

Activity	AnalysisInstrument	ChemicalAnalysisMethod	Diseas	EOConstituentCompound	ExtractionMethod	Location	PlantMaterialHistory	Plant	PlantPart	TargetOrganism
	THIS DICTIONARY IS UNDER DEVELOPMENT (Not Current Priority)	THIS DICTIONARY IS UNDER DEVELOPMENT	Peter to Provide			Peter to Provide Dictionary or Data Source for development.		I need to quantify the results in this one. Will do so today.	THIS DICTIONARY IS UNDER DEVELOPMENT (Gita thinning down a list now)	Peter to Provide Dictionary or Data Source for development.
				Essential Oils (EOs) are the concentrated hydrophobic liquid containing volatile chemical compounds extracted from plants. Essential oils are also known as volatile oils, ethereal oils, aetherolea, or simply as the oil of the plant from which they were extracted, such as oil of clove. Qualitative (constituent compounds) and quantitative (%) analysis of the chemical composition of the tested Essential Oils (Extracts?), with each known compound linked to its IUPAC International Chemical Identifier (InChI).						
	Description: A dictionary of 438 essential oil or constituent compound biochemical and/or biological activities, 340 of which resolved to wikidata IDs, and 336 with descriptions of 250 characters or less.	Description: A dictionary of [24] makes/models of Gas chromatography–mass spectrometry equipment used to identify different substances within a test sample — in this case, Essential Oils mentioned in the 186 test articles downloaded from PubMed.		Description: A dictionary of 2114 constituent chemical compounds extracted from Essential Oils mentioned in the 186 test articles downloaded from PubMed. Of the 2114 entries, 1010 had their names normalized and tagged with corresponding Wikidata IDs, the other 1104 remain to be resolved as no Wikidata IDs currently exist for them.	Description: A dictionary of 73 terms for Essential Oil extraction methods.		Description: A dictionary of 82 entries relating to the plant material history leading up to the extraction of Essential Oils mentioned in selected literature chosen from the 186 test articles downloaded from PubMed. The entries include key words and phrases describing: growth conditions, plant life stages, plant material selection, post-harvest treatment methods, and extracted plant material products. Of the 82 entries, 60 were resolved to WikidataIDs.	Description: A dictionary of 1678 constituent chemical compounds extracted from the Essential Oils of [XX] plants mentioned in the 186 test articles downloaded from PubMed. Of the 1678 entries, 1678 had their names normalized and tagged with corresponding Wikidata IDs, the other 112 remain to be resolved.	Description: A dictionary of [XX] part(s) of a plant from which Essential Oils — mentioned in the 186 test articles downloaded from PubMed — were extracted.	Description: A dictionary of [55] organisms mentioned [as subjects of experiment?] in the 186 test articles downloaded from PubMed.
Filename: Activity.xml	Filename: AnalysisInstrument.xml			Filename: EOConstituentCompound.xml	Filename: ExtractionMethod.xml	Filename: Location.xml	Filename: PlantMaterialHistory.xml	Filename: Plant.xml	Filename: PlantPart.xml	Filename: TargetOrganism.xml
File Location: https://github.com/petermr/CEVOpen/blob/master/dictionary/Activity/Activity.xml	File Location: https://github.com/petermr/CEVOpen/blob/master/dictionary/AnalysisInstrument/AnalysisInstrument.xml			File Location: https://github.com/petermr/CEVOpen/blob/master/dictionary/EOConstituentCompound/EOConstituentCompound.xml	File Location: https://github.com/petermr/CEVOpen/blob/master/dictionary/ExtractionMethod.xml	File Location: https://github.com/petermr/CEVOpen/blob/master/dictionary/Location/Location.xml	File Location: https://github.com/petermr/CEVOpen/blob/master/dictionary/PlantMaterialHistory/PlantMaterialHistory.xml	File Location: https://github.com/petermr/CEVOpen/blob/master/dictionary/Plant/Plant.xml	File Location: https://github.com/petermr/CEVOpen/blob/master/dictionary/PlantPart/PlantPart.xml	File Location: https://github.com/petermr/CEVOpen/blob/master/dictionary/TargetOrganism/TargetOrganism.xml
id: DAVE.activity.n where n is a serialized number	id: DAVE.activity.n where n is a serialized number			Missing	DAVEid: DAVE.activity.n where n is a serialized number		DAVEid: serialized identifiers for use in this project	id: ID correlated to the row number of this table (not including first two column heading rows)		id: CM.TargetOrganism.n where n is a serialized number
				title: type of data to be normalized and tagged with Wikidata ID.				title: type of data to be normalized and tagged with Wikidata ID.		title: type of data to be normalized and tagged with Wikidata ID. In this case, “TargetOrganism”
				name: The name is a human readable string describing the concept.						name: a human readable string describing the concept.
term: The name is a human readable string describing the concept.	term: The name is a human readable string describing the concept.			term: The term is the precise string used to identify the concept. Name and Term are often the same.	term: The name is a human readable string describing the concept.		SearchTerm: the precise string used to identify the concept to be normalized and tagged with Wikidata ID.	term: The term is the precise string used to identify the concept.		term: the precise string used to identify the concept. (Name and Term are often the same.)
					Acronym: The acronym for the term, if any.		PlantMatHistCat1:			
					Apparatus: Apparatus used to conduct the extraction method described by the term.		PlantMatHistCat2:			
wikidataid: Unique identifier linked to wikidata.org — a free and open knowledge base that can be read and edited by both humans and machines.	wikidataid: Unique identifier linked to Wikidata.org — a free and open knowledge base that can be read and edited by both humans and machines.			wikidata: Unique identifier linked to Wikidata.org — a free and open knowledge base that can be read and edited by both humans and machines.	wikidataid: Unique identifier linked to Wikidata.org — a free and open knowledge base that can be read and edited by both humans and machines.		wikidataID: Unique identifier for each normalized dictionary term, linked to Wikidata.org — a free and open knowledge base that can be read and edited by both humans and machines	wikidata: Unique identifier linked to Wikidata.org — a free and open knowledge base that can be read and edited by both humans and machines.		wikidata: Unique identifier for each normalized dictionary term, linked to Wikidata.org — a free and open knowledge base that can be read and edited by both humans and machines.
description: short description of the activity sourced from wikidata and/or wikipedia	description: short description of the activity sourced from wikidata and/or wikipedia			desc: data source	description: short description of the activity sourced from wikidata and/or wikipedia			desc: data source		description: Short description of the plant part being identified in that row
No. of source papers: 186	No. of source papers: 1			No. of source papers: 186	No. of Entries (The header is not counted): 73		No. of source papers: 186	No. of source papers: 186		No. of source papers: 186
No. of entries (Headers are not counted): 438	No. of Entries (Headers are not counted): 24			No. of Entries (Headers are not counted): 2114	No. of terms describing EO Extraction Methods resolved in Wikidata: 71		No. of Entries (Headers are not counted): 82	No. of Entries (Headers are not counted): 1678		No. of Entries (Headers are not counted): 55 (also not including 5 rows with data only in the “query” column
No. of unique activity names (including alternate spellings or synonyms): 438	No. of unique entries (including alternate spellings or synonyms): 23			No. of unique compound names (including alternate spellings or synonyms): 2114	No. of unique Wikidata IDs (including synonyms): 37		No. of unique entries (including alternate spellings or synonyms): 82	No. of unique plant names (including alternate spellings or synonyms): 1678 - DUPLICATES = ??		No. of unique entries (including alternate spellings or synonyms): 55
No. of activities resolved in wikidata (including alternate spellings or synonyms): 340				No. of Chemical Compounds resolved in Wikidata: 1010	No. of entries with no Wikidata IDs: 2		No. of unique WikidataIDs resolved to SearchTerms: 38	No. of Chemical Compounds resolved in Wikidata: ????		No. of Target Organisms resolved in Wikidata: 5/
Number of unique wikidata ids attributed to activities (normalizing for alternate spellings and synonyms): 250				No. of Chemical Compounds NOT resolved in Wikidata: 1104	No. of source articles with no Analysis Type found: 64		No. of processes NOT resolved in Wikidata: 22	No. of Chemical Compounds NOT resolved in Wikidata: 112		No. of Chemical Compounds NOT resolved in Wikidata: 3
No. of entries without wikidataid: 98										
No. of entries with descriptions: 336										
No. of entries without descriptions: 102										
	<ul style="list-style-type: none">I’m not sure if Gas chromatography–mass spectrometry equipment always go together, or if Gas chromatography can be one machine and mass spectrometry another.I believe there are duplicates yet to be normalizedPotential errors: For example, I believe the Wiley references (lines 20 and 1. are references to Spectral Databases.Once this dictionary is finalized, please update main description and the file location in this document and delete this note.Currently all entries for “desc”....<ol style="list-style-type: none">are from one source<ul style="list-style-type: none">are described as being “Hacked” from a few papers PMR 20190904 ... is that the best way to describe the method of extration from the source material?As of Jan 27,2020 this equipment list contains only 24 models. Should we need to train software by providing some makes/models, many can be found here:<ol style="list-style-type: none">https://www.thomasnet.com/nsearch.html?cov=NA&heading=96153267&typed_term=Gas+chromatography&searchterm=Gas+Chromatography+%28GC%29+Instruments&what=Gas+Chromatography±%28GC%29+Instruments&WTZO=Find+Suppliers&searchsource=suppliers						<ul style="list-style-type: none">I still need to deal with the duplicates, and once de-dupbed, will update the figures in this document description as well as the results.Duplicates (some of them more than once): 48<ul style="list-style-type: none">(See oil OilPlantCopy.xlsx for duplicates I’ve highlighted in yellow boxes.)Artemisia lobeliiArtemisia roxburghianaArtemisia caeruleascensBunium persicumCedrus libaniCitrus limonCitrus paradisiCitrus sinensisConyza canadensisCurcuma longaCymbopogon citratusCymbopogon nardusHedychium gardenianumHyssopus officinalisHelichrysrum stoechasLavandula latifoliaMelaleuca alternifoliaMentha piperitaMyrtus communisMicromeria cristataNepeta betonicifoliaNeolitsea dealbataOcimum basilicumOcimum micranthumPistacia lentiscusPourouma cecropiifoliaPsidium guajavaPsidium guineenseRosmarinus officinalisSalvia aucheriSalvia euphraticaSalvia moorcroftianaSalvia officinalisSalvia sclareaSideritis raeseriSphaerantia discolorTanacetum cadmeumTanacetum polycephalumTeucrium chamaedrysTeucrium montanumValeriana officinalisVetiveria zizanioides	<ul style="list-style-type: none">We need to normalize the headings across all Dictionaries<ul style="list-style-type: none">This is the third case where the column heading “description” means something other than “data source / method of input”CapitalizationIn this case, is the column heading “id” related to Essoil? I don’t know how to describe it here. The format is: <i>CM.TargetOrganism.n</i> where n is a serialized numberI don’t know how to describe the column headings for “Wikipedia” or “query” in this case		