This is a reproduction of a library book that was digitized by Google as part of an ongoing effort to preserve the information in books and make it universally accessible.

Google™ books

https://books.google.com

Z
5071
U58Btg
1991

Digitized by Google

CORNELL UNIVERSITY LIBRARY
3 1924 059 848 014

Digitized by Google

United States
Department of
Agriculture

National
Agricultural
Library

Technical
Services
Division

AGRICOLA – Guide To Subject Indexing

December 1990

Digitized by Google

Digitized by Google

United States
Department of
Agriculture

National
Agricultural
Library

Technical
Services
Division

December 1990

AGRICOLA – Guide To Subject Indexing

By Martha W. Hood
Indexing Branch

LIBRARY
AT
CORNELL UNIVERS

National Agricultural Library
Beltsville, Maryland
December 1990

Digitized by Google

Z
5071
U58Btg
1991

National Agricultural Library Cataloging Record:

Hood, Martha W.

AGRICOLA—guide to subject indexing.

1. AGRICOLA (Information retrieval system)—Abstracting and indexing. I. National Agricultural Library (U.S.) II. Title.

aS494.5.I47

Digitized by Google

INTRODUCTION

The Indexing Branch of the National Agricultural Library (NAL) and its cooperators currently index over 2,000 journal titles and 500 monographs per year for AGRICOLA (AGRICultural OnLine Access), the library's bibliographic database. AGRICOLA contains records for agricultural literature citations of journal articles, monographs, theses, patents, computer software, audiovisual materials, and technical reports. Indexing records comprise approximately 85 percent of the AGRICOLA database, the remaining 15 percent being cataloging records. In addition to the online database, AGRICOLA records are used to produce the printed *Bibliography of Agriculture,* CD-ROM versions of the database, and various bibliographic series and other products of NAL. Since NAL is the U.S. input center for the decentralized agricultural database of the Food and Agriculture Organization of the United Nations, NAL indexing and cataloging records also appear in AGRIS online and the printed *AGRINDEX.*

This guide outlines the principles to which NAL indexers adhere in subject indexing. It supersedes all sections of the Indexing Manual issued in March of 1985 which dealt with subject indexing. (Sections of the March 1985 Indexing Manual dealing with descriptive indexing are still in force.) The guide sometimes refers to, and is to be used in conjunction with, the continuing series *National Agricultural Library Notes to Indexers.*

Comments and questions about this guide are welcome and should be directed to:

Martha W. Hood
National Agricultural Library
Indexing Branch, Room 011
10301 Baltimore Boulevard
Beltsville, MD 20705

Digitized by Google

Digitized by Google

BASIC PROCESSES IN SUBJECT INDEXING FOR AGRICOLA

1. Identifying concepts which represent the subject and purpose of a document.

2. Deciding which of these concepts are important for retrieval of this document by a user of AGRICOLA.

3. Expressing concepts needed for retrieval in the indexing languages used in AGRICOLA:
   3.1. the classification scheme (AGRICOLA Subject Category Codes, MARC tag 072), and
   3.2. the controlled vocabulary (*CAB Thesaurus*, MARC tags 650 and 651).

4. Using uncontrolled vocabulary (identifiers, MARC tag 653) for concepts not represented or represented insufficiently specifically in the *CAB Thesaurus*.

Each of these processes will be dealt with in detail in the pages that follow.

Digitized by Google

1. IDENTIFYING CONCEPTS WHICH REPRESENT THE SUBJECT AND PURPOSE OF A DOCUMENT.

Follow these steps for every document to be indexed! You may occasionally get by with less, but over the long haul your indexing will surely reflect whether you take this systematic approach vs. a scattershot approach which quickly gleans a few terms from the document's title.

1.1. Understand the **TITLE.**

1.2. Read the **INTRODUCTION** to the document for background information and to locate the **author's statement as to the purpose of the article.** This statement of purpose is often the last sentence of the introduction. (But see 2.15.) Compare the statement of purpose with the title of the document to see that they correlate.

1.3. Read the **ABSTRACT** as an aid for focusing your examination of the text. (You may want to reexamine the abstract at a later stage in your indexing to verify that items in the abstract are actually substantively discussed within the text.)

1.4. Scan the **MATERIALS AND METHODS** section for specific names of organisms, geographic locations, procedures, etc. Information in the Materials and Methods section is often more specific than elsewhere in the document.

1.5. Note **ILLUSTRATIONS, CHARTS, GRAPHS,** and **TABLES** and their legends or captions. Also note headings, phrases, or words in **BOLDFACE** or **OTHER SPECIAL TYPEFACES.**

1.6. Scan the **RESULTS.** Don't overlook significant negative results.

1.7. Read the **CONCLUSIONS** and/or **SUMMARY** and note actual findings, as differentiated from speculations or implications. It may be helpful to compare the findings to the statement of purpose in the **INTRODUCTION** to see to what extent the purpose was realized.

1.8. Scan the **BIBLIOGRAPHIC REFERENCES** for hints and confirmation if, at this point, you are still having problems understanding the article or if you are looking for some specific item of information (e.g., the meaning of an acronym).

1.9. Scan any supplied **KEYWORDS** and/or **RUNNING TITLES** to see that you have considered all concepts indicated there. However, use these with care; they are not always well chosen. Don't allow them to overrule your own good judgment.

1.10. As a final aid in determining slant, consider the following:
a) the type of journal in which the article appears,
b) the organization with which the author is affiliated, and
c) the funding source for the research.

Digitized by Google

2. DECIDING WHICH OF THESE CONCEPTS ARE IMPORTANT FOR RETRIEVAL OF THIS DOCUMENT BY A USER OF AGRICOLA.

In making these decisions, consider both **EXHAUSTIVITY** and **SPECIFICITY.**

EXHAUSTIVITY: You should index as many concepts as you deem necessary to adequately represent the document in all its significant aspects.

SPECIFICITY: You should index these concepts at a level of detail consistent with the intended audience of the document and the treatment of the subject in the document. Of course, the specificity of descriptors available in the thesaurus is also a consideration.

It is vitally important to keep in mind that we are talking here of CONCEPTS, not WORDS. Always remember that the same concept can be expressed by many different words or combinations of words.

For a longer discussion of indexing exhaustivity and specificity, consult *Indexing Policy Memorandum No. 1*, July 17, 1989.

2.1. Give a high priority to concepts the author considers important as evidenced by the manner and frequency of their treatment in the document. (See 1.1 - 1.9.)

2.2. However, also index concepts that you know to be important, on the basis of your own experience and judgment, even if they are not emphasized by the author.

2.3. Index only information which warrants the time and expense of retrieval, i.e., index concepts which are substantively discussed, not just mentioned in passing.

2.4. Index methods, techniques, and procedures only if they are one of the main topics of the document. The appearance of the name of a method, technique, or procedure in the title of an article is not sufficient justification for indexing it. There should be some textual discussion to the effect that the method is new, or modified, or applied differently, etc., for it to be indexed.

2.5. Index significant negative results. E.g., an article reporting a study of antibiotic residues in beef should be indexed with descriptors for ANTIBIOTIC RESIDUES, BEEF, etc., and a Category Code Q202, even if no residues are detected.

2.6. Many concepts are composed of more than one factor, or component. Index each concept completely, i.e., for all its factors. E.g., the concept "spleen weight" must be indexed with descriptors for both SPLEEN and WEIGHT.

2.7. Generally, do not index redundantly, i.e., do not index the same concept at more than one level of specificity. E.g., do not index a document on bark beetles as pests with both BARK BEETLES and INSECT PESTS.

Digitized by Google

2.8. However, do index the same concept both broadly and narrowly if it is so treated in the article. E.g., a document may discuss analytical techniques in general, but concentrate upon gas chromatography in specific detail; index with both ANALYTICAL METHODS and GAS CHROMATOGRAPHY.

2.9. Give serious consideration to reflecting the level of specificity in the document in your indexing. E.g., index camel milk with the descriptor CAMEL MILK, not MILK; index dairy effluents with the descriptor DAIRY EFFLUENTS, not EFFLUENTS.

2.10. However, be aware that to do so may not always be appropriate or necessary. E.g., a Materials and Methods section might identify the cattle used in an experiment as Holsteins, but whether to index the breed concept depends upon whether this aspect of the study is emphasized in the document.

2.11. There is no limit to the number of concepts which can be indexed for a particular document, and exhaustivity is often correlated with indexing a large number of concepts per document. However, this is not necessarily true. The interplay of exhaustivity and specificity determines how many concepts are indexed. E.g., under some circumstances, it may be desirable to index a document covering many aspects of a subject with just one broad, general concept; such a document is exhaustively indexed at the proper level of specificity with only one concept.

2.12. In indexing, treat equally concepts which are treated equally in the document. E.g., a document discusses five specific parameters of blood chemistry, giving approximately equal weight to all five; either index all five individually or index to a broader concept that encompasses all five (BLOOD CHEMISTRY), but do not index just two or three of the five. If, however, two of the five are emphasized significantly more than the other three, it is appropriate to index those two specifically, and also use BLOOD CHEMISTRY. (See 2.8.) For further discussion of accurate representation of concept emphasis, consult *Indexing Policy Memorandum No. 2*, July 17, 1989.

2.13. Generally, do not index concepts which are present only through implication or speculation. E.g., a document on metabolism of omega-3 fatty acids should not be indexed with the concept DISEASE PREVENTION, even though an understanding of omega-3 fatty acid metabolism may ultimately aid in reducing the incidence of cardiovascular disease.

2.14. However, sometimes an implied concept must be indexed to provide the proper context for the other concepts. E.g., a document discussing pathogenicity of a parasite for a pest insect should be indexed with the concept BIOLOGICAL CONTROL even if the author does not use this term.

Digitized by Google

2.15. Documents usually contain references in the text to related work of other authors. In particular, the all-important **INTRODUCTION** may contain a brief literature review as a background to or a rationale for the present study. Do not be misled into indexing concepts that represent work not conducted as part of the research described in the document at hand.

Digitized by Google

3. EXPRESSING CONCEPTS NEEDED FOR RETRIEVAL IN THE TWO INDEXING LANGUAGES USED IN AGRICOLA: THE CLASSIFICATION SCHEME AND THE CONTROLLED VOCABULARY.

The classification scheme used for indexing AGRICOLA is the *AGRICOLA Subject Category Codes with Scope Notes (Modified AGRIS)* and the controlled vocabulary is the *CAB Thesaurus*.

These two indexing languages are independent of one another in the sense that each is a distinct system used to represent the subject content of a document. The Category Codes assigned to a document should represent that document as completely as possible and without any necessity to refer to the descriptors assigned, and vice-versa.

Naturally, this is not to say that there will not be relationships between the codes and the descriptors, but to point out that they are used separately, not in combination with one another, to represent the document. Indeed, there may well be what appears at first blush to be redundancy in assignment of codes vis-a-vis descriptors and vice-versa.

Descriptors representing the same concepts represented by the assigned Category Code(s) should always be assigned, though most often the descriptors will be more specific than the codes to which they relate. To put this another way, a user should be able to comprehend why each Category Code was assigned by examining the assigned descriptors.

The reverse, however, is not necessarily true; descriptors may be used to bring out aspects of the article which are not represented by the codes because these aspects are not treated in sufficient depth or with sufficient emphasis to merit assignment of a code.

In expressing concepts in the indexing languages, aim for as much CONSISTENCY as possible in applying Category Codes and descriptors to represent concepts. Use the same code(s) and the same descriptor(s) to represent the same concept each time you index that concept. Our indexing will be valuable to AGRICOLA users only insofar as it is CONSISTENT.

Digitized by Google

3.1. AGRICOLA SUBJECT CATEGORY CODES (MARC TAG 072)

NAL, as the U.S. input center, is a major contributor to AGRIS, a decentralized agricultural database maintained by FAO, the Food and Agriculture Organization of the United Nations. It is therefore important that AGRICOLA and AGRIS Category Codes be compatible. AGRIS codes were found to be insufficiently specific for indexing AGRICOLA and the scheme was expanded, but in a manner which assured easy conversion of specific AGRICOLA codes into broader AGRIS codes.

In the AGRICOLA Category Code scheme, the field of agriculture is divided into 21 broad subject areas. Each of these areas is further subdivided, but even the subdivisions are very broad. Each of the subdivisions carries an alphanumeric notation consisting of one letter (which remains the same within one of the 21 broad subject areas) and three numbers. Each code also has a 'title,' known as a 'section heading' (e.g., notation: J800 = section heading: Soil Conservation).

There are some hierarchical relationships within the scheme, e.g.:

L830 Animal Diseases – General
L831 Animal Diseases – Fungal
L832 Animal Diseases – Bacterial
L833 Animal Diseases – Viral

The scope of each code is described by fairly lengthy notes, perhaps 100 words or more. Included among these Scope Notes may be cross-references indicating other codes which might be appropriate instead of, or in addition to, the code at hand.

There is an alphabetical index to the section headings and scope notes.

The Category Codes are used for:

1. searching the database for broad subject areas, and
2. arranging entries in the print version of the AGRICOLA database, the *Bibliography of Agriculture.*

3.1.1. Assign up to three (3) Category Codes per document.

3.1.2. Assign the codes in descending order of importance; use the first Category Code to represent the major topic discussed. (Remember that this code determines where the document appears in the *Bibliography of Agriculture.*)

3.1.3. Use relator codes (K001, L001, L002, L003, M001) to narrow the focus of large categories.

3.1.3.1. Never place relator codes in first position.

3.1.3.2. Never use relator codes without assigning at least one other code.

Digitized by Google

3.1.4. Assign Category Codes on the basis of the subject of the work actually conducted in the study at hand; do not assign them on the basis of what the subject can be used for.

3.1.5. Follow the instructions in Scope Notes that sometimes specify the order in which certain codes should be applied.

3.1.6. **Never** assign codes on the basis of the alphabetical index; always consult the Scope Note itself before assigning a code.

3.1.7. Emphasize pests and diseases over the normal processes of plants and animals.

3.1.8. Be alert to problems in, or the need for changes to, the Scope Notes. Write a brief, informal memo (on 8 1/2 x 11 inch paper) to the Category Code Coordinator detailing the problem. Keep a copy of this note in your Category Code notebook. It will be useful when you are asked to comment on new versions of the Scope Notes.

Digitized by Google

3.2. CAB THESAURUS (MARC TAGS 650 – SUBJECT DESCRIPTORS – AND 651 – GEOGRAPHIC DESCRIPTORS)

NAL uses a controlled vocabulary, the *CAB Thesaurus*, which is compiled and maintained by CAB International in cooperation with NAL. This thesaurus is also used to index *CAB Abstracts*. It includes approximately 56,000 terms/phrases, of which approximately 47,400 are DESCRIPTORS, or preferred terms, and 8,600 are NON-DESCRIPTORS, or lead-in vocabulary. These terms are arranged in a single alphabetical (word-by-word) sequence.

Although there is a large number of precoordinated descriptors in the thesaurus, the basic approach is that of post-coordination, where concepts are combined (coordinated) at the retrieval stage rather than at the indexing stage.

Each descriptor is accompanied by a WORD BLOCK which shows the descriptor's relationship with other descriptors in the thesaurus. There are three types of relationship represented in the word blocks:

HIERARCHICAL RELATIONSHIPS

BT Broader Term
NT Narrower Term

The **BT/NT** relationship is basically a generic/specific or part/whole relationship, although the *CAB Thesaurus* does contain hierarchies based upon relationships other than these two.

Hierarchies may contain up to seven (7) levels and all levels are displayed with each descriptor.

ASSOCIATIVE RELATIONSHIPS

rt related term

This is used to represent relationships other than hierarchical. The **'rt'** relationship is approximately equivalent to a 'see also' relationship.

PREFERENTIAL RELATIONSHIPS

uf used for
USE use

Digitized by Google

The **'uf'** and **USE** relationships are reciprocals of one another employed to relate synonyms, quasi-synonyms, and terms which stand in various other relationships to one another. The term following a **'uf'** designation is a non-descriptor, a non-preferred term. The term following a **USE** designation is a descriptor, a preferred term. The **USE** relationship is approximately equivalent to a 'see' relationship. **USE** instructions are **not** optional, and must be followed.

Two other elements may be present in a word block:

- a SCOPE NOTE, which may define or limit or give instructions for the use of a descriptor; and
- a HISTORY NOTE, which indicates usage of the descriptor over time, e.g., when it was added to the thesaurus, what descriptor(s) was used to index the concept before it was added, etc.

3.2.1. If the term to which you first refer is a descriptor, always examine its complete word block before assigning it to a document.

3.2.2. Choose the most specific suitable descriptor in the word block.

3.2.3. If the most suitable specific descriptor in the word block is any other than the descriptor to which you first referred, examine the word block of the descriptor which you judge to be most suitable. **The importance of this step cannot be overstated.**

3.2.4. If the term to which you first refer is a non-descriptor, proceed to the word block for the preferred descriptor and examine it completely. Never assign a descriptor on the basis of a **USE** reference without consulting the word block for the descriptor. **Again, the importance of this step cannot be overstated.**

3.2.4.1. You may discover that the term to which you are being referred is in the wrong hierarchical context for the document at hand.

3.2.4.2. You may discover that a narrower term of the term to which you are being referred is more appropriate/specific.

3.2.5. Do not index the same concept at two different levels of the hierarchy, i.e., do not assign both a BT and one of its NT's to represent the same concept. (See 2.7; but see also 2.8 and *Notes to Indexers,* No. 9, November 1986).

3.2.6. Do not assign a descriptor to a concept that is inconsistent with the context of the descriptor's hierarchy. **This is a critically important principle in indexing.** If an indexer does not follow this instruction, s/he runs the risk of seriously misleading a user and negating the value of using a controlled vocabulary. We should be able to guarantee our users that we apply descriptors as consistently as possible. See: *Notes to Indexers,* No. 7, June 1986.

Digitized by Google

3.2.7. It will often be necessary to assign more than one descriptor to index a single concept. E.g., both COMPUTER TECHNIQUES and MAPPING are needed to represent the concept of "computer mapping."

3.2.7.1. On the other hand, do not use two descriptors to represent a concept when a precoordinated descriptor for the concept exists. E.g., for "amino acid biosynthesis," use AMINO ACID METABOLISM, not AMINO ACIDS and BIOSYNTHESIS.

3.2.8. If you don't readily find a descriptor or combination of descriptors which represents a concept, be persistent.

3.2.8.1. One good source is the KWOC index of the thesaurus.

3.2.8.2. Another possibility is to search the AGRICOLA database on CD-ROM. Search for the concept free-text (i.e., as it occurs in titles and abstracts) and look to see what descriptor(s) other indexers have used to represent it.

3.2.9. Presently, there is only one rule concerning the order in which descriptors are arranged. (See 3.2.10.) Many indexers do consider concept importance, word proximity, and other factors in arranging descriptors, but such considerations are not mandatory.

3.2.10. The first descriptor should represent the major thrust of the material being indexed and should belong to one of the following groups if at all possible:

3.2.10.1. names of animals.
3.2.10.2. names of plants.
3.2.10.3. names of foods.
3.2.10.4. names of feeds.
3.2.10.5. names of products.
3.2.10.6. names of international organizations.

See also *Notes to Indexers*, No. 8, October 1986.

3.2.11. Use geographic descriptors liberally. Consider their application for any of the following situations or subjects:

3.2.11.1. Geographical term in the title.
3.2.11.2. Map in the article.
3.2.11.3. Survey carried out.
3.2.11.4. Field research carried out.
3.2.11.5. Flora or fauna of a region studied.
3.2.11.6. New taxa or other taxonomic information are a topic.
3.2.11.7. Organizations or institutions are a significant topic.
3.2.11.8. The article is biographical in slant.
3.2.11.9. The article is or contains a case report or a case study.
3.2.11.10. Economics.
3.2.11.11. History.
3.2.11.12. Legislation, jurisprudence, or politics.

(continued next page)

Digitized by Google

3.2.11.13. Education/training.
3.2.11.14. Sociology.
3.2.11.15. Land and/or water resources.
3.2.11.16. Ecology, including pollution.
3.2.11.17. Meteorology/climatology.
3.2.11.18. Epidemiology.
3.2.11.19. Human ecology.

3.2.12. As with subject descriptors, apply the most specific appropriate geographic descriptor, e.g., NOVA SCOTIA not CANADA, MIZORAM not INDIA.

3.2.13. Do not apply geographic descriptors for experimental research **unless** the work involves field tests. If the work is conducted in a laboratory or a greenhouse, use of a geographic descriptor is usually not appropriate.

3.2.14. The geographic origin of a strain or line may sometimes be significant in a way that warrants use of a geographic descriptor even if the work is done in vitro. Note, however, that this does not apply to geographic words appearing in the nomenclature of strains of bacteria and other organisms.

3.2.15. For articles in which subjects (animals, plants, humans) originating from one country/area are studied in a country to which they have migrated or been imported, do not apply a geographic descriptor for the country of origin. Apply the geographic descriptor for the country in which the subjects were studied; indicate their geographic or ethnic origin in some other way. E.g., an article about food habits of Cubans living in Long Island would take NEW YORK and ETHNIC GROUPS as descriptors. If CUBAN-AMERICANS is not in the title, it could be used as an identifier. (See section 4.) However, this practice obviously should be applied only in the case of living subjects studied in a natural environment. E.g., an article about insect specimens collected in Ghana and studied in a laboratory at Harvard University would take GHANA, not MASSACHUSETTS, as its geographic descriptor.

3.2.16. For plants and animals, the *CAB Thesaurus* may contain descriptors for the scientific (Latin) name, the common name, or both, for the same organism. In applying these descriptors to agriculturally related plants and animals, follow these general guidelines:
3.2.16.1. Use the scientific name for plants through harvesting.
3.2.16.2. Use the common name for plants as commodities.
3.2.16.3. Use the common name for domestic animals.
3.2.16.4. Use the scientific name for pests and pathogens.
3.2.16.5. Use the scientific name for wildlife.
3.2.16.6. Certain articles may require both scientific and common names. Such articles may include, but are not limited to, those which emphasize the genetics/breeding of plants and animals, especially with respect to

Digitized by Google

genetic effects on the products of these plants and animals. E.g., an article discussing the fiber characteristics of cotton produced by these two different species of Gossypium and their hybrids would require COTTON; GOSSYPIUM HIRSUTUM; and GOSSYPIUM BARBADENSE.

3.2.16.7. For scientific names, use the greatest level of specificity available in the thesaurus in field 650. This might be order, family, genus, or genus and species. If necessary, reflect a greater level of specificity in field 653. (See 4.5.)

3.2.16.8. In cases where the use of a scientific name is mandated by these guidelines but is not available in the document at hand, it is the responsibility of the indexer to seek it out. Often a quick way to do so is to search the AGRICOLA database on CD-ROM. However, common names are highly variable and often very localized. Therefore, if an indexer cannot trace the scientific name at any level of the taxonomic hierarchy with some degree of certainty within a reasonable period of time, it may be necessary to use a term such as MEDICINAL PLANTS, RUBBER PLANTS, WEEDS, INSECT PESTS, etc., in field 650.

3.2.17. Be alert to the possible need to propose descriptors for addition to the thesaurus. Guidelines and forms for proposing terms are available as an Appendix to this Guide (Appendix B).

Digitized by Google

4. USING UNCONTROLLED VOCABULARY (IDENTIFIERS, MARC TAG 653) FOR CONCEPTS NOT REPRESENTED OR REPRESENTED INSUFFICIENTLY SPECIFICALLY IN THE CAB THESAURUS.

4.1. Identifiers are uncontrolled subject terms not found in the *CAB Thesaurus* either as descriptors or non-descriptors.

4.2. Identifiers are not validated at data entry against any authority list.

4.3. Use identifiers for significant local geographic concepts. E.g., add MILWAUKEE, WISCONSIN, to field 653 if the city itself is emphasized. (WISCONSIN will of course be entered in field 651.)

4.4. Use identifiers for clarification of some abbreviations, acronyms, initialisms, etc., in document titles. E.g., UTM in a title is explained by adding UTERINE MILK in field 653.

4.5. Use identifiers for taxonomic names. E.g., the species SENECIO ARVENSIS is not in the thesaurus and so must be entered in field 653. (SENECIO, which is in the thesaurus, will represent this species in field 650.)

4.6. Use identifiers for proper names not in the thesaurus. E.g., DISASTER ASSISTANCE ACT OF 1988; ALABAMA HOME ECONOMICS ASSOCIATION; COALITION FOR PEACEFUL SCHOOLS.

4.7. Do not repeat exactly title words or phrases as identifiers. However, if you think it may be useful, you may enter in field 653 some variant of a title word or phrase. E.g., for the title "Negatively taxed farm expansion investments," it was decided that NEGATIVE TAXES was a valuable entry in field 653.

4.8. Do not use as identifiers terms which appear in the *CAB Thesaurus* as descriptors even if the thesaurus descriptor represents a concept other than the concept associated with the term in the document at hand. E.g., the descriptor GRANULATION carries a **BT1** PLANT DISORDERS. It is not permissible to enter GRANULATION in field 653 for a document that discusses granulation in some other context, such as fertilizer technology.

4.9. Do not use as identifiers terms which appear in the *CAB Thesaurus* as non-descriptors. It is mandatory to follow the **USE** instruction in the thesaurus instead. E.g., the descriptor instructs: "sunlight **USE** solar radiation." It is not permissible to enter SOLAR RADIATION in field 650 and SUNLIGHT in field 653.

4.10. Always use some descriptor from the ***CAB Thesaurus*** which represents each concept being used as an identifier, even if the descriptor is very broad. E.g., when you add in field 653 EVOLSONIA TEXANA, a fossil plant that is not represented in the thesaurus even at the family level, add PLANTS in field 650.

Digitized by Google

4.11. Do not use long phrases as identifiers; identifiers should be in the same general format as descriptors. E.g., EFFECT OF ROTOTILLING ON GARDEN YIELDS is not an appropriate identifier.

4.12. Use identifiers liberally, but never as a short-cut to avoid searching the thesaurus for an appropriate descriptor.

4.13. Identifiers that you find yourself using frequently may be good candidates for proposal as new descriptors. (See 3.2.16.)

Digitized by Google

REFERENCES

AGRIS – Guide to Indexing. 23 (Rev.1) Rome: FAO-AGRIS, December 1982.

MEDLARS Indexing Manual. Part II. Bethesda, Md.: National Library of Medicine, 1981.

British Standard Recommendations for Examining Documents, Determining Their Subjects and Selecting Indexing Terms. BS6529:1984. London: British Standards Institution, 1984.

Digitized by Google

APPENDIX A

A CHECKLIST OF INDEXING ERRORS

It may be helpful to use this checklist to evaluate your own indexing. It gives indexers a good indication of the kinds of errors for which reviewers check during quality review in the Indexing Branch.

1. Failing to index a significant concept.

2. Indexing of a concept with an inappropriate term.
   2.1. Assigning term which is too broad.
   2.2. Assigning term which is too specific.
   2.3. Assigning term which is in the wrong hierarchical context.
   2.4. Assigning term which is inconsistent with scope note.

3. Indexing of a concept that should be ignored.
   3.1. Insufficient information on concept to warrant retrieval effort.
   3.2. Represents work not conducted as part of the research described in the document at hand.

4. Indexing of the same concept at more than one hierarchical level.

5. Failing to index all components/facets of a concept.

6. Failing to index significant negative results.

7. Failing to index with equal weight concepts which are treated with equal weight in the document.

8. Indexing for implications or speculations.

9. Indexing routine techniques, methods, or procedures.

10. Failing to index significant geographic concept.

11. Assigning geographic term which is insufficiently specific.

12. Assigning geographic term for laboratory/greenhouse research.

13. Failing to assign category codes in proper order.
    13.1. Failing to assign Category Codes in descending order of importance.
    13.2. Failing to assign Category Codes in order mandated by scope notes.

14. Using relator Category Codes in first position, or alone.

Digitized by Google

15. Failing to assign some descriptor to represent each concept which was assigned a Category Code.

16. Failing to assign some descriptor to represent each concept which was assigned an identifier.

17. First descriptor does not represent major thrust of article or belong to an acceptable group. (See 3.2.10 for acceptable groups.)

18. Failing to define title acronyms/abbreviations/initialism in either the descriptor or the identifier fields.

19. Assigning an identifier that repeats a term in the document title.

20. Assigning a thesaurus non-descriptor as an identifier.

21. Assigning an identifier inappropriate in format and/or length.

Digitized by Google

APPENDIX B

GUIDELINES FOR COMPLETING "PROPOSED CHANGES IN *CAB THESAURUS*" FORMS

1. Please type all submissions.

2. The tone of any text in the "Justification" and "Additional comments" sections of these forms should be neutral and business-like.

3. Follow conventions of capitalization and indentation used in the thesaurus itself, i.e., capitalize descriptors, use lower case for non-descriptors, etc.

4. Use a separate form for each proposed change.

"Present descriptor" section of form:

5. In this section of the form, enter any existing CAB descriptor(s) related to your proposed change together with page number(s).
   - 5.1. It is very important that you search diligently for such descriptors.
   - 5.2. Consider spelling variants, synonyms, and quasi-synonyms before proposing a new descriptor.
   - 5.3. Consider also that a combination of existing CAB descriptors may express a concept adequately.

6. If you are proposing a new **USE** reference to an existing CAB descriptor, enter that existing descriptor and its page number in this section of the form.

"Proposed descriptor" section of form:

7. When proposing a descriptor, observe the following:
   - 7.1. Descriptors should always be nouns, noun-phrases, or gerunds.
   - 7.2. Noun-phrases should be written to exclude prepositions unless doing so violates accepted English usage.
   - 7.3. Adjectives are used only with nouns, e.g., RED LIGHT. CABI is adamantly opposed to the use of adjectives standing alone as descriptors.
   - 7.4. Singular vs. plural: Use the singular form for processes, properties (attributes), and unique things. Use the plural for classes of things.
   - 7.5. Use natural word order. If deemed necessary, propose the inverted form as a non-descriptor.
   - 7.6. Avoid punctuation, including hyphens, if at all possible.
   - 7.7. Use scientific rather than lay terminology. The lay term may be proposed as a non-descriptor.

Digitized by Google

8. When considering proposal of a new descriptor, always ask yourself these two questions:
   8.1. Will this descriptor be used often enough to justify its addition to the thesaurus?
   8.2. Does this descriptor represent a concept which can be consistently and reliably differentiated from any other concept?

Do not propose the descriptor unless you can answer "yes" to both of these questions.

"Justification for change" section of form:

9. Authorities and references cited in support of proposed changes need not comply with any standard format, but the citation should be complete enough bibliographically that there will be no difficulty identifying the source, e.g., name of book, author or editor, place of publication, publisher, date, page number(s).

10. The first source to consult when justifying a proposal is *AGROVOC*, FAO's multilingual thesaurus of agricultural terminology. CABI, NAL, and FAO are all interested in compatibility between the *CAB Thesaurus* and *AGROVOC*; CABI will thus prefer an *AGROVOC* term in most cases over those from other sources. Always indicate on your form any related terms you find in *AGROVOC*.

11. For justification, consider other thesauri; specialized dictionaries, handbooks, monographs, and textbooks; and numbers of postings for the proposed concept in relevant databases. Generally, do not use all-purpose dictionaries.

"Hierarchy and rt" section of form:

12. You must supply at least one **BT1** (Broader Term) for each proposed descriptor (except in the rare instances when you may be proposing a new top term). This suggested **BT1** must be a valid descriptor in the *CAB Thesaurus* or a term you are proposing concurrently.

13. The relationship between a term and its **BT** should, in most cases, be a generic one, i.e., the broader term should denote a class of concepts of which the narrower term is **always** a member. The part-whole relationship is also acceptable in some cases, e.g., for descriptors representing geographic locations, those representing anatomical parts of an organism, those representing disciplines. The thesaurus presently contains hierarchies based on relationships other than these two types, but we wish to avoid adding any more.

14. Provision of **NT**'s (Narrower Terms) and **rt**'s (related terms) is optional. If you choose to do so, these should also be either valid descriptors in the thesaurus or terms you are proposing concurrently.

Digitized by Google

"Additional comments" section of form:

15. A term is assumed to include any of its common definitions. If certain definitions are to be excluded from the scope of the descriptor or if uncommon definitions are to apply, this should be made clear. The two primary methods of clarification would be hierarchical placement or a scope note. Please supply the scope note if one is needed.

16. If the Indexing Branch coordinator, CABI, or users will encounter any difficulty in finding a definition for the term, please supply one in this section of the form. Indicate the source for the definition if the source does not appear in the "Justification" section of the form.

Digitized by Google

Digitized by Google

PROPOSED CHANGES IN CAB THESAURUS

___New descriptor
___New non-descriptor (i.e., USE reference)
___Delete descriptor
___Delete non-descriptor
___Add rt (related term)
___Delete rt (related term)
___Change in hierarchy
___Problems with existing USE and *uf* references
___Spelling variance/error
___Substitution
___Word meaning/semantic variance
___Other (Please specify)_____

Please type. _____

Present descriptor, if any (CAB p._____).

Proposed descriptor or proposed USE reference.

Justification for change: Authorities, reference where proposed term was used, number of postings in relevant databases, etc.

Hierarchy and rt: Supply a BT1 either already existing in CAB or being proposed at this time (mandatory). Supply NT's and rt's either already existing in CAB or being proposed at this time (optional).

Additional comments relevant to the proposed change: Definition, scope note, etc.

Submitter's name, affiliation, and phone number.

* U.S. Government Printing Office : 1991 - 518-954/03309
Digitized by Google

Digitized by Google

Digitized by Google

Digitized by Google

Digitized by Google

Digitized by Google