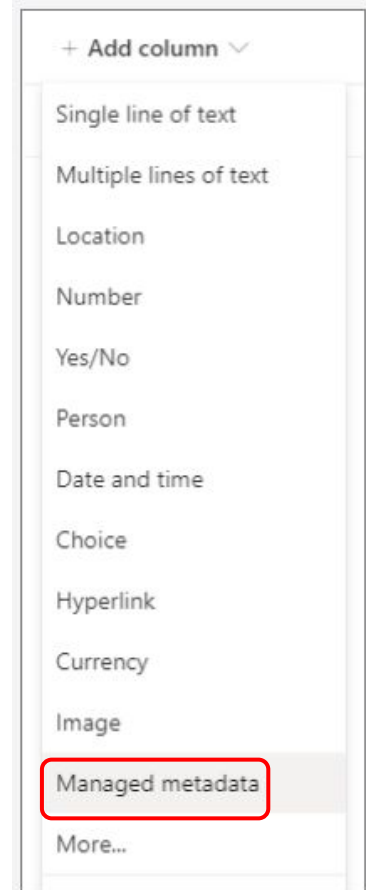
The screenshot shows the SharePoint Term store interface. The left sidebar contains navigation options: Microsoft Search, Content services, Term store (selected), and Content type gallery. The main area is titled 'Term store' and includes a search bar for terms. Below the search bar, a taxonomy tree is displayed under 'Global term groups', with 'ESG Knowledge Model' expanded to show 'ESG risk management' selected. The right pane shows the 'ESG risk management' term details, including tabs for 'General', 'Usage settings', and 'Advanced'. The 'Translations and synonyms' section contains a table with the following data:

Language ↑	Translation	Synonyms	Description
English	ESG risk management	ESG due dilligence, ESG Risk	Investing responsibly, or considering environmental, social and corporate governance (ESG) criteria in

SharePoint  
Term store for  
managing  
metadata/  
taxonomies

# Solution: Controlled Metadata & Taxonomies

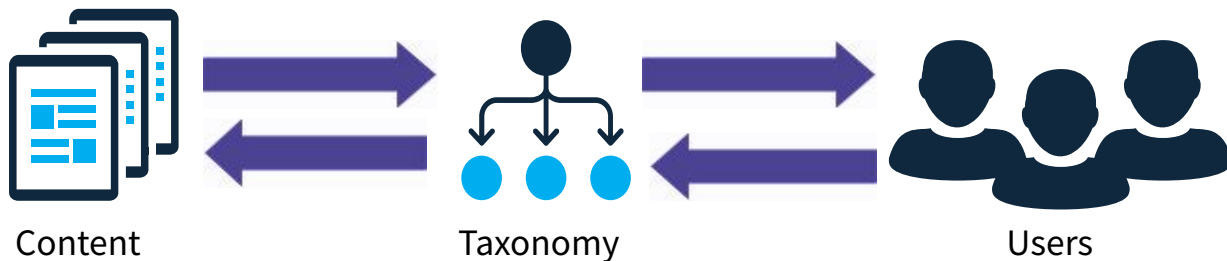
	In SharePoint	Description
Metadata	Column type Content type	Organized data about data/content, grouped into consistent types/elements/properties/fields/columns, filled with a specific value for each content item.
Taxonomy	Taxonomy/ Managed metadata/ Term group	Metadata type that uses a controlled vocabulary and is structured into subsets and/or hierarchies



A taxonomy is a source of values for a metadata type (e.g. “managed metadata”).

# What is a Taxonomy?

- ▶ A taxonomy is a controlled vocabulary organized into a hierarchical structure.
- ▶ Concepts/terms are used to tag/index/categorize content to make it easier to be found and retrieved
  - ▷ supporting better findability than search alone
- ▶ The taxonomy is an intermediary that links the user to the desired content.



- ▶ Terms are based on unambiguous concepts, rather than using text strings.
- ▶ “Things, not strings.”

# What is a Taxonomy?

## Controlled and organized

### 1. Controlled:

A kind of controlled vocabulary or knowledge organization system, based on unambiguous concepts, not just words:  
**things, not strings**

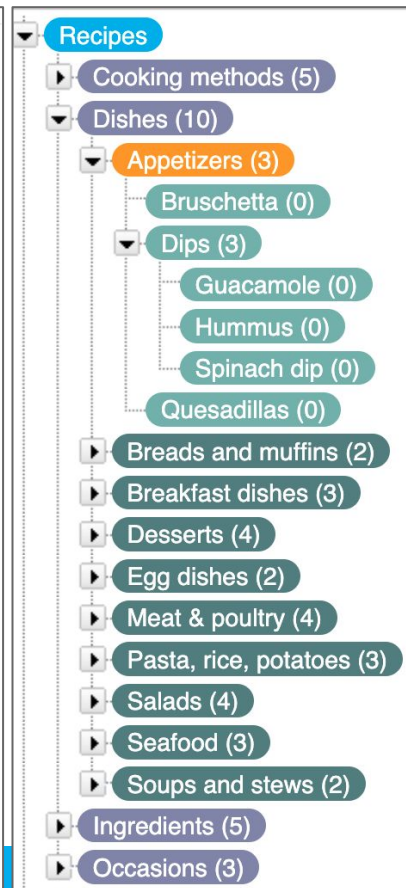
### 2. Organized:

Concepts are arranged in a structure of hierarchies, categories, or facets to organize them.

*controlled*

<b>Preferred Label</b>	
<input type="radio"/> Vorspeisen	de
<input type="radio"/> Appetizers	en
<b>Alternative Labels</b>	
<input type="radio"/> Hors d'oeuvres	en
<input type="radio"/> Starters	
<input type="button" value="+"/>	
<b>Hidden Labels</b>	
<input type="radio"/> Appetisers	en
<input type="radio"/> Horderves	
<input type="button" value="+"/>	
<b>Scope Notes</b>	
<input type="radio"/> Dishes usually served as appetizers	en
<input type="button" value="+"/>	
<b>Definitions</b>	
<input type="radio"/> A small dish of food or a drink taken before a meal or the main course of a meal to stimulate one's appetite	en

*organized*



## 1. Controlled vocabulary aspect



Brings together different wordings (synonyms) for the same concept

- ▶ Helps people search for information by different names

## 2. Classification and structure aspect



Organizes information into a logical structure

- ▶ Helps people browse or navigate for information
- ▶ Provides context and meaning for concepts for indexing and retrieval

# Benefits of Taxonomies

A screenshot of a search interface. At the top, a search bar contains the text 'data' and an orange search button with a magnifying glass icon. Below the search bar, a list of search results is displayed. Each result consists of a bold title and a subtitle in italics. The results are: 'Metadata' (Context: Topics), 'Linked Open Data' (Context: Topics), 'Linked Data enrichment' (Context: PoolParty Product), 'Output data unit' (Context: PoolParty Product), 'Data processing' (Context: PoolParty Product), and 'Data engineering' (Context: Topics).

Controlled:  
PoolParty product Help  
<https://help.poolparty.biz>

## Topics (8)

- ▶ Standards - 1029
- ▶ Activities and methods - 1028
- ▼ Data - 1000
  - Metadata - 345
  - ▶ Unstructured data - 213
  - ▼ Structured data - 117
    - ▶ Spreadsheet - 99
    - Relational data - 19
  - ▼ Open data - 72
    - Linked Open Data - 52
  - Linked Open Data - 52
  - Master data - 42
  - Data structures
  - Sensor data - 7
  - Streaming data
- ▼ Knowledge organization systems - 898
  - ▶ Taxonomies - 598
  - ▶ Knowledge graphs - 405
  - Ontologies - 318

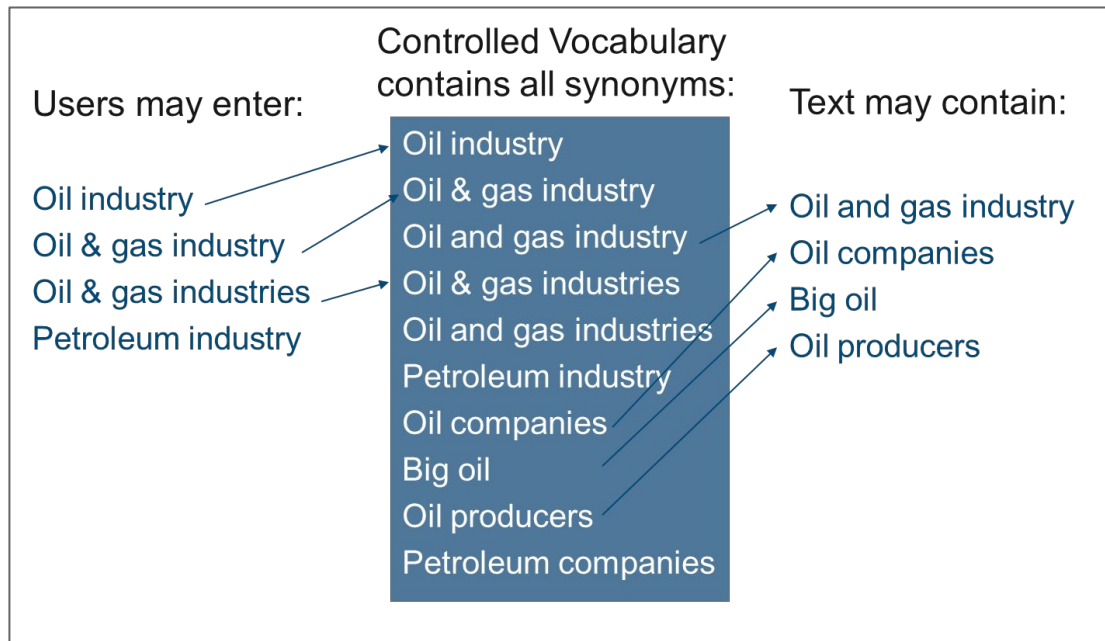
Organized:  
Semantic Web Company  
Knowledge Hub Topics



# Benefits of Taxonomies

## Bringing together synonyms for the same term

- ▶ Retrieves content containing different synonyms for the same thing.
- ▶ Supports searchers using different synonymous search strings find the same relevant content.



## Benefits over search alone

- ▶ Tagging and retrieval based on concepts, not just words/phrases, for better precision (accuracy) and recall (comprehensiveness)
- ▶ Grouping terms by type, for search refinement filters (facets)
- ▶ Organizing terms into guiding hierarchies
- ▶ Supporting retrieval of tagged non-text content
- ▶ Enabling discover of related content tagged with the same concepts
- ▶ Supporting term definitions
- ▶ Supporting multilingual concepts
- ▶ Potential to link to other taxonomies and thus other content/data

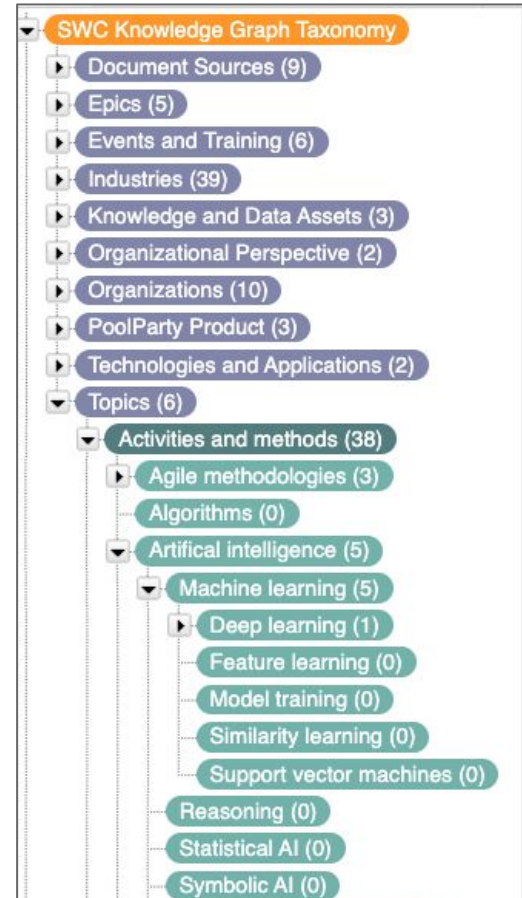
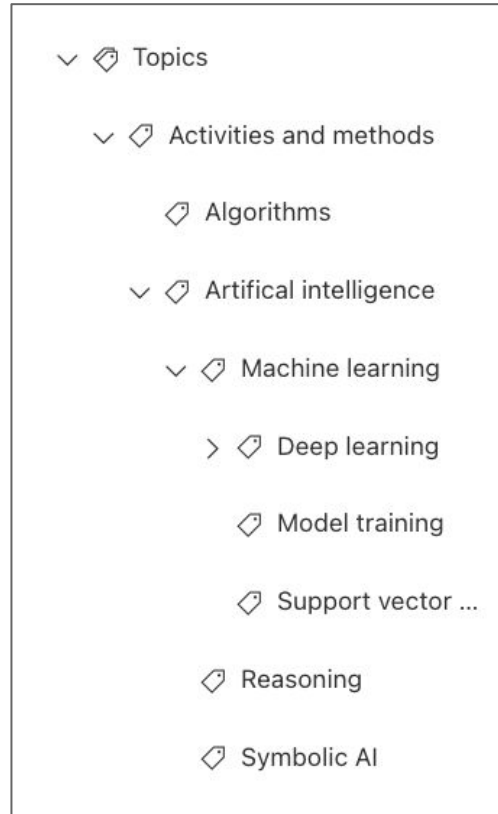
# **Taxonomy Use and Management in SharePoint**

SharePoint can make use of taxonomies various ways

1. Search support and as facets or post-search refinement/filters
2. Metadata terms for filtering or sorting columns of content items (documents, images, etc.) in document libraries
3. Implemented in navigation as menu labels, headers, and page URL names
4. Also informally, as a hierarchical model for naming and organizing libraries and multi-level folders

# Taxonomies in SharePoint

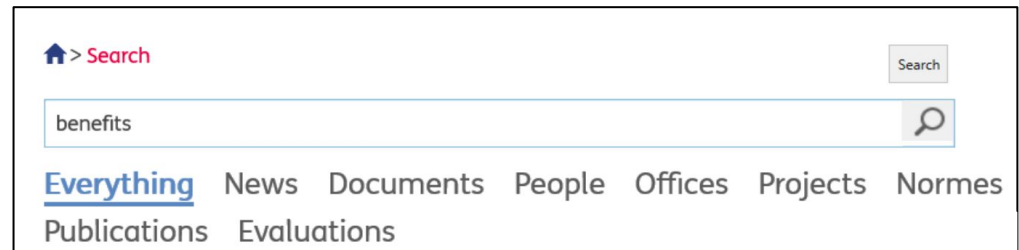
Taxonomy hierarchy in the SharePoint Term Store and in a dedicated taxonomy management system (PoolParty)



## Advantages of SharePoint

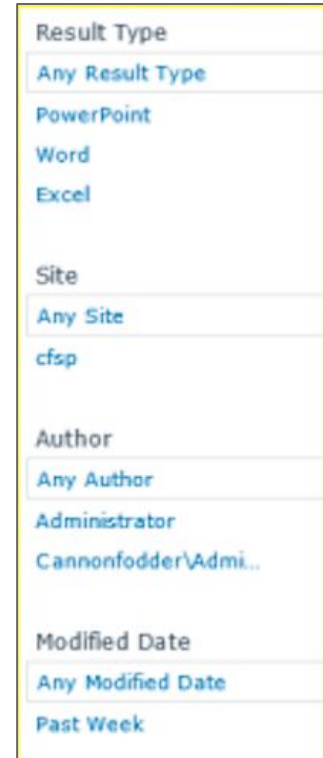


- ▶ Departments and teams can create their own sites.
- ▶ Access to sites can be shared or not (usually shared).
- ▶ Different user access allows some to edit and more to view.
- ▶ Editing users can create taxonomies to organize their team's content.
- ▶ Additional, uncontrolled keywords can be shared across the enterprise for reuse.
- ▶ Search can be executed at the site or enterprise level.

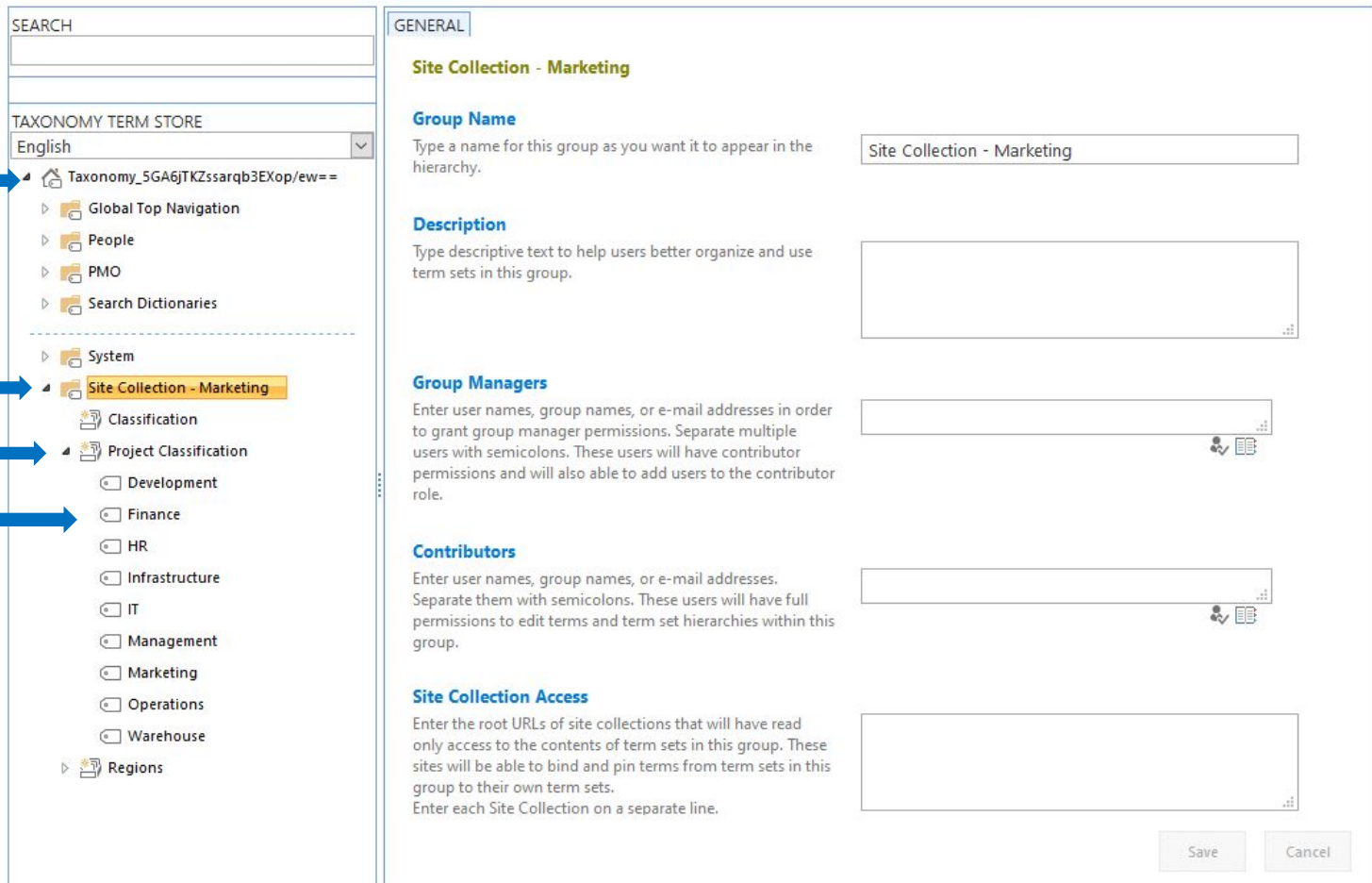


## Disadvantages of SharePoint

- ▶ Different departments have their own term sets for tagging
  - ▷ Term sets in different sites use different terms for the same concept, and the same term for different concepts.
  - ▷ Search across the enterprise thus cannot leverage a taxonomy.
- ▶ Search refinement filters are only the default (file type, creator, date, etc.) and not based on taxonomy facets



# Taxonomy Management in SharePoint



The screenshot shows the SharePoint Taxonomy Management interface. On the left, the 'TAXONOMY TERM STORE' is displayed with a tree view. A search bar is at the top. The tree view includes categories like 'Global Top Navigation', 'People', 'PMO', 'Search Dictionaries', 'System', and 'Site Collection - Marketing'. Under 'Site Collection - Marketing', there are 'Classification' and 'Project Classification' term sets. 'Project Classification' has terms like 'Development', 'Finance', 'HR', 'Infrastructure', 'IT', 'Management', 'Marketing', 'Operations', and 'Warehouse'. 'Regions' is also listed. On the right, the 'GENERAL' tab is active, showing the configuration for the 'Site Collection - Marketing' group. It includes fields for 'Group Name' (filled with 'Site Collection - Marketing'), 'Description', 'Group Managers', 'Contributors', and 'Site Collection Access'. 'Save' and 'Cancel' buttons are at the bottom right.

Term Store

Group

Term Set

Term

Group - set of terms with shared users

Term set - controlled vocabulary or facet, can be used as a filter



# SharePoint Term Store Capabilities

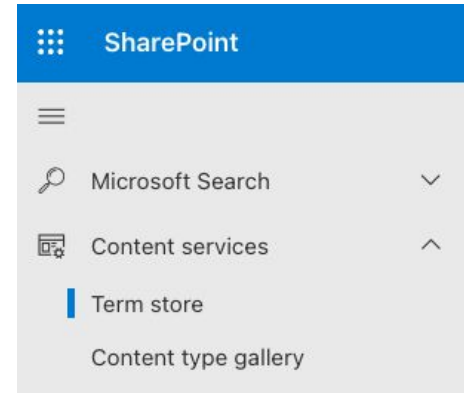
Supporting term definitions/notes for clarification

Data management			
General	Usage settings	Advanced	
Translations and synonyms			
+ Add			
Language ↑	Translation	Synonyms	Description
English	Data management	Data wrangling, Integrated data management	Data management comprises all the disciplines related to managing data as a valuable resource.

# SharePoint Term Store Capabilities

## Taxonomy management in SharePoint Term Store features

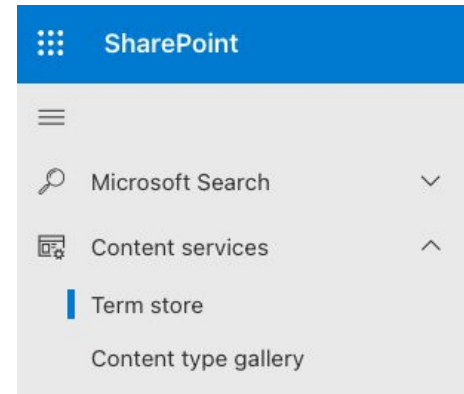
- ▶ Multiple taxonomies or multiple facets
- ▶ Unique identifiers on terms
- ▶ Hierarchical taxonomies, many levels deep
- ▶ Synonyms, also called “Other labels,” in support of tagging
- ▶ Scope notes or definitions in a “Description” field
- ▶ Type-ahead search on taxonomy terms and on synonyms for those tagging
- ▶ Broader categories, which are not terms for tagging, to group terms
- ▶ Customizable sort order of terms at the same level
- ▶ Term store reports (with added Syntex license)



# SharePoint Term Store Weaknesses

## Taxonomy features, which the SharePoint Term Store lacks

- ▶ Synonyms supported in search (not only in tagging)
- ▶ Polyhierarchy: more than one broader in same term set
- ▶ Associative (non-hierarchical) relationships
- ▶ Thesaurus quality standards (ANSI/NISO Z39.19)
  - ▶ e.g. preventing duplicates, circular references
- ▶ Additional notes and attributes on concepts
- ▶ SKOS interoperability, SKOS export (only SKOS import with added Syntex license)



A light grey background with a network diagram consisting of various sized circles connected by thin lines, representing a complex web or data structure.

# Improving Taxonomy Management in SharePoint

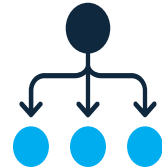
# Challenges with Metadata & Taxonomies

---

1. Poorly designed and managed metadata and taxonomies
2. Poorly applied metadata or taxonomy tagging
3. Taxonomies and metadata residing within separate applications, siloed

## 1. Poorly designed and managed metadata and taxonomies

- ▷ Metadata or taxonomy terms vague or not clearly distinguished
- ▷ Taxonomies too specific or too general
- ▷ Too many, or too few metadata columns or term sets
- ▷ Incorrect hierarchies
- ▷ Too many or too few synonyms for terms



### ▶ Addressed by:

- ▷ Elevate taxonomy development to a funded project or role
- ▷ Obtain taxonomy training and/or contract a taxonomy consultant
- ▷ Manage taxonomies in a full-featured taxonomy management system, connected to SharePoint and other applications

## 2. Poorly applied metadata or taxonomy tagging

- ▷ Insufficient tagging - content cannot be found
- ▷ Over-tagging - taxonomy terms retrieve too many irrelevant results
- ▷ Inconsistent tagging - search results aren't comprehensive

### ▶ Addressed by training and auto-tagging

- ▷ For more comprehensive, consistent, and faster tagging
- ▷ Auto-tagging technologies: text mining, NLP, NER
- ▷ Auto-tagging with human review to identify improvements



## 3. Siloed taxonomies and metadata in separate sites and applications

- ▶ Separate taxonomies for separate SharePoint sites
- ▶ Separate taxonomies for each CMS or other application
- ▶ Taxonomies managed by different groups
- ▶ Lack of central taxonomy governance



Term store	
<input type="text" value="Search terms"/>	
<input checked="" type="checkbox"/> Taxonomy	
<input type="checkbox"/> Global term groups	
<input type="checkbox"/> Site level term groups	

### ▶ Addressed by:

- ▶ Centrally managing taxonomy and metadata in dedicated taxonomy management system, as middleware, connected to other systems
- ▶ Documentation and following a taxonomy governance plan to manage taxonomy change



Centrally managed taxonomies in an integrated dedicated taxonomy management system

## Benefits

- ▶ Full taxonomy features
- ▶ Consolidation and alignment of enterprise-wide taxonomies
- ▶ SKOS and ANSI/NISO compliant taxonomies
- ▶ Superior taxonomy management features and UI

## Additional possible options:

- ▶ Semantic auto-tagging
- ▶ Semantic search and UI enhancements

# Superior Taxonomy Management Features

What can be built on top of SharePoint to extend Term Store functionality in dedicated taxonomy management middleware:

- ▶ Synonyms supported in search in addition to tagging
- ▶ Supporting thesaurus quality standards (preventing duplicates, circular references, same terms as both preferred and synonyms)
- ▶ Managing multiple linked taxonomies
- ▶ Supporting associative (non-hierarchical relationships)
- ▶ Supporting the display of hierarchy within facets.
- ▶ Supporting polyhierarchy
- ▶ Supporting additional attributes for terms
- ▶ Full compliance with W3C standards for vocabularies (SKOS) for import/export interoperability
- ▶ Link to content tagged in different systems (other than SharePoint), with a central frontend
- ▶ Optimized taxonomy management UI

# Superior Taxonomy Management Features and UI

The screenshot displays the Superior Taxonomy Management interface, which is divided into several key sections:

- Term store (Left Panel):** A search interface for terms, currently showing a taxonomy tree. The tree is expanded to 'Machine learning', which is highlighted in grey. Other visible categories include 'Technologies and Applications', 'Topics', 'Activities and methods', 'Artificial intelligence', 'Reasoning', 'Symbolic AI', 'Tensor network theory...', 'Content management', 'Data governance', 'Data management', and 'Database querying'.
- Machine learning Detail View (Right Panel):** A detailed view of the 'Machine learning' concept. It includes:
  - General Information:** Title 'Machine learning', a URL, and actions like 'Add to Collection', 'Add to Blacklist', 'Add to ExactMatch', and 'Delete Concept'.
  - Details Tab:** Shows the concept is part of the 'SKOS' scheme and 'SWC KG - GS Scheme'.
  - Broader Concepts:** Lists 'Artificial intelligence' as a broader concept.
  - Narrower Concepts:** Lists 'Deep learning', 'Feature learning', 'Model training', 'Similarity learning', and 'Support vector machines' as narrower concepts.
  - Related Concepts:** Lists 'Algorithms' and 'Data mining and machine learning software' as related concepts.
  - Top Concept of Concept Schemes:** A section for top concepts.
  - Preferred Label:** 'Machine learning' (en).
  - Alternative Labels:** A list of alternative labels such as 'Adaptive machine learning', 'Computer machine learning', 'Dictionary learning', 'Feature discovery', 'Learning algorithm', 'Learning algorithms', 'List of open-source machine learning software', 'ML', and 'Statistical learning'.
  - Hidden Labels:** A section for hidden labels.
  - Scope Notes:** A section for scope notes.
  - Definitions:** A definition: 'Machine learning is the subfield of computer science that gives computers the ability to learn without being explicitly programmed' (Arthur Samuel, 1959). Evolved from the study of pattern recognition and computational learning theory in artificial intelligence, machine learning explores the study and construction of algorithms that can learn from and make predictions on data -
- Central Panel:** A hierarchical tree view of the 'SWC Knowledge Graph Taxonomy'. The 'Machine learning' node is highlighted in orange. Other nodes include 'Document Sources', 'Epics', 'Events and Training', 'Industries', 'Knowledge and Data Assets', 'Organizational Perspective', 'Organizations', 'PoolParty Product', 'Technologies and Applications', 'Topics', 'Activities and methods', 'Agile methodologies', 'Algorithms', 'Artificial intelligence', 'Deep learning', 'Feature learning', 'Model training', 'Similarity learning', 'Support vector machines', 'Reasoning', 'Statistical AI', 'Symbolic AI', 'Tensor network theory', 'Content management', 'Contextual advertising', 'Data analytics', 'Database querying', 'Data engineering', 'Data governance', 'Data management', 'Design thinking', 'E-learning', 'Enterprise architecture', and 'Graph embedding'.

“Optimizing Your Taxonomy in SharePoint Online: Search Filters” Mergan Salerno, Enterprise Knowledge White Paper (September 2022)

[https://enterprise-knowledge.com/wp-content/uploads/2022/09/Optimizing\\_Your\\_Taxonomy\\_in\\_SharePoint\\_Online\\_Search\\_Filters.pdf](https://enterprise-knowledge.com/wp-content/uploads/2022/09/Optimizing_Your_Taxonomy_in_SharePoint_Online_Search_Filters.pdf)

“How to Manage Term Stores in SharePoint Online”

<https://www.mrsharepoint.guru/term-store-sharepoint/>

“Term Store: Introduction to Management Metadata” Microsoft Learn

<https://learn.microsoft.com/en-us/sharepoint/managed-metadata>

“Taxonomies in SharePoint,” Heather Hedden, Accidental Taxonomist Blog (August, 2017)

<http://accidental-taxonomist.blogspot.com/2017/08/taxonomies-in-sharepoint.html>

*The Accidental Taxonomist*, 3rd edition, Heather Hedden (2022)

<https://accidental-taxonomist.com>

# Questions/Contact

---

## Heather Hedden

Knowledge Manager

Semantic Web Company Inc.

One Boston Place, Suite 2600

Boston, MA 02108

857-400-0183

[heather.hedden@semantic-web.com](mailto:heather.hedden@semantic-web.com)

[www.linkedin.com/in/hedden](http://www.linkedin.com/in/hedden)

Semantic Web Company [www.semantic-web.com](http://www.semantic-web.com)

PoolParty software [www.poolparty.biz](http://www.poolparty.biz)

PoolParty for SharePoint <https://www.poolparty.biz/poolparty-for-sharepoint>

