Common Names / Scientific Names for Global Agricultural Concept Scheme (GACS):
Historical policies, current evidence in thesauri, discussion points and a proposal for future policy

Lori Finch, NAL
March 2, 2015

The problem

The unfortunate truth is that taxonomic names change. Common names are ambiguous. Also, there is not a single taxonomic authority or database that covers all organisms for agriculture that is up-to-date, reliable and open source. It is a huge undertaking to do this work and requires expertise and people. If there was one golden source, the partners (CABI, FAO and NAL) would not feel compelled to create a subset of organisms that are important to agriculture in their respective thesauri. Certainly, CABT has striven to fill this need with a bounty of organism terms. However, we all struggle with keeping content up-to-date with so many changes due to molecular genetics. Use a phylogenetic or taxonomic viewpoint? In an ideal future, lexicographers would simply "download" and refresh our nomenclatural data from an authoritative source, relying on the individual taxonomists to produce golden data in a standard format that is reliable, current and in alignment and agreement with cohorts from all over the globe. There are some efforts that are trying to be this golden data set, such as Catalog of Life, Encyclopedia of Life or notably, www.globalnames.org . The golden set is still elusive. The partners face the tasks of trying to keep this data fresh and current so that it is useful for indexing the current scientific literature and providing links back to previous literature that uses an older name for searchers. GACS can be a step forward to reduce the effort the partners expend to keep taxonomic names current. GACS can incorporate any advances in information sharing in this field.

GACS mapping has revealed that common names and scientific names are not treated the same in each thesaurus, and has prompted the question of modeling. CABT, who has the lion's share of organism names, uses a Subject Category for "Organism names". One will find both common group names (e.g. , vertebrates, fishes), functional names (e.g., aquatic animals) and scientific names (e.g. Lampetra) as preferred terms intermingled in this hierarchy (see Appendix 1). AGROVOC and NALT have somewhat a similar approach in that common group names and functional names are separate from scientific names (see Appendix 2 and 3).

Indexing policy and guidelines have shaped these terminologies. Because of this relationship, it is prudent to recall these policies. The historical indexing policies for common names and scientific names which were instituted at NAL came from AGRIS Indexing Guidelines. These policies are stated with examples from NALT in Appendix 4. A noteworthy principle is that common names were used in lieu of scientific names since we expect users of our information systems to use these common names. Since CABI also supports the use of common names, so it seems we are really in agreement but have taken different approaches to organization of concepts.

Suggestions on topics for discussion on common and scientific names

1. Decide organization, such as using a singular subject category called “organisms”. This is a top-down approach, but we may alter this decision after dealing with the details.
2. Decide what types of common names are allowed as prefLabel or altLabel. Additionally consider the hierarchical and associative relationships to visualize the hierarchy and word block. Here is a list of names which need consideration:
    a. Common name of species that have 1:1 relationship to a species
    b. Common name of species that has a 1:many relationship to species name
    c. Common names of group that are taxonomically-derived and have 1:1 relationship with a scientific name
    d. Common names of groups that are not taxonomically-derived and have 1:1 relationship with a scientific name.
    e. Common names of groups that have a 1:many relationship with scientific name.
    f. Common names of groups of domesticated animals
    g. Common names of laboratory animals
    h. Common name for species, ambiguous with name of product
    i. Functional common names
3. Develop a decision making tool so that we can enable others to follow our policy. “A policy is only good if it is understandable, executable and enforceable”.

See Appendix 5 that demonstrates a decision making tool with examples. This spreadsheet was created with these assumptions:

a) singular subject category of organisms
b) limited intermingling of common names and scientific names in hierarchies
c) separation of functional names
d) continued use of common names of animals as prefLabel
e) continued use of common name for laboratory animals as prefLabel

4. Decide if “technical categories” or extended SKOS can explicitly state taxonomic rank, scientific name, common name, common group name, etc. CABT and AGROVOC have experience and examples and may be desired for GACS, for example:
    a. CABT: dogs has “technical categories”: COM Common Name (Organisms).
    b. AGROVOC: Oryza sativa <produces> rice
    c. CABT: Orzya sativa <harvested product> rice

Appendix 1: Excerpt from CAB Thesaurus Subject Category “Organism Names” showing the presence of common names, functional names and scientific names. Note intermingling of common names, functional names and scientific names.

: : organisms
: eukaryotes
animals (UF Animalia)
. aquatic animals
. . aquatic invertebrates
. . . aquatic arthropods
. . . . aquatic insects
. . . . water mites
. . . freshwater invertebrates
. . . . free living nematodes
. . . . . Anatonchidae
. . . . . Anatonchus
. . . . . Anatonchus alleni
. . . . . Anatonchus bathybius

: : Chordata
: vertebrates (UF Vertebrata)
fishes (UF Pisces)
. aquarium fishes
. brackishwater fishes
. Cephalaspidomorphi
. . Petromyzontiformes
. . . Petromyzontidae
. . . . Lampetra
. . . . . Lampetra fluviatilis
. . . . lampreys (RT Lampetra, Petromyzon)
. . . . Petromyzon
. . . . . Petromyzon marinus

Appendix 2: Excerpt from AGROVOC showing the presence of common (group) names, functional names and latin names. Note “vertebrates” as well as “Vertebrata”, note “amphibians” as well as “Amphibia”. Scientific names generally separate from common names.

- animals
  - Carnivorous animals
  - Insectivorous animals
  - Invertebrates
  - Young animals
  - aquatic animals
  - fighting animals
  - hyperprolific animals
  - monogastric animals
  - noxious animals
  - useful animals
  - vertebrates
    - amphibians
    - birds
    - fishes
      - Bony fishes
      - Brackishwater fishes
      - Cartilaginous fishes
      - Cyclostomes
      - Diadromous fishes
      - Freshwater fishes
      - Saltwater fishes
    - mammals
    - reptiles
  - wild animals

- Animalia
  - Acanthocephala
  - Amphibia, Reptilia
  - Annelida
  - Arthropoda
  - Aschelminthes
  - Brachiopoda
  - Bryozoa

- Chaetognatha
- Chordata
  - Aves
  - Mammalia
  - Pisces
  - Thaliacea
  - Urochordata
  - Vertebrata
    - Amphibia
      - Frogs
      - Salamanders
      - Toads

Appendix 3: Excerpt from NALT showing the presence of common names, functional names and latin names. Scientific names are in Subject Category Taxonomic Classification of Organisms and does not contain any common names as prefLabel in hierarchy. Common (group) names and functional names are separate in an “organisms” hierarchy. Note: “Amphibia” and “amphibians” like AGROVOC.

: : : Taxonomic Classification of Organisms
: : Animalia
: Chordata
Amphibia
. Anura
. . Bufonidae
. . . Amietophrynus
. . . . Amietophrynus pantherinus
. . . . Amietophrynus pardalis
. . . Bufo
. . . . Bufo americanus
. . . . Bufo bufo
. . . . Bufo calamita
. . . . Bufo marinus
. . Dicroglossidae
. . . Fejervarya

: : : Biological Sciences
: : organisms
: animals
vertebrates
. amphibians
. . caecilians
. . frogs
. . salamanders and newts
. . toads
. birds
. . aviary birds
. . birds of prey
. . budgerigars
. . canaries
. . chickens

Appendix 4: Review of the indexing policies and use of common and scientific names for animals and plants, and their impact

ANIMALS

Historically, NAL, AGRIS and CABI had similar indexing policy for animal names:

1) For domestic animals, use the common name
    a. E.g., swine UF Sus scrofa domestica
    b. E.g., zebu UF Bos indicus
    c. E.g., chickens UF Gallus gallus domesticus
    d. E.g., bison UF Bos bison
    e. E.g., reindeer UF Rangifer tarandus tarandus
    f. E.g., dogs UF Canis familiaris
2) For wild animals, use the Scientific name
    a. E.g., Branta canadensis UF Canada goose
    b. E.g., Procyon lotor UF raccoon
    c. E.g., Cervus elaphus UF red deer
    d. E.g., Loxodonta africana UF African elephant
    e. E.g., Bos javanicus UF banteng

Since we expect people to search on the popular, common name of domesticated species, this policy helped match a searcher's terms to the indexing terms and help retrieval of the concept. However, we recognize that common name can be ambiguous (and various in different geographic regions) and scientific names are generally unambiguous.

In a thesaurus, implementing this policy can be difficult as it is not always easy to decide if a species is domesticated or not. For example, some deer or reindeer are domesticated and farmed but many are wild. Note the inconsistency in red deer and reindeer in the above examples. In the policy, zoo animals were always considered wild animals and so the scientific name was used. Also, the common name of laboratory animals was used (e.g., rats instead of Rattus).

Due to the domestic / wild animal policy and the need to keep Scientific names in their own hierarchy, there is also a divergence in NALT for the handling of nontaxonomic and taxonomic animal groups (Note the inconsistencies):

1) The "organisms" hierarchy is used for common names and animal groups
    a. E.g., germ-free animals
    b. E.g., plant parasitic nematodes
    c. E.g., root-knot nematodes (note: RT Meloidogyne)

d. E.g., cattle UF Bos taurus
e. E.g., camels UF Camelus (so Camelus is not in the Taxonomic Classification of Organisms hierarchy)
f. E.g., trout – (note: ambiguous 1:many relationship)
g. E.g., bass – (note: ambiguous 1:many relationship)
h. E.g., eel – (note: 1:1, RT Anguillidae)
i. E.g., corals – (note: 1:1, RT Anthozoa)
j. E.g., vertebrates UF Vertebrata (so Vertebrata is not in the Taxonomic Classification of Organisms hierarchy)
k. E.g., mites (note: RT Acari)
l. E.g., ticks (note: RT Acari)
m. E.g., bot flies (note: RT Oestridae)
n. E.g., snakes (note: RT Serpentes)
o. E.g., alpacas, (note: RT Lama)
p. E.g., llamas, (note: RT Lama)
q. E.g., subterranean termites (note: RT Rhinotermitidae, Termitidae)
r. E.g., gophers (note: RT Geomyidae)

2) The "taxonomic classification of organisms" hierarchy is used for the taxonomic animal groups, and common names are treated as cross references when there is 1:1 relationship.
    a. E.g., Anthozoa UF sea anemones
    b. E.g, Nematoda UF nematodes
    c. E.g., Echinoidea UF sand dollars
    d. E.g., Tunicata UF urochordates
    e. E.g., Aves RT birds (but "birds" is in organism hierarchy)
    f. E.g., Isoptera UF termites (see "subterranean termites" above)
    g. E.g., Coleoptera UF beetles
    h. E.g., Araneae UF spiders

When an individual species has a common name that is unique to that species, it has been used as UF for the common name:

a. E.g., Leptinotarsa decemlineata UF Colorado potato beetle
b. E.g., Dolichovespula arenaria UF aerial yellowjacket

c. E.g., Oreamnos americanus UF mountain goat

There are some exceptions, such as when they are domesticated….which only adds to the confusion:

a. E.g., Bactrian camels UF Camelus bactrianus
b. E.g., dromedaries UF Camelus dromedariu

PLANTS

Historically, NAL, CABI and AGRIS had the same policy for plant names:

1) Use the Scientific name when it is preharvest, e.g. “The corn was treated with alachlor” would be assigned “Zea mays” as we knew that the “corn” in the sentence here is the corn plant in the field being treated with herbicide. This example is easy to know this the corn plant but it is not always this apparent unless the reader has expert knowledge.
2) Use the product name when it is postharvest, e.g. “The corn was treated with fungicides to prevent spoilage” as we knew this was the “corn grain” that was harvested that was being treated.

The same was used for fruits:

1) Use the Scientific name when it is preharvest, e.g., “The apple canopy was sprayed with insecticides” would be assigned “Malus domesticus” as we knew the “apple” in the sentence was the tree growing in the orchard.
2) Use the product name when it is postharvest, e.g., “The apples were stored in modified atmospheres to prevent ripening” as we know the apples were the fruit that was harvested.

Historically, specific common names of plants were always nondescriptors for the Scientific name. We do not have the same problem with plants that we have with animals. There are almost no examples of the common group or product name being used for the Latin name in NALT. (Only example I found is vascular plants UF Tracheophyta)

Examples:

1) Acer rubrum UF red maple
2) Xanthocyparis nootkatensis UF Alaska yellow cedar
3) Pieris japonica UF Japanese pieris

As with animals, the Scientific names are in a separate hierarchy from the products and are usually connected by RT relationship.

Examples:

1) apples RT Malus domestica
2) persimmons RT Diospyros kaki, RT Diospyros virginiana
3) alfalfa RT Medicago sativa
4) wheat RT Triticum
5) quinoa grain RT Chenopodium quinoa
6) cilantro RT Coriandrum sativum
7) coriander seed RT Coriandrum sativum

However, there is ambiguity with terms where the product and common name of the plant are the same:

Examples:

1) corn
2) cilantro
3) ginger
4) pepper – very ambiguous
5) rice

Groups of plants, nontaxonomic, are covered mostly in the “organisms” hierarchy:

1) alpine plants
2) ground cover plants
3) crops
4) scions

Groups of plants, taxonomic in nature but may be seen as “form or habit types”, have their common name listed in the “organisms” hierarchy with RT to Scientific name:

1) grasses RT Poaceae
2) ferns and fern allies RT Equisetopsida, Isoetopsida….
3) Mosses and liverworts RT Marchantiopsida, Bryopsida
4) Conifers (note: no relationship to any taxonomic name in NALT)

But there are exceptions:

1) Magnoliophyta UF angiosperms
2) Coniferophyta UF gymnosperm

Appendix 5: Decision-making spreadsheet for determining prefLabel vs. altLabel of organisms names, examples, and the resulting equivalence, hierarchical and/or associative relationships.

Name	Condition	Example	altLabel	prefLabel	Equivalence result	Hierarchical result	Associative result
Common name of a species	Has 1:1 relationship to species name	Colorado potato beetle / Leptinotarsa decemlineata	Common name is altLabel	Species name is prefLabel	Leptinotarsa decemlineata UF Colorado potato beetle	Leptinotarsa . Leptinotarsa decemlineata	N/A
		Bactrian camel / Camelus bactrianus	Common name is altLabel	Species name is prefLabel	Camelus bactrianus UF Bactrian camel	Camelus . Camelus bactrianus	N/A
		Northern red oak / Quercus rubra	Common name is altLabel	Species name is prefLabel	Quercus rubra UF Northern red oak	Quercus . Quercus rubra	N/A
	Has 1:many relationship to species name	Bluebells / Campanula / Barleria / Mertensia/ Eustoma / Hyacinthoides / etc.	Common name is altLabel and must be disambiguated in some manner	Species name is prefLabel	Mertensia virginica UF Virginia bluebells ; Hyacinthoides non-scripta UF English bluebell	Mertensia . Mertensia virginica Hyacinthoides . Hyacinthoides non-scripta	N/A

Name	Condition	Example	altLabel	prefLabel	Equivalence result	Hierarchical result	Associative result
Common name of a laboratory animal / domesticated species / wild animal	Has 1:1 relationship to species or taxonomic name	dogs / Canis lupus	Scientific name is altLabel	Common name is prefLabel	dogs UF Canis lupus familiaris UF Canis familiaris	Canis . Canis lupus . dogs	dogs RT laboratory animals
		tigers / Panthera tigris	Scientific name is altLabel	Common name is prefLabel	tigers UF Panthera tigris	Panthera . tigers	N/A
		swine / Sus scrofa domestica	Latin name is altLabel	Common is prefLabel	Swine UF Sus scrofa domestica	Sus scrofa . swine	farmed animal species RT swine
Common name of laboratory animal / domesticated species / wild animal	Has 1:many relationship to taxonomic names	Rats / Rattus / Rhizomys / Sigmodon / Dipodomys / etc.	Common and scientific names are prefLabel	Common name is prefLabel	rats	Rodentia . rats . mice . Muridae . Heteromyidae . Bathyergidae laboratory animals . rats	Rattus RT rats Dipodomys RT rats
		Rabbits / Oryctolagus / Sylvilagus / Nesolagus / etc.	Common and scientific names are prefLabel	Common name is prefLabel	rabbits	Leporidae . rabbits . Oryctolagus . Sylvilagus . Pentalagus	Oryctolagus RT rabbits
		Elephants / Loxodonta / Elephas	Common and scientific names are prefLabel	Common name is prefLabel	elephants	Elephantidae . elephants . Loxodonta . Elephas	Elephants RT Loxodonta RT Elephas

Name	Condition	Example	altLabel	prefLabel	Equivalence result	Hierarchical result	Associative result
Common name of a group, derived from taxonomic name **	Has 1:1 relationship to taxonomic group	vertebrates / Vertebrata	Taxonomic name is altLabel	Common name is prefLabel	vertebrates UF Vertebrata	vertebrates . birds . . Anseriformes . mammals . . ruminants	N/A
		cervids / Cervidae	Taxonomic name is altLabel	Common name is prefLabel	cervids UF Cervidae	animals . chordates . . vertebrates . . . mammals . . . . cervids . . . . . Alces . . . . . Cervus	N/A
		nematodes / Nematoda	Taxonomic name is altLabel	Common name is prefLabel	Nematodes UF Nematoda	invertebrates . nematodes . . Chromadoridae	N/A
Common name of a group, derived from taxonomic name **	Has 1:many relationship to taxonomic group	<No examples found>					

Name	Condition	Example	altLabel	prefLabel	Equivalence result	Hierarchical result	Associative result
Common name of a group, name NOT derived from taxonomic name / popular	Has 1:1 relationship to taxonomic group	birds / Aves	Taxonomic name is altLabel	Common name is prefLabel	birds UF Aves	N/A	N/A
		termites / Isoptera	Taxonomic name is altLabel	Common name is prefLabel	termites UF Isoptera	Termites . Hodotermitidae . Kalotermitidae	
		Ticks / Ixodida	Taxonomic name is altLabel	Common name is prefLabel	ticks UF Ixodida	Acari . ticks . . Argsidae . . Ixodidae	
	Has relationship to many taxonomic groups (no 1:1 relationship)	Mites / Astigmata / Oribatida / Mesostigmata / Prostigmata		Common name is prefLabel	N/A	Acari NT mites NT Astigmata NT ticks NT Mesostigmata NT Prostigmata NT Oribatida	mites RT Astigmata RT Oribatida RT Mesostigmata RT Prostigmata
		trout /Oncorhynchus / Salmo / Salvelinus		Common name is prefLabel	N/A	Salmonidae NT trout NT Salmo NT Oncorhynchus NT Salvelinus NT Coregonus	trout RT Salmo RT Oncorhynchus RT Salvelinus

Name	Condition	Example	altLabel	prefLabel	Equivalence result	Hierarchical result	Associative result
Common name for species, ambiguous with name of product	Has 1:1 relationship to species name	rice* / Oryza sativa	Disambiguate common name / product name	Species name AND product name are prefLabel	Oryza sativa UF rice plant Rice UF rice grain <is a> product	rice BT products Oryza sativa BT Oryza	Rice RT Oryza sativa
		Coriander* / cilantro* / Coriandrum sativum	Disambiguate common name / product name	Species name AND products are prefLabel	Coriandrum sativum UF coriander plant UF cilantro plant	cilantro BT herbs coriander seed BT spices Coriandrum sativum BT Coriandrum	Coriandrum sativum RT cilantro RT coriander seed
	Has 1:many relationship to species name	Pepper* / Capsicum annuum / Piper nigrum	Disambiguate common name / product name	Species names AND products are prefLabel	Piper nigrum UF pepper (Piper) Capsicum annuum UF pepper (Capsicum)	black pepper BT spices sweet peppers BT vegetables Piper nigrum BT Piper Capsicum annuum BT Capsicum	Piper nigrum RT black pepper Capsicum annuum RT sweet peppers

*Note: need better semantics to help disambiguate common name / product

**Note: NAL prefers to use the taxonomic name for this section, but uses common name here for recommendation since it may be more useful friendly and has 1:1 relationship with taxonomic name. CABT uses common name as prefLabel sometimes and scientific name as prefLabel other times. Is there a policy helping to decide when to do one or the other at CABI? Examples in CABT: prefLabel Nematoda (not nematodes), prefLabel invertebrates (not Invertebrata), prefLabel arthropods (not Arthropoda), prefLabel Arachnida (not arachnids), prefLabel Amphibia (not amphibians). Is there a level at which taxonomic name is prefLabel? Could we use rank to set policy?

Name	Condition	Example	altLabel	prefLabel	Equivalence result	Hierarchical result	Associative result
Common name is functional name**	Has some relationship to a taxonomic name	Aquatic animals / animals / Animalia	N/A	Functional name is prefLabel, taxonomic name follows previous rules	N/A	Organisms . <functional names> . . aquatic organisms . . . aquatic animals* . . . aquatic plants* . animals . . aquatic animals . . vertebrates	N/A
		Benthic plants / plants / Plantae	N/A	Functional name is prefLabel	N/A	Organisms . <functional names> . . aquatic organisms . . . aquatic plants . . . . benthic plants	N/A
		Annuals / plants / Plantae	N/A	Functional name is preLabel	N/A	Organisms . plants . . annuals . . bedding plants . . benthic plants . . perennials	
	Has no relationship to taxonomic name	Examples: autotrophs, hosts, microorganisms, parasites, symbionts, etc.	N/A	Functional name is prefLabel	N/A	Organisms . <functional names> . . autotrophs . . hosts . . microorganisms . . parasites	

*Note: Prefer to keep functional names separated from scientific names hierarchy. However, aquatic animals ARE animals, so we may have to allow these as in hierarchy for animals but keep separated with a node label or other mechanism.

**Note: CABT uses Subject Category: OG Organism Groups for many functional names.