

Building Taxonomies from the Bottom Up and Top Down

Heather Hedden

Library Science Talks

Zentralbibliothek Zürich

CERN Scientific Information Service

Association of International Librarians and Information Specialists (AILIS), Geneva

29 September 2026

Outline

- § Introduction to Taxonomies
 - § Features and Uses of Taxonomies
- § Taxonomy Creation Approaches
 - § Bottom-up Taxonomy Creation
 - § Top-down Taxonomy Creation
- § Conclusions

Introduction to Taxonomies

Knowledge management
Knowledge
Knowledge-based systems
Knowledge transfer
Knowledge workers
Knowledge economy
KnowledgeWare Inc.
Knowledge acquisition (Expert systems)
Knowledge Adventure Inc.
Knowledge Networks Inc.

data

Graph databases
Context: Technologies and Applications

Data engineering
Context: Underlying Concepts

Linked Data Harvesting
Context: Features & Components

Data processing
Context: Underlying Concepts

Linked Data publishing
Context: Features & Components

Gale Academic Onefile

Topic

☐ Advanced Content (73)
☐ Agile, Design Thinking, & Facilitation (60)
☐ Artificial Intelligence (25)
☐ Change Management & Communications (23)
☐ Company (19)
☐ Content & Brand Strategy (5)
☐ Enterprise Learning (23)
☐ Enterprise Search (49)
☐ Knowledge Graphs & Data Modeling (108)
☐ Knowledge Management Strategy & Design (243)
☐ Taxonomy & Ontology Design (113)
☐ Technology Solutions (98)

Article Type

☐ Blog (324)
☐ Podcast (71)
☐ Presentation (56)
☐ Case Study (40)
☐ White Paper (31)

– Computer and information sciences

Artificial intelligence

- Artificial neural networks
- Expert systems
- Genetic programming
- Machine learning

– Artificial intelligence

Machine learning

- Decision tree learning
- Deep learning
- Ensemble methods
- Machine learning algorithms
- Relevance vector machines
- Supervised machine learning
- Support vector machines
- Unsupervised machine learning

PLOS One

Enterprise Knowledge
website
knowledge
base

Taxonomies help people find information.

© 2026 Hedden Information Management

3



WORLD TRADE
ORGANIZATION

Search

Home
About WTO
News and events
Trade topics
WTO membership
Documents, data and resources
WTO and you

GOODS

- Agriculture
 - Agriculture negotiations
 - Cotton
- Anti-dumping
 - Trade Remedies Data Portal
- Balance of payments
- Customs valuation
- Fisheries subsidies
- Goods Council
- Goods schedules
- Import licensing
- Information Technology Agreement
- Market access for goods
 - Market access negotiations
 - Quantitative restrictions
- Non-tariff measures
- Pharma Agreement
- Preshipment Inspection
- Rules of origin
- Safeguard measures

- Sanitary & phytosanitary measures
- State trading enterprises
- Subsidies & countervailing measures
 - Trade Remedies Data Portal
- Tariffs
- Technical barriers to trade
- Textiles
- Trade facilitation
- Trade-Related Investment Measures (TRIMs)

SERVICES

- Services negotiations
- Services schedules
- I-TIP services
- TS4D platform

INTELLECTUAL PROPERTY

- Specific TRIPS issues

DISPUTE SETTLEMENT

- Disputes chronologically
- Disputes by country/territory
- Disputes by agreement
- Disputes by subject
- Find disputes cases
- Dispute Settlement Body
- Appellate Body

REGIONAL TRADE AGREEMENTS

- Transparency Mechanism
- RTA Database

DOHA DEVELOPMENT AGENDA

- Subjects treated
- Trade Negotiations Committee
- Groups in the negotiations

COVID-19 AND WORLD TRADE

BUILDING TRADE CAPACITY

- Development
 - Aid for Trade
- Enhanced Integrated Framework
- Standards and Trade Development Facility (STDF)
- Technical assistance and training
 - E-Learning website
- Trade Facilitation Agreement Facility

TRADE MONITORING

- Trade monitoring reports
- Trade policy reviews

WTO REFORM

MORE TOPICS

- Accessions
- Civil aircraft
- Competition policy
- Electronic Commerce and Digital technologies

Environment

- Environmental Database
- Climate change
 - Plastics pollution and environmentally sustainable plastics trade
- Trade and environmental sustainability
 - Fossil fuel subsidy reform
- Government procurement
- Investment
- Notifications Portal
- Preferential trade arrangements
- Small business and trade
 - Informal Working Group on MSMEs
- Trade finance
- Trade for Peace
- Women and trade
 - Informal Working Group on Trade and Gender

JOINT INITIATIVES

- Electronic Commerce
- Investment facilitation for development
- Services Domestic Regulation

Introduction to Taxonomies

WIPO				English ▾		IP Portal login
Understand & Learn ▾ Find & Explore ▴ Protect & Manage ▾ Partner & Collaborate ▾ About WIPO ▾						
Patent & Technology Information	Trademark Information	Design Information	Geographical Indication Information			
PATENTSCOPE	Global Brand Database	Global Design Database	Lisbon Express Database			
Patent Analytics	Madrid Monitor	International Designs Bulletin	Global Brand Database for GIs			
International Patent Classification	Article 6ter Express Database	Hague Express Database				
ARDI – Research for Innovation	Nice Classification	Locarno Classification				
ASPI – Specialized Patent Information	Vienna Classification					
Plant Variety Information (UPOV)	IP Laws, Treaties & Judgments	IP Resources	IP Reports			
PLUTO Plant Variety Database	WIPO-Administered Treaties	WIPO Standards	World Intangible Investment Highlights			
GENIE Database	WIPO Lex - IP Laws, Treaties & Judgments	IP Statistics	WIPO Technology Trends			
		WIPO Pearl (Terminology)	Global Innovation Index			
		WIPO Publications	World Intellectual Property Report			
		Country IP Profiles				

Introduction to Taxonomies

Taxonomy definition:

a collection of controlled vocabulary terms organized into a hierarchical structure

– ISO 25964-1:2025 *Thesauri and interoperability with other vocabularies — Part 1: Thesauri for information retrieval, management, and use. Section 3 Terms and definitions*

What is a taxonomy not?

- | | |
|---------------------------------|---|
| § Not a navigation scheme | Ø For searching, not just browsing |
| § Not just any metadata or tags | Ø Needs to be controlled, structured, related |
| § Not a classification scheme | Ø For tagging, not classifying |
| § Not business glossary | Ø For finding, not defining |

Introduction to Taxonomies

How taxonomies compare to other knowledge organization systems

Compared to Classification Systems:

- § Taxonomies resemble classification systems with their hierarchy.
- § Taxonomies don't have "miscellaneous" or "other" categories.
- § Information taxonomies comprise "concepts," not actual things, and are digital.
 - § A concept can be in more than one hierarchy (subclass to more than one class).
 - § Content can be tagged with more than one concept.

Dewey Decimal Classification

000 GENERAL WORKS 020 Library and Information Science 030 General Encyclopaedias 050 General Periodicals 060 General Organizations	600 TECHNOLOGY 610 Medical Sciences 620 Engineering 630 Agriculture 640 Domestic Sciences 650 Business and Management 660 Chemical Technology 670 Manufactures 680 Building Construction
100 PHILOSOPHY 110 Metaphysics 120 Psychology, occultism 130 Philosophy 140 Psychology 150 Psychology 160 Logic	700 THE ARTS 710 Landscape and Civic Art 720 Architecture 730 Sculpture, Plastics 740 Drawing, Decorative Arts 750 Painting 760 Prints and Book Making 770 Photography 780 Music 790 Recreation, Performing Arts
200 RELIGION 220 The Bible 230 Christian Doctrine 290 Comparative and Other Religions	800 LITERATURE 810 American Literature 820 English Literature 830 German Literature 840 French Literature 850 Italian, Portuguese Literature 860 Spanish, Portuguese Literature 870 Latin and Other Italic Literatures 880 Classical and Modern Greek Literature 890 Other Literature
300 SOCIAL SCIENCES 310 Statistics 320 Political Science 330 Economics 340 Law 350 Public Administration 360 Social Welfare 370 Education 380 Public Service 390 Customs and Folklore	900 HISTORY, GEOGRAPHY

Compared to Thesauri:

- § The required associative (related term) relationship is less common in taxonomies.
- § Taxonomies have an over-arching hierarchical structure, but many thesauri do not.
- § There are published standards for thesauri, but taxonomies are less strict.

pesticides	UF: fumigants BT: agrochemicals NT: fungicides RT: insecticides	plant products	NT: cereals fruits spices vegetables RT: plants
pests	NT: pest insects plant pests pesticides	plants	RT: plant products
pigs	UF: hogs BT: porkers sows livestock	porkers	USE: pigs
plant pests	BT: pests	poultry	BT: livestock NT: chickens ducks geese turkeys RT: eggs
		sheep	BT: livestock RT: wool

Compared to Subject Heading Schemes:

- § Subject heading schemes have evolved to have more features of thesauri.
- § Taxonomies and thesauri do not support structured indexing/cataloguing of heading + subdivision combinations (e.g. *Accounting — Law and legislation*)

LIBRARY

Search results for "Accounting" (10 results)

Accounting 1. The art and science of recording financial transactions 2. The system of recording financial transactions 3. The system of recording financial transactions 4. The system of recording financial transactions 5. The system of recording financial transactions 6. The system of recording financial transactions 7. The system of recording financial transactions 8. The system of recording financial transactions 9. The system of recording financial transactions 10. The system of recording financial transactions	Related terms Accounting Accounting Accounting Accounting Accounting Accounting Accounting Accounting Accounting Accounting
---	--

Introduction to Taxonomies

Taxonomy Structural Types

Leisure and culture

- . Arts and entertainment venues
 - . Museums and galleries
- . Children's activities
- . Culture and creativity
 - . Architecture
 - . Crafts
 - . Heritage
 - . Literature
 - . Music
 - . Performing arts
 - . Visual arts
- . Entertainment and events
- . Gambling and lotteries
- . Hobbies and interests
- . Parks and gardens
- . Sports and recreation
 - . Team sports
 - . Cricket
 - . Football
 - . Rugby
 - . Water sports
 - . Winter sports
- . Sports and recreation facilities
- . Tourism
 - . Passports and visas
- . Young people's activities

Hierarchical Taxonomy Example

Career Level

- Student
- Entry Level
- Experienced
- Manager
- Director
- Executive

Function

- Customer Service & Support
 - Delivery
 - Engineering
 - Finance
 - General Management
 - Legal & Regulatory Affairs
 - Marketing & Advertising
- [\[more\]](#)

Industry

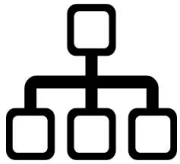
- Agriculture
 - Apparel & Fashion
 - Automotive
 - Aviation & Aerospace
 - Banking
 - Biotechnology
 - Broadcast Media
 - Chemicals
- [\[more\]](#)

Faceted Taxonomy Example

Features of Taxonomies

Two major approaches to (or features of) taxonomies

1. A hierarchy of terms/topics/categories arranged with narrower topics/subcategories displayed under their broader/parent categories.



- ∅ To guide users to find the desired topic (and its linked content of pages or documents)
- ∅ Similar to website navigation, but more topical, not just based on page titles
- ∅ Similar to a classification system, but allows multiple classifications

2. A controlled vocabulary of metadata tags/labels to apply to pages, posts, or documents, so that they can be more precisely and comprehensively retrieved (than by search algorithms alone on keywords in text).



- ∅ Implemented as search suggestion terms, search refinement filters, or related topics and searches

Features of Taxonomies

Taxonomy focused on controlled

data

Graph databases
Context: Technologies and Applications

Data engineering
Context: Underlying Concepts

Linked Data Harvesting
Context: Features & Components

Data processing
Context: Underlying Concepts

Linked Data publishing
Context: Features & Components

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 - 21

Sensor data - 7

Streaming data - 7

▼ Knowledge organization systems - 898

> Taxonomies - 598

> Knowledge graphs - 405

Taxonomy focused on organized

© 2026 Hedden Information Manag

10

Features of Taxonomies

Controlled and organized in taxonomy management systems

Search...

Content Type

Level

Role

Skill

Business skills

Languages

Soft skills

Communication skills

Client relations skills

Interviewing skills

Public speaking skills

Writing skills

Observation skills

Organizational skills

Social skills

Technical skills

Automation

Programming

Communication skills

Concept Class

Skill

Preferred Labels

Create a preferred label

Communication skills

Alternative Labels ²

Create an alternative label

alternative label > Communicating well en

alternative label > Communications skills en

Metadata

Add metadata field

definition

The ability to communicate effectively with other people for business purposes. en

Organized

Controlled

Recipes

Cooking methods (5)

Dishes (11)

Appetizers (2)

Bruschetta (0)

Dips (0)

Breads and muffins (2)

Breakfast dishes (3)

Desserts (4)

Ingredients (4)

Dairy Products (4)

Fruits (16)

Grains (10)

Vegetables (20)

Occasions (4)

Lists

Collections

GraphEditors

Appetizers

<http://advanced.poolparty.biz/FoodandRecipes/279>

+ Create Narrower Concept...

Delete Concept

+ Add to Exclude-list

+ Add to ExactMatch

Details

Notes

Documents

Linked Data

Triples

Visualization

Quality Management

SKOS

+

Broader Concepts

Narrower Concepts

Related Concepts

Top Concept of Concept Schemes

Preferred Label

Alternative Labels

Hidden Labels

Scope Notes

Organized

Controlled

Select a narrower concept

has narrower > Client relations skills

has narrower > Interviewing skills

has narrower > Public speaking skills

has narrower > Writing skills

PoolParty

Semaphore

© 2026 Hedden Information Management

Features and Uses of Taxonomies

1. Controlled vocabulary feature



Brings together different wordings (synonyms) for the same concept (for tagging).

- Helps people search for information by different names

2. Classification and organized structure feature



Organizes information into a logical structure.

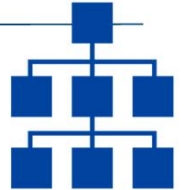
- Helps people browse for information
- Provides context and meaning for concepts, to support indexing/tagging, retrieval, and AI applications

Helps improve search results

- § Retrieving only correct results
- § Retrieving all relevant results
- § Provides consistency to retrieve across different, separate content repositories

Uses of Taxonomies

What you can do with a taxonomy

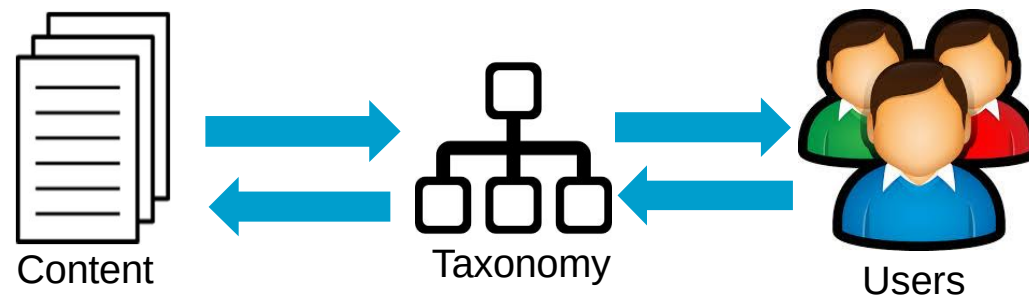


- § Search: find content about.... (search string matches taxonomy concepts)
- § Topic browse: explore subjects arranged in a hierarchy and then content on the subject
- § Faceted (filtering/refining) search: find content meeting a combination of basic criteria
- § Tagging: index content consistently so that retrieval is comprehensive and accurate
- § Discovery: find other content tagged with same concepts as tagged to found content; explore broader, narrower, and (sometimes) related taxonomy topics
- § Content curation: create feeds or alerts based on user profile (personalization) and taxonomy concepts tagged
- § Metadata management: for retrieval, identification, comparison, analysis, etc.
- § Knowledge graph development: when combined with an ontology, to retrieve data and content across different repositories

Uses of Taxonomies

What is a taxonomy for?

- § Concepts/terms are used to tag/index/categorize pages or content to make them easier to be found and retrieved
 - Ø supporting better findability than search alone
- § The taxonomy is an intermediary that links the user to the desired content.
- § The taxonomy should suit the context of the users and the content.



- § Taxonomy design should take into consideration the users and the content.

Taxonomy Creation Approaches

Generic (Traditional) Taxonomy Design

Bottom Up

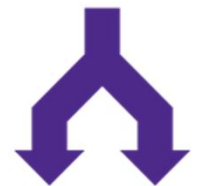
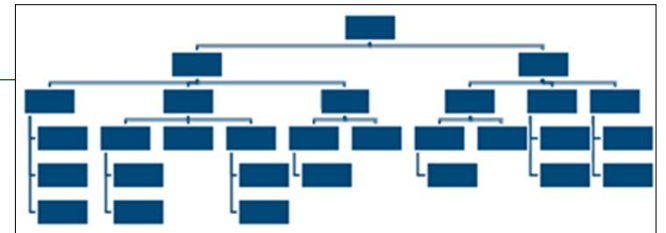
1. Identify the specific concepts needed for tagging the content.
2. Then group the concepts into categories and then broader categories, which form the hierarchies and/or facets.

Top Down

1. Identify the top concepts/terms, hierarchies and/or facets needed to cover the subject domain.
2. Then add narrower concepts, and further narrower concepts to build out the hierarchies to the level of detail desired.

They can be undertaken simultaneously.

In combination, they can meet (and overlap) in the middle.



Taxonomy Creation Approaches

Modern “Business” Taxonomy Design: Content and User-Focused

Bottom-Up Methods Focus on:

- § The concepts, terms, entities within the content that will be searched for and retrieved with the taxonomy
- § Content analysis, tagging/indexing
- § The back end



Top-Down Methods Focus on:

- § Users, use cases, the application, the subject domain scope
- § Information architecture, user experience, taxonomy displays
- § The front end



Both approaches are needed in combination.

Methods for creating a new taxonomy or expanding an existing one

Taxonomy Creation Approaches

Value of Business Taxonomy Design Approaches

Value of Bottom-Up Methods:

- § Alignment - Taxonomy concepts are reflected and accurately distributed across content.
- § Completeness - The taxonomy concepts are applicable to the complete set of content across the system.



Value of Top-Down Methods:

- § Usability - The structure and language of the taxonomy resonate with and are intuitive to end users.



Bottom-Up Taxonomy Creation

To ensure that the taxonomy is applicable to relevant content and data sources and meets standards.

Content-Focused Methods

- § Manual content analysis/survey
- § Automated term extraction
- § Column headers from data in spreadsheets
- § Extracted list of document titles

Content Management-Focused methods (content managers, taggers)

- § Review of existing controlled vocabularies and glossaries
- § Term lists suggested by content creators/editors
- § Keyword tagging analysis



Bottom-Up Content-Focused Methods

Manual survey of sample content

- § From a representative sample of the content to be tagged with the taxonomy
 - § Of different types, sources, file formats, etc.
 - § Look for terms especially within titles, section headings, lead paragraphs.
 - § Look for main idea terms, as if you were indexing or categorizing the content.
 - § Keep it general for the content item/document/page as a whole
 - § Consider desired search strings to retrieve the content item.
 - § Consider different aspects: activity, location, event, person type,
- Ø Manual survey can often get the main idea more reliably than automated methods.



	L	N	O	P	Q	R	S	T	U	V
		<u>Population (Patients)</u>	<u>Medical procedure</u>	<u>Music interaction (Recommend changing this to cover other</u>	<u>Health Aspect or Condition</u>	<u>Music type or genre</u>	<u>Study method</u>	<u>Outcome Measures</u>	<u>Setting</u>	<u>Other</u>
1	Abstract setting as an adjunct to current treatment modalities. Previous studies have indicated that some types of music may benefit patients by reducing pain and anxiety, and may have an effect on physiological measures. OBJECTIVE: To evaluate the scientific foundation for the implementation of a complementary therapy, harp playing. The research questions for this pilot study were: Does live harp playing have an effect on patient	Post-operative patients (Post-surgical patients)	Vascular and thoracic surgery	Post-operative inpatient music listening	surgical pain, anxiety	Live harp playing	Quasiexperimental, longitudinal (repeated)	visual analog scale, patient satisfaction questionnaire, heart rate, systolic and diastolic blood pressure, respiratory rate, and oxygen saturation	Hospital Vascular Thoracic Unit	Patient satisfaction
10	common in cardiac surgery patients. Studies have suggested that music can decrease anxiety in hospitalized patients. Primary Study Objective This study focused on the efficacy and feasibility of special music, which included nature sounds, for pain and anxiety. METHODS/DESIGN: In this randomized controlled trial, postoperative cardiovascular surgery patients were	Post-surgical patients (post-operative patients)	Cardiac surgery	Post-surgery music listening	Surgical pain, anxiety, relaxation	Special music including nature sounds	Randomized control trial	Visual analog scales	Hospital Cardiovascular surgical unit	Nature sounds, Patient satisfaction
11	experiencing medical procedures. Complementary and alternative medical (CAM) therapies have become increasingly important in treating children's painful conditions, yet it is still unclear whether CAM has a place in acute pediatric pain analgesia. This review aims to present an overview of the available published evidence. Most research has examined the efficacy of hypnosis,	Pediatric patients (children)	Medical procedures		acute medical procedure pain		review			Complementary and alternative medical therapies (Complementary and alternative medicine)
12	becoming an increasingly common treatment for a variety of hematologic disorders. The treatment process is not benign. Both physiologic and psychological regimen-related side effects are common, painful, and even life threatening. Music therapy is the prescribed use of music to aid in the prevention or amelioration of physical, psychological, or cognitive problems. Relaxation imagery, systematic review of the efficacy of music therapy (MT) on pain and anxiety in children undergoing clinical procedures. METHODS: We searched 16 electronic databases of published and unpublished studies, subject bibliographies, reference lists of relevant articles, and trials registries. Two reviewers independently screened	Medical procedure patients	Bone marrow transplant	Music listening sessions	bone marrow transplant pain and nausea		Case controlled study	Visual analog scale		relaxation imagery
13		Pediatric medical procedure patients,		Music listening sessions; Music therapist involved;	Clinical procedure pain; clinical		Review of multiple trials (Literature)			Active music therapy, Passive

Identified terms then grouped into types for developing the Berklee Music & Health taxonomy

Bottom-Up Content-Focused Methods

Automatic term extraction

Methods:

- § Using text analytics technologies, such as natural language processing (NLP) and statistical analysis
- § The system considers combinations of general frequency, frequency within a document, and multiple word co-occurrences.
- § All extracted terms are suggestions only and should be reviewed for inclusion.



Issues:

- § Term extraction, intended for auto-tagging, is usually too specific/granular for most taxonomy development. Too many terms will be extracted.
- § Candidate term extraction is most practical when the tool is integrated into taxonomy management software.
- § Useful for *enriching* a taxonomy already started from manual identification of terms.

Thesaurus

Employers (31)

Industries (144)

Job roles (5)

Locations (4)

Skills (3)

Corpora

Job skills

Candidate Concepts

cyber security

risk assessment

Blacklist

Job skills

corpus:08373cca-cd7a-4984-a1b5-b86dcfc3c579

Metadata & Statistics

Extracted Concepts

Extracted Terms

Corpus Documents

Search Terms

Search

WSI Filter

All

Search

Reset

Extracted Terms

Select All

Deselect All

Add Candidate Concepts

Add to Blacklist

Export Documents

Term	Relevance	CTS	MIS	Frequency			
Process Improvement	10.02	0	17.1	14	☒		
Digital Asset Manager	32.86	0	16.89	6	☒		
Managing Director	27.69	0	16.75				
social media	36.69	0	16.73				
video games	2.77	0	16.61				
communication skills	10.71	0	16.61	7	☒		
annual budget	6.88	0	16.61	5	☒		
Economic evaluation	3.09	0	16.25	3	☒		
Content Classification Specialist	1.35	0	16.21	4	☒		

Term extraction for candidate taxonomy concepts in PoolParty taxonomy/thesaurus management tool.

Bottom-Up Content-Focused Methods

Data ingested from spreadsheets

- § Column headers tend to be metadata types, which could become: taxonomy concepts schemes, top concepts, ontology classes, the basis for ontology relations, or attributes
- § Values in cells are metadata values, which could become: taxonomy concepts or attribute values (ontology literals).

	A	B	C	D	E
1	Company	Industry	Country	Address	Phone number
2	Volkswagen	Automotive	Germany	Dieselstrae 28, 38446 Wolfsburg, Germany	49-5361-90
3	Accenture	Information technology and services	Ireland	Unit 1, Grand Canal Square, Grand Canal Quay, Gr	353 1 407 6000
4	BASF	Chemicals	Germany	Carl-Bosch-Strasse 38, 67056 Ludwigshafen/Rheir	49 621 600
5	Tata Consultancy Services	Information technology and services	India	TCS House, Raveline Street, Mumbai 400001, Maha	91-22-6778 9999
6	Nestlé	Food and beverages	Switzerland	Av. Nesté 55, 1800 Vevey, Switzerland	41 21 924 1111
7	Royal Dutch Shell	Oil and gas	Netherlands	Carel van Bylandtlaan 16, 2596 HR The Hague, The	31 70 3779111
8	Toyota Motor	Automotive	Japan	1 Toyota-Cho, Toyota City, Aichi Prefecture, Japan	81 565 282121
9	Apple Inc	Computer hardware	United States	1 Infinite Loop Cupertino, CA 95014	1 408 996-1010
10	Samsung Electronics	Consumer electronics	South Korea	416, Maetan 3-Dong, Yeongtong-Gu, Suwon, Gyeor	82 31 2001114

Bottom-Up Content-Focused Methods

Document titles extract/download and analysis

- § Content management systems/SharePoint allow downloading of all content page titles into a spreadsheet.
- § Sort alphabetically to group some similar pages together
- § Use AI, such as Copilot with Excel files, for initial quick analysis.

Example prompt:

The attached spreadsheet contains page titles for the Intranet of an X organization. I would like to create a taxonomy of concepts to tag these pages to support retrieval. The taxonomy shall have four facets: Document Type, Audience, Activities, and Topics. An example of a term in each facet is [examples]. The Document Type and the Audience facets should be comprehensive to include all such terms mentioned, and they should be flat lists. The Topics and Activities facets should each comprise hierarchies of 2-3 levels, and they should each group multiple pages so don't need to include specific terms from all of them. Insert the generated taxonomy into the second worksheet with just one term per row...

	A	B
1	Annual Strategy and Operating Priorities	
2	Board of Directors Resource Center	
3	Climate Resilience Project Toolkit	
4	Communications Brand and Media Center	
5	Community Health Program Resources	
6	Country Office Directory and Contacts	
7	Cybersecurity Awareness and Reporting	
8	Data Protection and Privacy Resources	
9	Diversity Equity Inclusion and Belonging	
10	Donor Compliance and Reporting Guidance	
11	Education Initiatives Across Our Regions	
12	Emergency Response Coordination Hub	
13	Employee Benefits by Country	
14	Executive Leadership Team Updates	
15	Finance Policies and Procedures	
16	Fundraising Campaign Planning Toolkit	
17	Global Programs and Regional Offices	
18	Grant Management Resources and Templates	
19	Humanitarian Programs Learning Library	
20	Information Technology Help Desk	
21	Internal Job Opportunities Worldwide	
22	International Travel Request Center	
23	Learning and Professional Development	
24	Monitoring Evaluation Accountability and Learning	
25	New Employee Global Onboarding	
26	Our Mission Vision and Values	
27	Partnership Development and Due Diligence	

Bottom-Up Content Management-Focused Methods

Existing controlled vocabularies and term lists

- § Important to find and review
- § Sources may include siloed taxonomies in current or legacy systems: CMSs for documentation or marketing, DAM systems, Web CMS, SharePoint term store, metadata stored in spreadsheet files.
- § If created for a different purpose (use case), their usefulness may be limited.
- § Consider them as a source for only some terms.
Don't take all of the legacy terms



Glossaries

- § Purpose of glossaries is different from taxonomies/controlled vocabularies
- § Terms needed to be defined might not be appropriate for tagging, search, and retrieval
- § Some terms may be too specific. Useful broader terms may be missing.



Bottom-Up Content Management-Focused Methods

Requested terms from content creators/editors

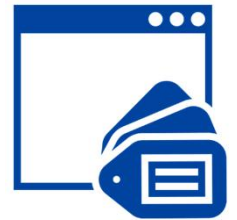
- § Usually from subject matter experts (SMEs)
- § In spreadsheets, allowing columns for narrower terms, synonyms, notes, and comments
- § Define the scope to a very specific subject domain.
- § Provide example terms to SMEs.
- § Provide instructions, guidance, and clarify specificity:
 - Ø number range of terms, such as 10-50
 - Ø need terms for tagging and search: common topics the content is about
 - Ø do not need a classification scheme, do not need terminology/glossary
- § Meet to review, discuss, and clarify suggestions
- § Taxonomist edits suggestions and sends back to SME for approval
- § Possible multiple iterations

	A	B	C	D	E	F	G
1	L1	L2	L3	L4		Comments	Synonyms
23		Transport and Logistics Services				RT customs and border management	
24		Freight Services			Key:		
25		Passenger Services			Red - Added		
26		Transport Integrators			Blue - Moved		freight forwarders, travel agencies,
27		Transport Logistics Providers			Purple - Changed		Port Services, Transport Terminals
28		Customs and Border Management			Cross-out - to Delete	Moved from under Transport Policy and Regulation	
29		Transport and Sustainable Development				Changed from Transport and Deve Sustainable Transport, Transport and Sustainability	
30		Transport and Poverty					
31		Transport-Impact-on-Trade				move under Transport Economic	
32		Transport-and-Economic-Geography				move under Transport Economic	
33		Transport and Urban Development					
34		Transport and Millennium Development Goals (MDGs)				[will probably remove and make a synonym to Transport and Sustainable Development-HH]	
35		Transport-and-Social-Responsibility					
36		Gender and Transport					
37		Transport and HIV-AIDS				RT Disability	
38		Universal Access in Transport					
39		Transport Safety				Move from under Transport Policy and Regulation	
40		Traffic and Road Safety				changed by Heather Hedden	Traffic Safety, Road Safety, Highway Safety
41		Transport Impact on the Environment				Change from Transport and the Environment; move down under Transport and Sustain	
42		Carbon Emissions and Transport					Low-Emissions Transport
43		Transport Infrastructure and Environment					
44		Transport-Information-Systems				Move under Transport Infrastructure	
45		Transport Economics				Transport Finance	
46		Congestion					
47		Fuel Taxes					
48		Road Funds					
49		Road Tolls				Toll Roads, Highway Tolls, Road Pricing, Road User Fees	
50		Transport and Economic Geography					
51		Transport Efficiency				Transport Productivity	
52		Transport Employment					

Bottom-Up Content Management-Focused Methods

Keyword tag logs

- § A very useful source for terms to add to the taxonomy
 - More useful source than search logs
- § Many content management systems support uncontrolled tagging.
- § Find out if and to what extent uncontrolled tagging has been done.
- § Methods:
 - Obtain system tagging reports for a period, such as 6-12 months.
 - Sort by numeric count in descending order to focus on high-frequency terms
 - Identify and group synonymous tags.
 - Edit tags into proper concept label format.



Top-Down Methods - Involving Users

To ensure that the taxonomy is applicable to relevant to the users' needs, their terminology, and their use cases

Methods

- § Brainstorming workshops
- § Interviews
- § Questionnaires
- § Review of AI-generated taxonomy ideas



Sample users or stakeholders should review the draft taxonomy.

Top-Down Methods - Involving Users

Gather input from:

- § Brainstorming sessions of stakeholders
 - For designating vocabularies and facets; gathering top and sample terms
- § Interviews of sample users and stakeholders
 - For scoping, developing user cases, designating vocabularies and facets, and obtaining some sample terms.
- § Requested lists of search terms from subject-matter expert end-users
 - For gathering more detailed terms



Top-Down Methods - Involving Users

Brainstorming for user input

1. Identify and name things
2. Consolidate duplicates
3. Remove outliers
4. Group them into categories or facets



Using whiteboards/sticky notes or virtual collaborate tools (e.g. Miro or Mural)

Workshops

- § Larger number (6-15), usually in a room together with a facilitator
- § Good for intranet/enterprise taxonomies with representatives from different teams
- § Often a component of a longer, taxonomy training/introduction workshops

Brainstorming sessions

- § Smaller number (3-8) in-person or remote
- § Suitable for a taxonomy for any purpose

Top-Down Methods - Involving Users

Interviews for user input

- § Interviews of sample users are for multiple purposes:
 - § To obtain use cases to better design the taxonomy and its UX
 - § To obtain use cases to use later test the taxonomy
 - § To identify taxonomy facets and scope
 - § To collect *some* terms for the taxonomy
- § From different functions that deal with the content
- § 1-2 people at once (if from the same function), in person, phone, or video call
- § Have prepared sets of questions sent to stakeholders in advance
- § Different sets of questions for information users and for information curators (uploading/tagging)
- § For information users, different question about how they:
 - § find/discover information
 - § find desired content items
- § For information curators, questions about decisions and issues for tagging/categorizing content



Top-Down Methods - Involving Users

Questions asked in Berklee Music & Health taxonomy project of sample potential users

1. If you were to search a music and health database collection of articles, what kind of research would you do?
2. What kind of topics might you look for?
3. What would be any specific keywords you might search?
4. Given the different aspects of research studies in music & health, which of the following would you most likely *initially* want to search (enter into a search box) on?
 - Specific music therapy method, musical technique, or medical music intervention
 - Who is providing the musical intervention (physician, nurse, music therapist, etc.)
 - Who is choosing the music (patient, medical provider, music therapist)
 - Music type
 - Patient population/type
 - Health issue, disease, condition
 - Medical procedure, intervention, or care type
 - Objective, outcome, or target (pain reduction, anxiety reduction, quality of life, etc.)
 - Test type (Visual analog scale, Hospital Anxiety and Depression Scale, etc.)
5. For the aspects above that you initially search on, how specific would you want to search?
6. How would you look up “patient population/type” for your purposes (if at all)? By medical condition, setting (inpatient/outpatient), demographic, etc.?
7. When searching a database of articles, are you more likely to start with basic search or got to advanced search?

	A	B	C	D	E	F	G	H	I
1	Proff	Name	Topics	Keywords	Initial search aspects	Specificity	Population	Search type	Comments
2	MT	CH	Clinical trials, music therapy methods, Neuro-rehab standardized protocols, Day training, physical therapy, occupational therapy		music therapy method	Start broad, and then get specific. Example of specific: Walking for balance	By condition, such as stroke, traumatic brain injury, cerebral palsy; Not usually by age, but possibly aging/geriatric, not pediatric Not as inpatient outpatient, but rather as acute or chronic	Basic search	Might also look up by country, such as Korea
3	MT	JK	Trauma, mind-body health, adverse child experiences ACEs music and neurobiology, crisis music and disaster, community trauma, collective trauma	GIM – guided imagery, Bonny method of GIM, analytical music therapy, PTSD, anxiety, depression, mindfulness, psychoeducation, resilience	Health issues, including psychological issues Qualitative is important than quantitative studies, so not looking at test type. Whether involving a music therapist or not. Client selected music is of interest.		By diagnosis first. Demographic might be a factor; in person, setting is third	Could go either way. Less experienced so will probably go with defaults	Consider the population as "clients," not "patients"
4	MD	IY	Palliative care, pain management, quality of life, geriatric population, pediatric population, behavior, communication, stress, symptoms, music therapy impact,	Analgesics, PTSD, coping strategies, preparation for death,	1) Patient population, 2) condition (Alzheimer's, cancer, etc.) 3) Location - Hospice, nursing home, hospital; not initially but 4) impact: pain reduction, quality of life, etc.	Would start broad and stay broad and review the articles returned	chronically ill, hospitalized, patients with disabilities, developmental disabilities	Basic Search.	I would want review articles, rather than individual studies. Needs to be evidence-based. What's the impact?
5	PhD	PR	Music and wellness; music and biomedical markers; specific aging-related issues (by diagnosis), community music, chorus/choir, artistic endeavor and aging, cultural engagement	Aging related searches; cultural engagement; social connectedness; pain; enrichment. I currently search for "cultural engagement and aging", "music and wellness", "music and stress", etc.	intervention (physician, nurse, music therapist, etc.) ; Health issue, disease, condition; Medical procedure, intervention, or care type; Objective, outcome, or target (pain reduction, anxiety reduction, quality of life, physical rehabilitation, shortened hospital stay, etc.)	Articles indexed with music THERAPY, rather than general music interventions. I'd prefer the terms to be more specific than less; I also mine citations of the specific articles.	Notes from sample users for taxonomy terms from interviews: Berklee Music & Health taxonomy		
6	MD	DB	How music can support staff sustainability, pain management, stress, anxiety, depression. Distinguishing music therapy and music medicine.	pain, burnout (professional fatigue, compassion fatigue), resilience	1) Type of patients (cancer patients, hospice patients, palliative care patients) and 2) age demographic. Music genre is less important. Genre itself could be a topic.		By condition, site of care (facility, home) and age.	Advanced search	Search should be easy and accessible to many people, not just researchers.

Top-Down Methods - Involving Users Indirectly

Search Logs

- § Usually available for website, intranets, content management systems, etc.
- § Options of date ranges. Most recent 6-12 months is usually sufficient.
- § Reveals what users look for and what they call it.



Search Log Analysis Benefits/Goals:

- § Identify patterns of types things searched for which can suggest facets or top-term categories, broader terms
 - Search terms tend to be too specific to serve as concepts in faceted taxonomy facets.
 - Concepts in a faceted taxonomy are intended to *filter/limit* search results.
 - Concepts in a faceted taxonomy should not simply mirror the search strings driving search results.
- § Identify different names for the same thing which can become alternative labels.

Top-Down Methods - Involving Users Indirectly

Creating taxonomy concepts and alternative labels based on search log search strings



- § Sort by numeric count in descending order to focus on high-use search strings.
- § Identify and group synonymous search strings.
 - Choose the higher use or preferred organizational name as the preferred label.
 - Lower count terms should be examined for synonyms.
- § Edit search strings in proper concept label format.
 - Convert verbs to nouns/noun phrases.
 - Convert singular countable nouns into plural.
- § Omit vague search strings that cannot be determined.
- § Omit highly specific search strings, if the taxonomy is intended to be faceted for *filtering* search results.
- § Use AI (such as Copilot in Excel) to assist with analysis of very large logs.

Top-Down Methods - Involving Users Indirectly

Search log search terms – international organization employee intranet

	A	B	C	D	E	F
1	Search Term	Total Unique Searches	Include?	Type/facet	Term equivalencies	Other notes
2	shif	451	y	department/service		
3	dsa	449	n			on Tools menu
4	travel	414	y	topic		
5	staff regulations	380	y	document type		
6	igds	223	n			on Tools menu
7	pardev	220	y	department		
8	home leave	193	y	topic		
9	implementation agreement	191	y	document type		
10	procurement	178	y	department		
11	normes	157	y	department?		
12	exchange rate	151	y	topic		
13	infotec	147	y	department		
14	normlex	136	y	document database		add to Tools menu?
15	teleworking	135	y	topic		
16	dcomm	132	y	department		
17	actrav	126	y	department		
18	iris	126	y	document database	same as iuc	ERP system
19	cafeteria menu	123	n			Put it on the menu or home page
20	437	121	n		Travel policies	IGDS number
21	hrd	118	y	department		
22	iuc	114	y	document database	same as iris	
23	prodoc	112	y	department		
24	education grant	110	n			on Tools menu

Conclusions

Different design methods have value.

Bottom-Up Methods

- § Ensures the taxonomy is relevant for the content.
 - An appropriate number of search results are returned upon selection of a taxonomy concept (not too few or too many).
 - Content is not missed due to lack of taxonomy concepts.



Top-Down Methods

- § Ensures the taxonomy meets the user needs and expectations.
 - Improving results, so they can do their job better, saving time
 - Improving user satisfaction
- § Engages stakeholders, keeping them committed to supporting and improving the taxonomy.



Conclusions

Building a taxonomy based on...

- § Analysis of the actual content
- § Input from users

This context results in:

- § A more suitable, customized taxonomy based on context.
- § Greater engagement, commitment, and support for the ongoing use, maintenance, and expansion of the taxonomy by its users and other stakeholders.

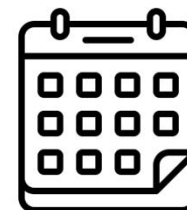
Taxonomy Resources

- § ISO 25964 Information and documentation — Thesauri and interoperability with other vocabularies. Part 1: ISO 25964-1: 2011 Thesauri for information retrieval
<https://www.iso.org/standard/53657.html>
- § ANSI/NISO Z39.19-2005 (2010) Guidelines for Construction, Format, and Management of Monolingual Controlled Vocabularies
www.niso.org/publications/ansiniso-z3919-2005-r2010
- § The Accidental Taxonomist Blog
<http://accidental-taxonomist.blogspot.com>
- § Accidental Taxonomist book websites
www.hedden-information.com/accidental-taxonomist/websites
- § Hedden Information Management past presentations
www.hedden-information.com/presentations
- § Hedden Information Management taxonomy training
www.hedden-information.com/courses-workshops
- § Taxonomy Talk, taxonomists community on Discord
<https://discord.com/invite/3qyMVYCAsw>



Upcoming Taxonomy Events

- § Bite-Sized Taxonomy Boot Camp virtual half-day event
7 October 2026 online
www.taxonomybootcamp.com/Bite-Sized
- § Taxonomy Boot Camp Conference
16 - 17 November 2026, Washington, DC, USA
www.taxonomybootcamp.com
- § Taxonomy Boot Camp London Conference
13 - 14 April 2027, London, UK
www.taxonomybootcamp.com/London/2027
(Call for speaker deadline is 9 October.)



Questions/Contact

Heather Hedden

Taxonomy/Ontology Consultant
Hedden Information Management
Carlisle, MA USA

www.hedden-information.com

accidental-taxonomist.blogspot.com

www.linkedin.com/in/hedden

heather@hedden.net



**Hedden Information
Management**

Making information findable

Hedden, Heather. (2022) *The Accidental Taxonomist*, 3rd edition.
Medford, NJ: Information Today Inc.

www.hedden-information.com/accidental-taxonomist

