

CAB Thesaurus Q1 2022 edition report

Anton Doroszenko, Thesaurus Manager
5th April 2022

Introduction

The Q1 2022 edition of CABT was released on 4th April 2022. It is a significant update compared with the previous quarterly edition. As always, the focus has been on the needs of [Plantwise](#), [Compendia](#), [VetMed Resource](#), and the [Global Health](#) and the [CAB Abstracts](#) databases.

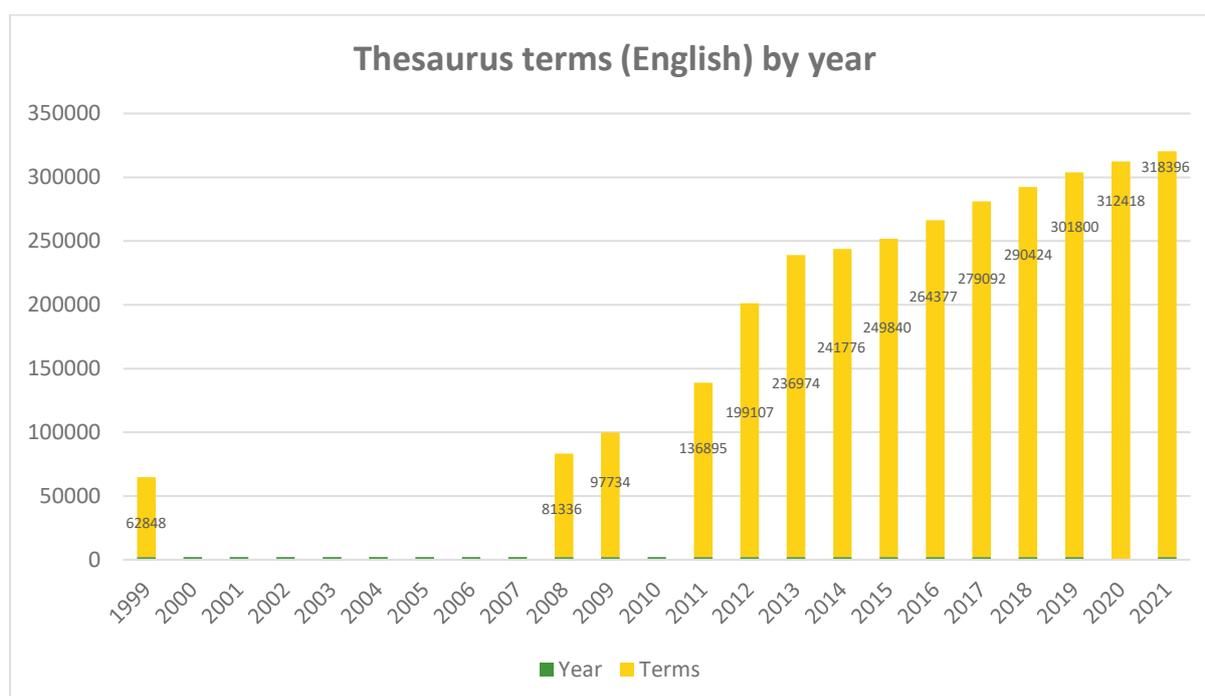
New Content and Features – some highlights

1. The number of terms/labels exceeds 3.1 million (3,145,893 to be exact).
2. In the last quarter (8th January-4th April) a total of 3718 terms were updated in some way to produce the current edition. Of these, 2179 terms, mostly organism names, were newly added in English.
3. Revision of dicot plant families continued, using the current [APG IV system](#) of classification. In Q1 2022, ten families in the Malpighiales, namely Achariaceae, Chrysobalanaceae, Clusiaceae, Erythroxylaceae, Hypericaceae, Ochnaceae, Pandaceae, Passifloraceae, Podostemaceae, and Rhizophoraceae, were revised. Additionally, Celastraceae was revised to complete the families included in Celastrales. These families include a large number of ornamentals, tropical forest trees as well as wild and cultivated fruit crops.
4. More ontological-style relationships were added. The thesaurus now includes 1372 vectors of 740 vectored agents (medical, veterinary and horticultural), 3363 biocontrol agents of 1380 hosts, and 3594 diseases caused by 2259 disease agents.
5. In Q1 2022, 413 translations from English were added, particularly into German (203), Spanish (38), Dutch (36), Portuguese (32), French (23) and Italian (19). This does not include scientific names of organisms.
6. The number of terms categorized as Miscellaneous was reduced from 2206 in Q4 2021 to 2191 currently (0.7% of total terms in English).
7. Since Q4 2021, English terms with history notes increased from 91.0% to 91.2%.

Some statistics and figures

The total number of terms/labels now stands at 3,145,893, which includes 189,271 distinct concepts (preferred terms) and 153,504 non-preferred synonyms in English. This is many times more than any other life sciences thesaurus in existence. Additionally, there are 109,110 translations of terms, excluding scientific names of organisms, from English into eleven European languages.

Growth in number of terms in English (total mid-year)



All terms by language

Language	Number in Q4 2021	Number in Q1 2022	% of UK English
Danish	270020	272080	84.1%
Dutch	292213	294288	90.9%
English - UK	321541	323595	100%
English – USA*	786	786	0.2%
Finnish	271824	273879	84.6%
French	275008	277079	85.6%
German	288018	290268	89.7%
Italian	270479	272546	84.2%
Norwegian	269836	271896	84.0%
Portuguese	293166	295246	91.2%
Russian †	2423	2438	0.8%
Spanish	296919	299005	92.4%
Swedish	270723	272787	84.3%

*Only where there is a different spelling from UK English †New in 2020 and excludes scientific names of organisms

Change in terms by technical category in English

Category	Number in Q4 2021	Number in Q1 2022	Change in terms
Scientific names	267275	269323	2048
Common names	16523	16523	0
Homographs	760	765	5
Abbreviations	1223	1224	1
Orthographic variants	1998	2041	43
Registered names	78	78	0

Change in terms by subject category in English

Category	Number in Q4 2021	Number in Q1 2022	Change in terms
Animal Breeds	1454	1454	0
Anatomical and Morphological Structures	993	993	0
Activities	226	226	0
Biogeographic Regions	19	19	0
Chemicals and Chemical Groups	8016	8015	-1
Climate Related	86	86	0
Commodities and Products	3343	3343	0
Disciplines, Occupations and Industries	945	945	0
Diseases, Disorders, and Symptoms	7182	7184	2
Geographic Entities	1554	1554	0
Habitats	80	80	0
Infrastructure	294	295	1
Institutions and Organisations	307	308	1
Miscellaneous	2206	2191	-15
Natural Processes	1144	1144	0
Organism Groups	1884	1885	1
Organism Names	285507	287555	2048
Publications	170	171	1
People Groups	500	500	0
Properties	2114	2118	4
Soil Types	371	371	0
Topographic Features	370	370	0
Techniques, Methodologies and Equipment	4411	4422	11
Time Periods	91	91	0
Vegetation Types	85	85	0

Taxonomic Ranks

29 ranks are included for scientific names of organisms. 'Unranked' is used for informal taxonomic names, such as clades. The table below lists some of the main ones.

Rank	Number of taxonomic names in Q4 2021	Number of taxonomic names in Q1 2022
Phylum	119	119
Class	204	204
Order	809	809
Family	4330	4333
Genus	31906	32124
Species	219286	221089

Thesaurus Subject Coverage – a brief outline

The main areas covered by the CAB Thesaurus are agriculture and human health. Other subject areas covered include food science, leisure, recreation and tourism.

In agriculture, the coverage includes crop production, forestry, horticulture, animal production, animal health, and aquaculture, and as well as their economic, social and environmental aspects. There is comprehensive coverage of the biology, ecology and biotechnology of plants, animals and micro-organisms of economic importance, including invasive species, pathogens, pests and parasites.

In human health and medicine, the emphasis is on human nutrition, community and public health, tropical diseases, communicable diseases, and medical mycology, entomology and parasitology.

Thesaurus classification scheme

1. GENERAL

- common terms - general processes, properties and characteristics
- named regions and countries
- research and methodology
- mathematics, statistics and computer science
- communication and information

2. PHYSICAL SCIENCES

- physics
- chemistry

3. EARTH SCIENCES

- geology
- geomorphology
- soil science
- hydrology
- meteorology and climatology

4. LIFE SCIENCES

- biology
 - microbiology
 - botany
 - zoology
 - cytology
 - embryology
 - molecular biology
 - biochemistry
 - physiology
 - genetics
 - biological structure and form
 - taxonomy
- ecology
- behaviour and psychology
- organisms

5. APPLIED SCIENCE AND TECHNOLOGY

- health and pathology
 - diseases
 - pathogenesis and epidemiology
 - health and health protection
 - pharmacology and toxicology
 - applied human and animal nutrition
 - applied genetics and breeding
 - agriculture, forestry and fishery
 - environment and natural resource management
 - technology and engineering
 - materials and equipment
 - materials handling and processing
 - transport
 - energy and power
 - food science

6. SOCIAL SCIENCES AND HUMANITIES

- social sciences
 - education
 - sociology
 - social welfare
 - policy, law and legislation
 - economics
- culture and humanities
 - leisure, recreation and tourism

Other Features of the CAB Thesaurus

Notes fields:

- Organism Name Author
- Term Definition
- History Note
- General Notes
- Scope Notes
- Source of Term
- Enzyme Commission Number

Non-standard associative relationships

- Crop Plant ↔ Harvested Product
- Disease Agent ↔ Disease Name
- Biocontrol Agent ↔ Host
- Disease Vector ↔ Vectored Agent
- Hybrid ↔ Hybrid Parent