

# ***Regional groupings for development information***

---

***Second General Meeting of the  
Informal Study Group on Exchange of Development Information  
Paris, October 4-6, 1993***

---

***Prepared by Ron Davies, Consultant***

503-11769

## 1 Introduction

Regional groupings are the groups of countries or geographic areas that organizations use in the description of their development activities. In the course of the use of the Common Exchange Format for Development Activity Information (CEFDA), particularly for the collection and distribution of development information on the DAI CD-ROM, problems have been identified concerning the use of these regional groupings. These problems arise principally from the fact that most development agencies use their own specific system of regional groupings which do not correspond exactly to the regions recommended in the CEFDA format. The inconsistencies between locally-defined regional groupings and the standard CEFDA groupings affects the way in which development information is used: relevant information may frequently be missed in the course of a search or attempts to find all information related to a region may retrieve a large amount of irrelevant information. In addition, users have a difficult time learning or adapting to the groupings recommended by the CEFDA format. This paper proposes changes to the CEFDA format and to the DAI CD-ROM that will improve access to development information using regional groupings.

## 2 Current problems and practice

### 2.1 Survey of practices of development agencies

In order to get more information about the current practices of development agencies concerning regional groupings, a survey of some seventy different organizations was sent out in March 1993. Complete results were received back from close to 40% of those organizations in time to be included in this paper. A variety of development agencies was represented: the breakdown was over 40% UN or UN-affiliated agencies, 35% bilateral agencies and 20% non-governmental organizations (NGOs) or other international organizations.

The most important conclusions concerning regional groupings resulting from this survey are:

- **With a few exceptions, organizations develop their own regional groupings to meet their own needs.** The organizations that follow some external standard in designating regions consist principally of a few UN agencies (which use UN Statistical Office groupings) and a few bilateral agencies (which use OECD groupings).
- **Regional groupings are not always global in scope.** In some organizations, regional groupings are not intended to be worldwide, and only certain countries or regions are included. For example, UN agencies may exclude non-member states, and non-governmental organizations (NGOs) with no activities or partners in a country may not include that country in its regional groupings.
- **Half of the organizations base their regional groupings on factors other than geographic proximity alone.** Other factors that are used in developing regional classifications are, in approximate order of importance: cultural/linguistic factors (e.g. *Arab world*); political groupings (e.g. *ASEAN countries*); economic groupings (e.g. *Least developed countries*); and administrative structures or

convenience of the development agency (e.g. Canada considered to be a country in the *European* region).

- **Regional classification schemes used by organizations rarely change.** Change, when it does come, is nearly always to accommodate changes in the political structure or designations of countries. (The only significant exception is institutional re-structuring or re-organization.)
- **Names are used more commonly to designate regions than codes.** More than three-quarters of the agencies responding indicated they used names almost exclusively to designate regions.
- **Regional classifications frequently allow a country to belong to more than one region.** One-third of the respondents indicated that regions may overlap in their geographic classification scheme. For example, Tunisia may be a North African country, an Arab country and a Francophone country. Not surprisingly, the organizations that assign a country to more than one region are the organizations that use cultural, linguistic, political or economic factors as well as geographic proximity in determining regions; they may also be organizations that draw on a thesaurus for all or part of their regional designations.
- **Development activities are sometimes or frequently described in terms of a regional grouping alone.** Two-thirds of the respondents indicated that a development activity of their agency sometimes or frequently took place at the level of a region only, and was not described in terms of countries. Retrieval of information on these activities using geographic criteria will depend upon the regional designation provided.

While many of the survey respondents did not mention problems with regional groupings on the DAI CD-ROM, several comments were noted during the survey and during informal evaluations of the first edition of the DAI CD-ROM. Requests were made to include additional regional groupings, or to use only locally-defined regional codes in the CEFDA format. It was also pointed out that the standard CEFDA regions were not well known to searchers, and that confusion was possible when a locally-defined region with a name similar to a CEFDA standard region (e.g. *West Africa*) was in fact composed of different countries.

## **2.2 The current CEFDA format**

CEFDA: Common Exchange Format for Development Activity Information<sup>1</sup> (or more simply the CEFDA format) is the recommended format for exchange of development information between two agencies as well as for contributions to the DAI CD-ROM. In both its initial and revised editions, the CEFDA format defines a single field to contain information about countries or regions that are the object of the development activity (*060 Country/region*). In this field, countries and regions are represented by codes: the codes for countries are the three-letter codes taken from the ISO 3166 standard<sup>2</sup>; the codes for regions are based on the United Nations Statistical Office (UNSO) groupings<sup>3</sup>. The *060 Country/region* field is an essential field, but descriptors which describe geographic areas may also appear in fields reserved for Macrothesaurus or local descriptors (*140 OECD Macrothesaurus descriptors* and *150 Local descriptors*).

The organization submitting data in the CEFDA format is responsible for translating locally-defined regional groupings into the equivalent CEFDA codes.

There are no clear, explicit guidelines as to how local region groupings are to be translated into the CEFDA codes; practices may vary from one organization to another. How much inconsistency this introduces into databases of CEFDA records is difficult to estimate, since the source data is only directly available to the originating agency. However these problems of translation have undoubtedly led to some difficulty in retrieving information relevant to a particular region.

### ***2.3 Current practice in the DAI CD-ROM***

The first edition of the DAI CD-ROM used a single index in which country and region names corresponding to the CEFDA-recommended codes were entered. In the second edition of the DAI CD-ROM, the Guided Search option distinguished between country searches and region searches. A country search located only activities that were focused on a particular country; a region search retrieved activities that were described in terms of that region or on any individual country in that region. This provided for an easy, inclusive region search, but did not allow for a search restricted to activities at a regional level.

## ***3 Toward an improved system of regional description***

Undoubtedly it is much easier to draw up a list of ideal characteristics for a regional system than it is to come up with a useful and practical grouping scheme. Nevertheless, such a list may help in evaluating alternatives for a regional grouping system. The ideal regional grouping scheme for the exchange of development information should be simple to understand and easy to use. It must suit bilateral exchange of development activity information as well as contributions to a common database like the DAI CD-ROM. It should lend itself to table-driven conversion routines, both for an organization exporting data in CEFDA format, and for an organization importing data into a local database or into the DAI CD-ROM. Geographic descriptions in the CEFDA format must allow the end user to distinguish between activities that take place at a country level (even when more than one country is involved) and activities that take place at a regional level (even when only a few countries are included in the region). It should use language-independent codes where possible. In searching, it should result in little loss of recall (i.e. retrieving all relevant activities) or precision (i.e. retrieving only relevant activities). It should also allow records to be sorted in a logical and useful order.

In trying to develop regional grouping schemes that would meet these criteria, a number of possible approaches to regional descriptions were considered and rejected. Because these approaches were mentioned in responses to the regional groupings survey, or in comments made concerning the first edition of the DAI CD-ROM, a brief digression is warranted in order to explain why these options were rejected.

One option suggested was to define a separate field for "region" in the CEFDA format, and have reporting organizations translate their local regional designations in each record into a list of codes of all the countries that belong to that region. (This option is more fully explained in the accompanying sidebar entitled "Translating locally-defined regions".) Organizations using CEFDA data would also translate their local regional groupings into lists of countries. During a search for a locally-defined region, the retrieval software would look for a match between

any country in the searcher's list of countries for a region and any country in the new "region" field in the CEFDA record. This option would allow for a great deal of flexibility in the way in which both the originating agency and the searcher defined regional groupings. However it would also result in the retrieval of a large amount of irrelevant material in many cases, and it would not be possible to sort CEFDA records reported in this format in a useful geographic sequence. For these reasons this option was rejected.

### ***Translating locally-defined regions***

An agency with a project in a locally-defined region *Southern Asia* (which corresponds to the UNSO-based regional definition) would include 9 occurrences of a new field *061 Region* in the CEFDA format record, as follows:

AFG BGD BTN IND IRN MDV NPL PAK LKA

An agency using a database such as the DAI CD-ROM for searching purposes, would create a locally-defined region such as *Middle East* which might include countries in North Africa and Western Asia and would be translated into the following codes:

AFG ALG BHR EGY IRN IRQ ISR JOR KWT LBN LBY MAR OMN QAT SAU  
SDN SYR TUN TUR ARE YEM

When a searcher was trying to find regional activities corresponding to the locally-defined region *Middle East*, the term *Middle East* would be expanded to the list of countries above, and records with any of those countries in the new *061 Region* field would be retrieved. However because locally-defined regions of development agencies overlap extensively, a search based on this country matching will retrieve information from many diverse areas. For example, a *Middle East* search as described above would retrieve development activities which were described by the originating agency with terms including *Africa, Arab countries, Asia, Asia and the Pacific, Middle East, North Africa and the Middle East, South Asia, West Asia*, as well as interregional or global projects.

In some cases, an exhaustive search like this would be very appropriate, but in other cases, this search would result in an overload of information of very little interest. Using this decomposition of regions into countries, however, the searcher has no control over how specific or general an individual search should be; the results of the search are always the same. It is also impossible to sort regional activities coded in this way into any useful or logical arrangement. For these reasons, the use of this approach is not recommended.

Consideration was also given to a second option (more fully explained in the sidebar "Amalgamating local regional groupings"), that would also see the elimination of standard groupings altogether. Instead reporting organizations would use only their own local regional groupings when describing their development activities in the CEFDA format. In order to provide some consistency and some assistance in searching a combined database like the DAI CD-ROM, these local regional groupings could be qualified with the name of the organization originating the term and then amalgamated into a single geographic file of terms. This geo-

### ***Amalgamating local regional groupings***

Using this approach, each organization exchanging information in the CEFDA format would use its own locally-defined regional groupings in creating CEFDA records. In addition, they would supply, with the CEFDA records, a file listing each locally-defined regional grouping and the countries or regions that it included in the grouping. In order to distinguish between different locally-defined regions, an organization responsible for a joint database, such as the DAI CD-ROM, would add the acronym of the organization as a qualifier to the regional grouping. These qualified regional groupings would then be amalgamated into a single geographic file organized along hierarchical lines, like a thesaurus.

For example, in a database containing information from a number of agencies, such as the DAI CD-ROM, many organizations would have a term *Africa*, which might be qualified as follows, where AAA, BBB and so on are the acronyms of the organizations contributing data:

Africa (AAA)  
Africa (BBB)  
Africa (CCC)  
Africa (DDD)  
Africa (EEE)

Each of these regional groupings would have as "narrower terms" in the hierarchy the specific countries or subregions that made up that region. One of the above terms might show up as follows in a hierarchical display:

Africa (AAA)  
NT Angola  
NT Benin  
NT Botswana  
NT Burkina Faso  
NT Burundi  
NT Cameroon  
NT Cap Verde  
etc.

Another might appear as:

Africa (CCC)  
NT North Africa (CCC)  
NT Africa South of Sahara (CCC)

Some of these different *Africa* regions might in fact cover identical areas. Some might be the same except for a few countries included in one, but not the other, definition. And some of these *Africa* regions might in fact be identical with other regional groupings such as *Africa South of Sahara (CCC)*. Each would be consistent only with the other regions and countries in its own system.

In the amalgamated geographic structure, a single country would belong to a number of different regional groupings. The following shows a thesaurus-like display for the country *Malaysia*, when the classification schemes of eight different organizations (here represented by the acronyms AAA through HHH) are included:

Malaysia

- BT ASEAN countries (AAA)
- BT Asia (BBB)
- BT Asia (CCC)
- BT Asia (DDD)
- BT Asia and the Pacific (EEE)
- BT East Asia and the Pacific (FFF)
- BT East/South-east Asia and Oceania (GGG)
- BT Far East (HHH)
- BT Southeast Asia (AAA)

The chief advantage of this amalgamated structure would be its flexibility. Contributing agencies would not have to worry about translating codes into standard codes, so there would be no loss of accuracy in the description. If the searcher could select for searching related terms from a thesaurus-like display (as with the DAI CD-ROM Macrothesaurus descriptors), then the searcher could easily expand a search to include different regions, according to how inclusive or how specific a search was needed. For example, using a search command much like the selection of related Macrothesaurus terms in the DAI CD-ROM software, the user interested in regional projects affecting Malaysia could easily select for searching all the different regions to which that country belonged. Similarly the user that was interested in all activities that took place in a certain geographic area could expand a region search to include all the countries in that region. In each case, the searcher has the flexibility to decide, for each individual search, how comprehensive or how specific the search should be.

While this approach works well for information contributed by only a few agencies, as the number of contributors grows, the structure rapidly becomes very complex and cumbersome. The example above is based on the locally-defined regions of only eight different organizations; any satisfactory solution for the CEFDA format and the DAI CD-ROM should be flexible enough to allow for direct contributions from dozens, if not hundreds, of organizations. In addition, some agencies have detailed and elaborate geographical thesauri which alone would add dozens of additional groupings and subgroupings to this amalgamated structure. Using locally-defined regions also does not solve the problem of sorting records by country or region, since activities from two agencies with different names for similar geographic regions may appear widely separated in any sort sequence. For these reasons, this approach is not recommended.

graphic file could be organized along hierarchical lines like a thesaurus, in order to make it easy to expand a search as required. A searcher could identify a region of interest and then have his or her search expanded to include all the countries of that region; or a searcher could start with a country and expand the search to include some or all of the differently defined regions to which the country might belong.

While this option would be practical when a small number of organizations contributed data to a joint database, direct data contributions by a moderate or large number of agencies to a database like the DAI CD-ROM would make this structure extremely complex and cumbersome. Even with a limited number of contributing organizations, some agencies also have extremely detailed geographic lists which would threaten to overwhelm this amalgamated structure. Without any system of standard groupings, sorting records into any useful sequence would be extremely difficult. For these reasons, this option was rejected.

## 4 Proposals for change

This paper proposes improvements to the description and use of development activity information that are less radical than the two approaches mentioned above. Through changes both to the CEFDA format and to the way in which information is distributed on the DAI CD-ROM, the use of regional groupings will be improved, though not all of the current problems will be resolved. Based on the survey results, the identification of common problems in the use of development information, and a brief survey of some literature on the problems of geographic retrieval of information, four different proposals were developed. These proposals are :

- Extension of the CEFDA standard for regional groupings to include frequently used non-UNSO groupings, and to re-define groupings where required to allow for better matching.
- Revision of the CEFDA format to clarify principles to be used in description of regional activities and in translation of locally-defined regional groups into standard CEFDA groupings.
- Inclusion in retrieval systems for development information, such as the DAI CD-ROM, of a hierarchical, thesaurus-like structure for countries and regional groupings that would allow searchers to expand their search strategies as required for each individual search.
- Inclusion in retrieval systems of a function to allow searchers to define commonly used search terms (such as locally-defined geographic groupings) that will be expanded at search time to include combinations of standard country and regions.

Each of these proposals will be discussed in more detail, with a brief example to indicate how the change would be reflected in practice.

***Proposal 1      The UNSO-based regional codes in the CEFDA standard should be revised to include new regions that will provide a better match for local regional groupings and to expand the current definition of UNSO-based regions, where it is warranted by the numbers of development activities.***

The CEFDA standard regional groupings should be expanded beyond the current set of UNSO-based regional groupings to include new groupings that would provide a better match for locally-defined regional designations for a significant number of development activity descriptions. The new regional groupings would include new geographic groupings, as well as groupings based on cultural/linguistic and political factors. In addition, current regional definitions should be adjusted, where appropriate, to allow for broader, more inclusive definitions, even if this means that some countries belong to several geographically-based regions.

Revisions to the current standard CEFDA groupings would fall into one of three categories:

- **New terms to allow for cultural/linguistic and political groupings**

In many cases development agencies have combined broad geographic areas in different ways. The addition of new "super-regions" helps to reduce the difficulties agencies have in describing regional activities without significantly changing the structure for activities which fit well within the current structure. For example, several agencies use a regional grouping *Latin America*. This region corresponds to the UNSO-based regions *Central America* and *South America*. A new regional grouping *Latin America* would assist in accurate description of development activities at this regional level. Also, a number of agencies have a category for the area of Africa that falls south of the Sahara. In some cases this regional grouping is called *Africa South of the Sahara*, but in many other cases, this regional grouping is simply called *Africa* by the agency, even though it does not correspond to the UNSO-based region *Africa*, which includes North African countries. In order to provide accurate description of activities in this region, as well as avoid inaccurate translation of locally-defined regions into CEFDA standard regions, a new region *Africa South of Sahara* should be defined to include the UNSO-based regions of *Eastern Africa*, *Western Africa*, *Middle Africa* and *Southern Africa*.

- **New terms to allow for cultural/linguistic and political groupings**

Many agencies use regions based on factors other than geographic proximity. For example, several agencies have a region for *Arab countries*, for *Francophone Africa*, or for *ACP countries*. The CEFDA standard should be expanded to include cultural, linguistic or political regions where a significant number of development activities would be described locally using one of these regional designations.

- **Expanded definitions of existing regions to include problem or borderline countries**

Problem or borderline countries are countries which are often included in different regions by different agencies, because they lie on the border between two regions, or because they have geographic, cultural or linguistic affinities with countries in more than one region. In the current CEFDA standard, countries belong to one and only one region (with the exception of countries in the former Soviet Union). Local regional definitions sometimes fail to match the CEFDA regional standards because they are different only by one or two countries. By expanding regions slightly, even at the cost of overlap, it will be possible for more agencies to use a CEFDA regional grouping which will correspond, by and large, to the locally-defined region.

For example, Sudan falls in the UNSO-based region of *Northern Africa*, though a number of development agencies include it in *Africa South of the Sahara*. In order that Sudan be included in searches for countries in the region *Africa South of Sahara*, Sudan should also be considered to be an *Eastern African* country. Other examples of "borderline" countries are Cyprus, which is in the UNSO-based region *Western Asia* but is considered by many agencies to be a *European* country, and Mexico, which should be both in the region *Central America* and the region *North America*.

The development and the on-going management of the expanded CEFDA standard regional codes would become the responsibility of the Steering Committee of the

Informal Group on Exchange of Development Information, with the technical support of the Coordinating Unit. The Steering Committee could perform a role similar to other agencies that manage a standard for shared information, making revisions to the standard as required. For example, an agency with problems in translating or searching using certain regional groups could apply to the Steering Committee proposing specific amendments to the CEFDA geographic authority file. Changes would be permitted where the Steering Committee felt it was warranted by a significant number of development activities. The Steering Committee could pursue other standards-related initiatives, such as training focal points in the use of the geographical authority list, or helping new contributors to establish correspondences between local regions and the CEFDA standard groupings.

***Proposal 2*** *The CEFDA format should be revised to provide more guidance in the use of regional groupings in the description of regional development activities, in the selection of standard groupings corresponding to locally-defined regions, and in the use of other fields to provide additional information about regional groupings.*

Currently the CEFDA format is somewhat ambiguous in terms of how an activity at the regional level should be described. The CEFDA format should be revised to indicate clearly that a project at a regional level should be described with a regional designation or designations, and examples (particularly *Example 4* in the section for field *060 Country/region*) should be revised to reflect this fact.

The CEFDA format should also be revised to provide guidelines to agencies that have difficulty in translating their locally-defined regions to the standard regions of the CEFDA format. With the expansion of CEFDA regional groupings to include more of the regional groupings used by many agencies, and to include new regional groupings, there will be fewer cases where the locally-defined regional grouping will not match a CEFDA standard grouping; however these changes will not eliminate the problem altogether. The CEFDA standard currently does not clearly indicate what an agency should do when faced with this kind of difficulty. There are at least three different possible approaches which an agency could adopt, based on what we might call the *approximation*, *analytic* and *promotion* principles. (These principles are explained in more detail in the sidebar "Three approaches to translating local regions".)

While each of these three approaches has disadvantages, the *promotion* approach is preferred because it is simple and can be consistently applied. While it may mean that in some cases, a relevant development activity is missed in a search because a search did not include a higher level region, the possibility of this occurring can be minimized by clearly recommending the principle of promoting local regional designations in the CEFDA standard and in the DAI CD-ROM documentation. Documenting this practice will help searchers of activity databases to be aware that they may need to include one or more levels of broader regional groupings if they hope to find all information relevant to a particular country.

Finally, the CEFDA format should encourage development agencies to use the field *160 Notes* for a qualification where the use of local groupings does not match the regional groupings of the CEFDA standard or to use field *150 Local descriptors* to include the name of the locally-defined region. The use of these qualifiers will help users to quickly identify activities that may not be relevant, or (depending on the capabilities of the retrieval software) to eliminate groups of irrelevant records from a search strategy.

### **Three approaches to translating local regions**

#### **Approximation approach**

Using the approximation approach, a development agency would approximate use of the standard CEFDA regions, by translating a locally-defined regional designation that did not have an exact CEFDA equivalent into the "closest" CEFDA regional code, even if the CEFDA region included countries not in the locally-defined region or vice versa. For example, if an agency had a regional project at the level of *South Asia* where this locally-defined regional code included *Myanmar*, the agency might decide to translate this into the standard CEFDA regional code for *South Asia*, ignoring the fact that the local region included more countries than the standard region.

The approximation approach implies inconsistency in the translation of local regions, since each development agency is free to decide, on a case-by-case basis, how close a match is required. It would be difficult or impossible for the information user, such as a searcher of the DAI CD-ROM, to know how close a match was really required. Relevant information would be consistently missed in searching, with little or no means for the searcher to ensure a complete search.

#### **Analytic approach**

Taking an analytic approach, the originating agency could translate the local regional grouping into a series of standard CEFDA regional and/or country grouping codes. The agency comes up with a number of codes, which in combination describe the same geographical area as the local regional designation. For example, an agency with a project for the locally-defined region *Middle East* might create a CEFDA format record with two occurrences of the *O60 Country/region* field, one with the code for *Northern Africa* and one with the code for *Western Asia* on the grounds that the area covered by the local region is the same as the area covered by the two standard CEFDA regions.

This approach means that activities relating to a single region are not given a single regional identifier, thereby obscuring the difference between regional and interregional activities. It involves a conversion from a single value to multiple values which may be more difficult to automate. It does not eliminate the need to use one of the other two approaches if component regions do not include all the countries of the locally-defined region.

#### **Promotion approach**

The "promotion" option involves translating a locally-defined regional designation into the lowest standard CEFDA region that includes all the countries in the locally-defined region. In other words, if an exact match for a locally-defined region is not found in the CEFDA standard, the originating agency moves up one level in the hierarchy of CEFDA regions until a region is found that includes all countries in the locally-defined region. For example, an agency that has a project in the locally-defined region *Andean region* for which there is no equivalent standard CEFDA region, would look for the next higher region that included all the countries in the local region. In this case, the agency would use the code for the standard region *South America*.

This "promotion" of a local region to a higher CEFDA regional designation encourages retrieving only relevant information in response to a region search, but it also means that users searching on a CEFDA region who do not also search broader regional groupings may be missing relevant information. While this solution is not ideal, it provides for consistent and unambiguous translation of regional designations, and is simple for both originating agencies and users to understand.

**Proposal 3** *The Country/region index in activity databases, such as the DAI CD-ROM, should be associated with a hierarchical thesaurus-like structure which will permit easy selection of related country/regional designations in searching.*

Retrieval systems, such as the DAI CD-ROM, should include a hierarchical, thesaurus-like structure for use in searching using countries and regional groupings. Countries would display as the lowest level of the hierarchy, with one or more regions as their "broader terms". These regions in turn would have the countries that composed them as "narrower terms", and would have as "broader terms" more broadly-defined regions. For example, *Algeria* would have a "broader term" *Northern Africa*, and possibly other broader terms such as *Arab countries* and *Middle East*. The region *Northern Africa* would have as "narrower terms" all the countries in that region (including *Algeria*) and as a "broader term" the region *Africa* to which it belonged.

The definition of this hierarchical structure of groupings and countries would then allow the use of search commands to expand automatically the selection of a search term to other terms of interest. In the same way that the current DAI CD-ROM software supports selection of associated terms from a thesaurus display, the user would be able to select related countries or regions from a thesaurus display of the country/region index.

In the case of the country/region hierarchy, two additional functions could materially assist the user. A "select all" command could be added which would select, with a single command or keystroke, all the associated terms for the displayed term. In addition, an "explode" command would select all terms underneath the term in the geographical hierarchy, down to the lowest level of the hierarchy. For example, if a user "exploded" the term *Africa*, the search would combine *Africa* with all the subregions of Africa (i.e. *Northern Africa*, *Western Africa*, etc.), and all the countries that belonged to any of those subregions. One or both of these functions would replace the need for an automatic expansion of a region into its component countries as in the second edition of the DAI CD-ROM, which prevents retrieving information on a region alone.

**Proposal 4** *Users of the DAI CD-ROM should have a facility to define their own regional groupings for search purposes. These regional groupings would be automatically converted into a search for one or more standard regions or countries when a search was performed.*

In retrieval systems, such as the DAI CD-ROM, provision should be made for users to define search strategies that make use of the user's local definition of regions, which may include country designations or CEFDA-standard regional designations in any combination. This facility will allow users to do a search using locally-defined regional designations and retrieve information with the degree of completeness and specificity that they feel is appropriate. For example, a searcher could define a term *Least developed countries* and list all the countries which they wished to include in this category. A searcher for an organization which did not use a Middle African region in its local regional scheme could define a local *West Africa* region, which would include some of the Middle African countries, as well (perhaps) as the region *Middle Africa* itself. In both cases, it would be up to the user to define the scope of this search-time regional designation; searchers could decide how inclusive or how specific they wanted the search term to be.

There are at least two different ways to implement this facility. One possibility would be to allow the user to add new terms to the hierarchical geographic file referred to above. During installation of the DAI CD-ROM, for example, this file could be transferred to the user's hard disk, and a function provided to allow users to enter new records in this geographic file. Presumably this process would have to be repeated with each new edition of the DAI CD-ROM, since country and regional designations might have changed from one edition to another.

Alternatively, this same search facility could be implemented as a "saved search" capability, where the user builds a search strategy that corresponds to the locally-defined region, and saves the search strategy under the name of the locally-defined region. This saved search could then be called up at any time, and used in collaboration with Boolean operators as with any other search. While saved searches would be written to the hard disk, there would be no need, when using this approach, to copy the geographic file to the hard disk. If the saved search is saved in a format that can be edited (such as a simple text file), the user could simply edit the saved search file if new editions of the DAI CD-ROM should use revised geographic designations.

Which of these two approaches is better depends primarily on the facilities available in the CD-ROM search software. The simpler, more effective and more practical option should be implemented. In any case, this facility should be optional, so that an organization using the DAI CD-ROM which is satisfied with the standard groupings, which has limited search expertise, or which does not search the DAI CD-ROM often, will not be required to use these new facilities, and may continue to use the simple, standard grouping scheme.

## **5 Conclusion**

The problem of regional groupings will never be entirely resolved as long as regional groupings differ from agency to agency. Because the regions into which each organization divides the world are so closely tied to the mandate, philosophy, and administrative structure of that development agency, standards will not soon (if ever) be widespread in this area. The proposals presented here do not resolve the problem of reconciling very different regional groupings in data collected from a large number of different agencies. In particular they do not address the unfamiliarity of the CEFDA standard groupings for many users, an unfamiliarity which can only be overcome through documentation, training and continuing experience with CEFDA-format databases. However, these proposals, though limited, do suggest improvements in the way in which development activity information is exchanged and used that should improve the quality and value of development activity information.

## **Appendix A Some proposed revisions to CEFDA regions**

### ***New terms to allow for combinations within an existing region***

Africa South of Sahara (QMM)	Includes Eastern Africa, Western Africa, Middle Africa, Southern Africa
North and Central America (QNM)	Includes North America and Central America
Latin America (QNN)	Includes Central America and South America
Latin America and Caribbean (QNO)	Includes Latin America and Caribbean
Far East (QRM)	Includes East Asia and Southeast Asia

### ***New terms to allow for combinations across existing regions***

Far East and the Pacific (QXA)	Includes Far East and Oceania
North America and Europe (QXB)	Includes North America and Europe
Asia and the Pacific (QXC)	Includes Asia and Oceania
Europe and CIS (QXD)	Includes Europe and CIS
North Africa and Middle East (QXE)	Includes North Africa and Western Asia

### ***New terms to allow for cultural/linguistic and political groupings***

Arab countries (QVA)  
French-speaking countries (QVB)  
French-speaking Africa (QMN)  
English-speaking Africa (QMO)  
Portuguese-speaking Africa (QMP)  
ACP countries (QVC)  
ASEAN countries (QVD)

### ***Terms with adjusted scope to include problem countries in several regions***

Northern America (QND)	Adjusted to include Mexico
Eastern Africa (QMB)	Adjusted to include Sudan
Western Asia (QRE)	Adjusted to include Afghanistan
Southern Europe (QSD)	Adjusted to include Cyprus

## ***Bibliography***

Jackson, Philip. "A Thesaurus for Enhanced Geographic Access" LASIE, vol. 22 no. 3 (Nov./Dec. 1991), p. 49-60.

Niehoff, Robert and Greg Mack. "The Vocabulary Switching System: Description of Evaluation Studies" International Classification vol. 12, no. 1 (1985), p. 2-6.

Rada, Roy. "Connecting and Evaluating Thesauri: Issues and Cases" International Classification vol. 14 no. 2 (1987), p. 63-69.

Zeng, Lei. "Achieving Compatibility of Indexing Languages in Online Access Environment" Encyclopedia of Information Science, Vol. 50 Suppl. 13. Exec. ed. Allen Kent. New York: Marcel Dekker, 1992. p. 1-24.

## ***Notes***

<sup>1</sup> Di Lauro, Anne. CEFDA: Common Exchange Format for Development Activity Information. Revised ed. [Ottawa: IDRC], 1992. Copies of the CEFDA format are available free of charge from the Coordinating Unit for the Exchange of Development Information, International Development Research Centre, P.O. Box 8500, Ottawa, Canada K1H 3G9 (Telephone: (613) 236-6163 ext. 2574; Facsimile: (613) 563-3858).

<sup>2</sup> International Organizations for Standardization. Codes for the Representation of Names of Countries. 3rd ed. Geneva: ISO, 1988. (ISO 3166-1988). Updates to this standard are issued by the ISO 3166 Maintenance Agency Secretariat, Berlin.

<sup>3</sup> United Nations Statistical Office. Standard Country or Area Codes for Statistical Use. New York: UN, 1982. (ST/ESA/STAT/SER.M/49/Rev. 2)