DOCUMENT RESUME

ED 231 363 IR 050 237

AUTHOR	Dierickx, Harold; Hopkinson, Alan
TITLE	Reference Manual for Machine-Readable Descriptions of Research Projects and Institutions.
INSTITUTION	United Nations Educational, Scientific, and Cultural Organization, Paris (France).
REPORT NO	PGI/82/WS/10
PUB DATE	May 82
NOTE	146p.; Adapted from the Reference Manual for Machine-Readable Bibliographic Descriptions. For a related document, see ED 229 016.
PUB TYPE	Reference Materials - General (130)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Communications; *Computer Oriented Programs; Databases; Definitions; Documentation; Guidelines; *Information Dissemination; *Information Storage; *Information Systems; International Programs; Online Systems; *Research Projects; Specifications; Standards
IDENTIFIERS	Information Exchange Procedures; Machine Readable Data

ABSTRACT

This reference manual presents a standardized communication format for the exchange between databases or other information services of machine-readable information on research in progress. The manual is produced in loose-leaf format to facilitate updating. Its first section defines in broad outline the format and content of applicable records. A matrix shows all mandatory and optional data elements and sub-elements. Names and detailed definitions of each mandatory and optional data element are given in the second part as well as guidance on how the data element content is to be selected and entered onto the machine-readable medium. The third section provides detailed specifications of the carrier format or record structure, character sets and coding, transliteration, physical standards for magnetic tapes, and other aspects which are primarily of concern to computer system designers. The fourth part presents general guidelines for providing documentation on individual implementations of the reference manual. Topics covered include computer-related aspects; subject coverage; indexing, abstracting and editorial policies and practices; selected choices from among alternatives allowed by the reference manual specifications; and extensions for user-defined data fields. Additional information which is needed for the form and presentation of individual data elements and for transliteration is given in appendices. A 26-item bibliography, including relevant international standards, is also provided. (Author/ESR)

PGI/82/WS/10
Paris, 1982

REFERENCE MANUAL FOR MACHINE-READABLE DESCRIPTIONS OF RESEARCH PROJECTS AND INSTITUTIONS

Adapted from the Reference Manual for
Machine-Readable Bibliographic Descriptions

by

Harold Dierickx and Alan Hopkinson

General Information Programme and UNISIST

United Nations Educational, Scientific and Cultural Organization

UNISIST International Centre for Bibliographic Descriptions (UNIBID)

Recommended catalogue entry:

Dierickx, Harold
Reference Manual for Machine-Readable Descriptions of Research Projects and Institutions / Adapted from the Reference Manual for Machine-Readable Bibliographic Descriptions by Harold Dierickx and Alan Hopkinson [for the] UNISIST International Centre for Bibliographic Descriptions (UNIBID) .- Paris : Unesco, 1982.- v, [166] ; 30 cm.-

I - Title
II - Hopkinson, Alan
III - UNISIST International Centre for Bibliographic Descriptions (UNIBID)
IV - Unesco - General Information Programme and UNISIST

CONTENTS

LIST OF TABLES AND DIAGRAMS

CHECKLIST OF PAGES

As of May 1982, the Reference Manual consists of the following pages:

Title page i
Title page verso ii
iii
iv
v

0.1 - 0.4
1.1.1 - 1.1.4
1.2.1 - 1.2.7
1.2.1 - 1.2.7
2.0.1 - 2.0.2
2.001.1
2.R01.1
2.R02.1
2.R03.1
2.R11.1 - 2.R11.3
2.R12.1
2.R13.1 - 2.R13.3
2.R14.1 - 2.R14.4
2.R15.1 - 2.R15.4
2.R16.1
2.R17.1 - 2.R17.2
2.R18.1
2.R19.1 - 2.R19.2
2.R20.1
2.R21.1
2.R31.1
2.R32.1 - 2.R32.2
2.R33.1 - 2.R33.3
2.R34.1
2.R35.1
2.R36.1 - 2.36.2
2.R37.1 - 2.R37.3
2.R38.1 - 2.R38.2
2.R39.1 - 2.R39.2
2.R40.1 - 2.R40.3
2.R41.1 - 2.R41.3
2.R42.1 - 2.R42.3
2.R99.1

3.1.1 - 3.1.11
* 3.2.1 - 3.2.4
4.1.1
4.2.1 - 4.2.3
4.3.1 - 4.3.2
4.4.1 - 4.4.3
4.5.1 - 4.5.2
4.6.1
App. A.1 - App. A.16
App. B.1 - App. B.9
* App. C.1 - App. C.4
App. D.0 - App. D.8
Bibl. 1 - Bibl. 4

* Pages 3.2.2, 3.2.4, App. C.1, App. C.2 and App. C.3 are dated "June 1981".

INTRODUCTION

0.1 Purpose of the Reference Manual

The present Reference Manual is based on the recommendations of a previous study on the development of a manual for standard description of research projects (UNIBID/RM/R/78/01, December 1978) and on recommendations by a number of expert meetings which studied an earlier draft of the Manual. Essentially it is an adaptation of the UNISIST Reference Manual for Machine-Readable Bibliographic Descriptions to suit the particular requirements for descriptions of research programmes or projects and of research institutions. It should be realized from the outset that the major objective of this Manual is to serve as a standardized communication or exchange format for the exchange of machine-readable information on research in progress between data bases or any other type of information services specialized in this field. It is not per se a format for internal (local) computer processing.

In particular the Reference Manual allows the following functions to be performed within or outside the context of its communication format function:

(1) creating descriptions of research projects or institutions with the inclusion of all necessary identifiable entry points for identification, filing, search and other forms of processing;

(2) providing a source for local systems design, including input and output procedures and computer processing formats;

(3) serving manual as well as computerized environments.

In order to avoid any possible misunderstanding some of the concepts used above to describe the major objectives and functions of the Reference Manual are defined below.

Any machine-readable format for the recording and dissemination of information has three major components:

(1) The carrier format : the file and record structure, i.e. the prescribed fixed pattern for arranging and locating records in a machine-readable file and data elements within individual records. An important element of the record structure are the content designators : symbols such as tags, identifiers and indicators which identify or delimit data elements or provide additional information about them.

(2) Data element names and definitions : the name and detailed specification of the content of each data element, including an indication of its components which must be separately identifiable in a computer-based system.

(3) Data element Sets : the specification of particular sets of data elements to be present in records describing particular types of information (e.g. research projects, research institutions).

A clear distinction must be made between a local or internal format and a communication or exchange format. Although both are composed of the three components listed above, they differ in the emphasis placed on each of these three constituent elements and particularly in the specific form the carrier format takes. These differences are the result of different objectives.. An internal format is mainly concerned with ensuring efficient, effective and economic storage and retrieval of information; in other words the chief purpose is efficient and economic computer processing. This calls for a suitable file and record structure which will normally be different from that used for an exchange format. The latter's main objective is to provide a record structure which is hospitable to the requirements of a wide variety of systems and, hence, allows for the unambiguous identification of discrete data elements, appearing at well-defined places in the record, necessary for automatic conversion from the exchange to the local format and vice versa.

0.2 Summary of Contents

The Reference Manual contains the specifications for a comprehensive, self-contained machine-readable communication format, composed of five parts, applicable to the field of information on research in progress.

Part 1 defines in broad outline the format and content of records containing information on research projects or institutions. A matrix shows all the data elements and sub-elements which are either essential (mandatory) or supplementary (optional) for the description of research projects and research institutions respectively.

Part 2 gives names and detailed definitions of each mandatory and optional data element and, where necessary, guidance on how the data element content is to be selected and entered on to the machine-readable medium (magnetic tape in the first instance).

Part 3 provides detailed specifications of the carrier format or record structure, character sets and coding, transliteration, physical standards for magnetic tapes; and other aspects which are primarily of concern to computer system designers.

Part 4 contains general guidelines for providing documentation on individual implementations of the Reference Manual. These guidelines cover all topics which need to be addressed in the documentation, including computer-related aspects, subject coverage; indexing, abstracting and editorial policies and practices, selected choices from among alternatives allowed by the Reference Manual specifications, and extensions for user-defined data fields.

Additional information which is needed for the form and presentation of individual data elements and for transliteration is given in Appendices A, B, C and D. These appendices have been taken without modification from the UNISIST Reference Manual for Machine-Readable Bibliographic Descriptions. The only change is that Appendix D is Appendix E in the manual for bibliographic descriptions. This purely editorial change is due to the fact that, in the other manual, there is an intervening Appendix D, relating to patent documents, which is not applicable to the present publication.

To the greatest extent possible, standards issued by the International Organization for Standardization (ISO) have been applied throughout. In particular, the bibliographic communication format upon which the Reference Manual is based is an implementation of the international standard ISO 2709 : Documentation - Format for bibliographic information interchange on magnetic tape.

0.3 Remarks on Use

It cannot be emphasized too strongly that the data elements defined in the Manual are not to be regarded as exclusive. The authors are well aware that for many applications the description facilities provided in the Manual must be supplemented with additional information. The Manual's purpose is to define a minimum set of data elements which can be agreed upon by services, to facilitate the exchange of information between them, and to enable them to present their computer-based products to the user in a more compatible and therefore more easily usable form.

It should also be realized that the scope of the Reference Manual is limited to defining the representation of its recommended essential and supplementary data elements as they should appear in a computer-readable record for exchange purposes between two or more computer-based systems. Nothing in the Reference Manual should be interpreted as attempting to lay down standards for input or display formats. A local system may choose any input format which is convertible by computer program to the exchange format and vice versa and the exchange format has been designed with the aim of retaining the highest degree of flexibility for deriving different types and arrangements of output, whether in the form of computer printout or printed publications such as various types of directories, registers or indexes.

Manual information systems need to consult in the first place the matrix in Part 1, Part 2 and the Appendices. Parts 1 and 5 are not entirely relevant to them and Part 3 is not of their concern at all, except if they plan to computerize at a later date. Computerized systems will need to consult the entire Manual. However, the Manual is not intended to be a tailor-made input manual for both computerized and non-computerized systems. It is to be regarded in the first place as a specification manual for technical management and system design staff to assist them in devising local systems in such a way that they can exchange files in either direction with other centres which have adopted

the Reference Manual format. In other words, it is not an input manual but a source for drafting local input manuals. It is simply not possible to draft a standard input manual which would be suitable for use without any adaptation or addition by any system. On the other hand, as the Manual is meant to define all essential and a fair number of supplementary data elements necessary for the description of research projects and institutions, it should be possible to prepare local input manuals based on it with a minimum of effort and adaptation.

Further, it should be borne in mind that, although the Manual is in the first place intended for use within computerized environments, it can also be used by non-computerized systems. Implementation of the Manual by these systems will make their computerization at a future stage easier.

When occurring for the first time citations of standards and other works in the text include at least the full title or equivalent, full references are listed at the end of the Manual. Since the citations are relatively few and appear also with a fair amount of detail in the text, no numbering system to link the brief citations and the full references has been used. This procedure will facilitate future updating in as much as adding references will not necessitate renumbering of the entire set.

0.4 Updating and Related Services

Future updating will be easy because the Manual is produced in loose-leaf form. Amendments and additions will be sent out as separate sheets to replace existing pages and to be inserted at the appropriate place in the loose-leaf volume. A check-list of every page and its date of issue is inserted and will be replaced by an amended version each time updates or additions are distributed.

The Reference Manual is being distributed with a questionnaire on the extent of actual or intended use, combined with a request form for receiving updates. Users and all those interested in the Manual are hereby reminded that distribution of future amendments, extensions, information on additional supporting services that may be developed, etc. will not be automatic. These will only be sent to those who have completed and returned the supplied questionnaire/request form to UNIBID.

PART 1: DEFINITIONS, OUTLINE OF RECORD FORMAT AND CONTENTS

CHAPTER 1.1: DATA ELEMENTS AND RECORDS

1.1.1 Records

For the purposes of this Reference Manual, a record is defined as a collection of information which pertains to a research project, programme or institution, and which may be stored in machine-readable form as a self-contained and unique logical structure. A record is likely to include a description of the research project, programme or institution in question, some form of classification and/or indexing applied to the subject dealt with by the project, programme or institution, and possibly additional information. The Reference Manual is concerned with all these aspects of the record, but, in addition to those given in Part 2, additional user-defined data fields may be required in order to carry such other information as may be needed for a particular application. Such local fields should be given tag notations Z01 to Z99 to distinguish them from the standard fields defined in the Manual.

From the computer system point of view, it should be noted that the definition of a record in this Manual applies to a logical record, with no special assumptions regarding the breakdown into physical records or blocks on a recording medium.

1.1.2 Research projects or programmes

For the purposes of this Manual a research project or programme is defined as any kind of research project or programme in progress in any basic or applied science, including the social sciences and the humanities.

Usually, a project has distinct funds for achieving a specific objective whereas a programme has a set of objectives and is composed of several projects. Nevertheless, the distinction between a project and a programme is not always clear. A programme may be considered as a project or as containing several projects. The connotations of the two terms are also different in different languages. Consequently, no attempt is made to include an exclusive definition of project and programme in the Manual; it is left to the individual information services to decide what is a project and what is a programme. From this point onwards the text usually refers to "projects" only although it should be understood as being applicable to "programmes" as well.

1.1.3 Research Institutions

A research institution is defined as any kind of organization which is directly engaged in fundamental or applied research (technological research) in any scientific discipline, including the social sciences and the humanities.

1.1.4 Description of research projects and institutions

The description of a research project or institution is a collection of information which is intended to provide a unique and unambiguous reference to them as may be found in registers on research projects or institutions. The description of a research project or institution should not be confused with the bibliographic descriptions found in the publications of documentation or bibliographic services covering completed research reported on in published or otherwise disseminated documents (a document is any published or unpublished item which may be described in a bibliographic record).

It is important to draw a clear distinction between description and record. The term "description" refers to the information which is required in order to define and identify (i.e. describe) a given research project or institution. A description is made up of a number of "data elements". The term "record", strictly speaking, refers to the structure within which the description is contained, whether stored in machine-readable form or not. A record is made up of a number of data fields which contain the data elements.

1.1.5 Data elements and data fields

A data element is a piece of information forming part of the description and having a specific functional relationship with the research project or institution to which the record refers. Examples of data elements are: project title, name of researcher, name of institution.

Data elements are separately identified within the machine record so that each element can, if desired, be idependently accessed and manipulated by computer program. This is achieved by dividing the record into a series of data fields, identified by field numbers or tags. Data fields are further subdivided into subfields introduced by subfield identifiers. Each data element or component of a data element which is to be separately identifiable normally occupies a given subfield of a data field.

1.1.6 Layout of data fields

The purpose of the present section is to facilitate correct understanding of Parts 1 and 2 of the Manual by means of providing a brief description of the layout of data fields within the context of the overall Reference Manual record structure. More details of the format and structure of the machine-readable record and its various components are given in Part 3.

The machine record has three distinct parts: a fixed-length label or leader, a variable-length directory, and variable-length data fields.

For the purpose of the present section there is no need to elaborate on the label, the contents of which is described in Part 3.

The directory may be regarded as a list of field numbers or tags identifying the data fields which are present in the record, and providing pointers to the location of the fields within the variable-length data part of the record. Thus the field number or tag which identifies the data field is not contiguous with the data field itself.

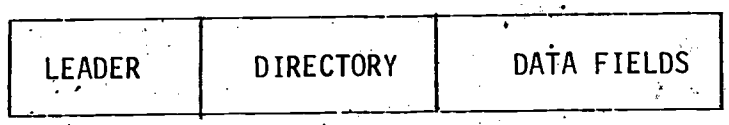
Each data field begins with two or more indicator characters, followed by one or more subfields, followed by a field separator.

The number of indicator characters at the beginning of each field is predetermined for a given implementation of the Reference Manual: the Manual requires a minimum of two, but additional indicators may be included at the user's discretion. Each subfield consists of a subfield identifier followed by a data string. The subfield identifier is a two-character code, of which the first character is the ISO symbol IS_1 (for convenience, represented throughout the Manual by the symbol "@").

The field separator is the ISO symbol IS_2. Whenever a data field is represented in the Reference Manual, however, the field separator is omitted, but should be understood to be always present at the end of the last or only subfield.

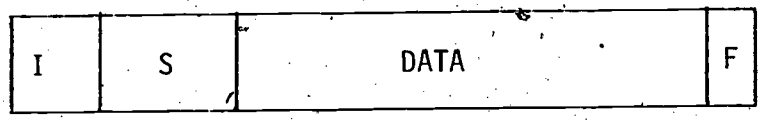
The following is a schematic representation of the record and data field layouts described above:

Record layout:

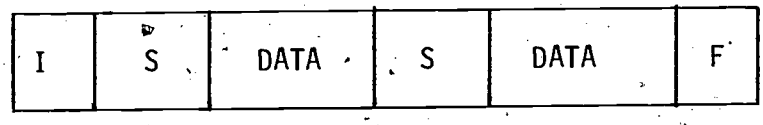


Data field layouts:

Single subfield



Two subfields



(I = indicators, S = subfield identifier, F = field separator)

Examples of data fields as represented in the Reference Manual:

Single subfield: Ø1@ØUNISISTØReferenceØManual
Two subfields: ØØ@1UNESCO@2ICSU

(Here the first two digits are indicators; "@Ø", "@1" and "@2" are subfield identifiers; "Ø" represents "blank" or "space"; note that the field separator is not shown).

1.1.7 Character sets

The Reference Manual is meant to provide an exchange format. A synonymous term for "exchange format" is "communication format". This exchange format is conceived to be receptive to any character set required for a given subject field, but within the limitations of existing ISO character sets and extensions thereof. Consequently, examples of data fields shown in the Manual frequently use a character set which is much wider than is provided by conventional computer coding systems. However, within the general framework of ISO standards, it is open to the user to determine what particular subset he needs to meet his functional requirements.

1.1.8 Summary

Much of the information given in the preceding paragraphs is amplified elsewhere in the Manual, notably in Part 3. The purpose of this section has been to introduce some of the terminology and conventions which are basic to Parts 1 and 2. Essentially the Reference Manual attempts to define an exchange format for the machine-readable record which describes a research project or institution. The record is a collection of data fields as described above.

The remainder of Part 1 is devoted to giving a listing of all essential (mandatory) and supplementary (optional) data elements, defined in detail in Part 2, and from which a minimum number must be selected to construct a valid record in accordance with the specifications of the present Manual.

CHAPTER 1.2: DATA ELEMENT MATRIX

This chapter constitutes a complete reference list of the essential (mandatory) and supplementary (optional) data elements which are defined in full in Part 2 and from which a selection must be made in order to construct a valid record describing a research project or institution.

The category "essential" or "mandatory" is defined as meaning that any data element so designated must be included in the description if it is known or derivable from the information supplied on the research project or institution. In this context, the term "essential" must not be taken to mean that it is necessarily valid in computer system design to incorporate checks which require the inclusion of all essential data elements in all records. Circumstances may arise in which an essential data element is absent. For example, nothing may be known about the completion date of a project. It is expected, however, that the absence of essential data elements will be exceptional.

The category "supplementary" or "optional" is defined as meaning that:

(a) Any data element so designated is regarded as being relevant and worthy of inclusion in the record describing the project or institution.

(b) The data element is not, however, an absolute requirement for a complete and unambiguous description and its inclusion is, therefore, optional, at the discretion of the individual information system.

It should be noted that a data field which is rated as "essential" may include optional subfields. The detailed data element definition in Part 2 will indicate what constitutes the essential portion of each field (e.g. mandatory field R11: PROJECT TITLE has an optional subfield "Abbreviated or other title").

Some users may find that information which is regularly included in their own descriptions is not present among the essential and supplementary data elements listed in this chapter. In this connection it must be stressed that the Reference Manual is not intended to be exclusive; it is expected that users will define additional local data fields, while standardizing on the basic set of data elements given in the Manual. In order to avoid confusion, it is recommended that local fields are tagged from Z01 through Z99 as required.

Only one matrix has been designed to cover both research projects (or programmes) and research institutions. It can be seen from the matrix that, whereas all data elements can be relevant for the description of a project, only a selection can be relevant for the description of institutions. In all there are 27 data elements and 27 corresponding data fields. Of these, eighteen are common to projects and institutions and none are specific to institutions only. Consequently, a separate matrix for institutions has not been included because this would amount to a mere repetition of the majority of the entries for projects. Instead, two separate columns have been entered for projects and institutions respectively under "status". In a computerized environment, a research institutions file could be generated automatically from a research projects file.

The organization of the matrix is straightforward. It shows for each data element the tag, the setting of the indicator positions, the subfield identifiers, and the status (essential or supplementary). Although the matrix provides in one place a general view of all data elements and their status, it is evident that not every conceivable detail can be included. For example, when a data element at the aggregate field level is shown as essential, it may not be possible to indicate specifically the status against each sub-element or variant. A case in point is field R13: NAME OF RESEARCHER. Whereas "Name of Researcher" is an essential data element, the status of each variant in the table cannot be indicated. On the other hand, field R18: CURRENT STATUS OF PROJECT, while being essential at the aggregate field level, can be shown to have sub-elements of different status: the status in coded form being essential, and free description of the status being optional. From these examples it follows that, for the detailed design of input and conversion procedures, the system designer will have to associate the summary information given in the matrix with the comprehensive definitions of data elements and data fields in Part 2.

DATA ELEMENT MATRIX : RESEARCH PROJECT AND INSTITUTIONS

NAME OF DATA FIELD/ELEMENT	FIELD TAG	INDICATORS	SUBFIELD IDENTIFIERS	STATUS	
				PROJ.	INST.
DATA ELEMENTS RELATING TO THE RECORDING OF THE RESEARCH PROJECT					
Record Control Number	ØØ1	-	-	E	E
Date of Input of Record	RØ1	ØØ	@1	E	E
Language of Record	RØ2	ØØ	@1	E	E
Inputting Organization	RØ3	ØØ	@1-3	S	S
Inputting organization: code			@1		
Inputting organization: text			@2		
Country code			@3		
DATA ELEMENTS RELATING TO THE RESEARCH PROJECT OR INSTITUTION					
Project Title	R11	Ø,Ø-4	@1,2	E	
Exact nature of title not specified		ØØ	@1		
Official title		Ø1	@1	E	
Enriched title		Ø2	@1		
Transliterated title		Ø3	@1		
Translated title (by the recorder)		Ø4	@1		
Language of title			@2	E	
Abbreviated or other title			@3	S	
Working Language(s) of Project	R12	ØØ	@1	E	
Name of Researcher	R13	Ø[1/],Ø-4	@1-4,9	E	S
Name: role not specified		ØØ	@1		
Name: principal researcher		Ø1	@1		
Name: other researcher		Ø2	@1		
Name: director or project head		Ø3	@1		
Name: other (role in subfield 9)		Ø4	@1		
Date of birth			@2	S	
Speciality			@3	S	
Country of nationality			@4	S	
Role			@9		
Name of Institution	R14	ØØ	@1-7	E	E
Name of institution			@1	E	E
Address of institution			@2	E	E
Country code			@3	S	S
Region code			@4	S	S
Abbreviation or acronym			@5	S	S
Date of establishment			@6	S	S
Working language(s) code(s)			@7	S	S

DATA ELEMENT MATRIX : RESEARCH PROJECTS AND INSTITUTIONS

NAME OF DATA FIELD/ELEMENT	FIELD TAG	INDICATORS	SUBFIELD IDENTIFIERS	STATUS	
				PROJ.	INST.
Affiliation of Researcher	R15	Ø-9[2],Ø	@1-6	S[3]	S[3]
Name of institution			@1		
Address of institution			@2		
Country code			@3		
Region code			@4		
Abbreviation or acronym			@5		
Date of establishment			@6		
Starting Date of Project	R16	ØØ	@1	E	
Completion Date of Project	R17	Ø;Ø-2	@1	E	
Date unknown		ØØ	@1		
Certain date		Ø1	@1		
Projected date		Ø2	@1		
Current Status of Project	R18	ØØ	@1,2	E	
Current status of project: code			@1	E	
Current status of project: note			@2	S	
Keyword(s)	R19	ØØ[4]	@1,2-...[4]	E[5]	E[5]
Source			@1	E[5]	E[5]
Keyword(s)			@2-...[4]	E[5]	E[5]
Subject Classification Code	R2Ø	ØØ	@1,2	E[6]	E[6]
Name of classification scheme and edition			@1	E[6]	E[6]
Classification number			@2	E[6]	E[6]
Abstract	R21	ØØ	@1,2	E[7]	E[7]
Abstract			@1	E[7]	E[7]
Language of abstract			@2	S	S
Academic Degree for which Research is undertaken	R31	ØØ	@1	S	
Institution Awarding Degree	R32	ØØ	@1-5	S	
Name of institution			@1		
Address			@2		
Country code			@3		
Region code			@4		
Abbreviation or acronym			@5		

DATA ELEMENT MATRIX : RESEARCH PROJECTS AND INSTITUTIONS

NAME OF DATA FIELD/ELEMENT	FIELD TAG	INDICATORS	SUBFIELD IDENTIFIERS	STATUS	
				PROJ.	INST.
Financial Support	R33	ØØ	@1-7	S	S
Source of support			@1		
Amount			@2		
Grant number			@3		
Abbreviation or acronym			@4		
Address			@5		
Country code			@6		
Region code			@7		
Resources	R34	ØØ	@1,2	S	S
Number of staff			@1		
Equipment			@2		
Budget Breakdown	R35	ØØ	@1	S	S
Related Document(s)	R36	Ø;Ø-2	@1	S	
No project documentation is available		ØØ			
Related documents are available		Ø1			
Official project report is available		Ø2			
Availability note			@1	S	
Related Research Project	R37	Ø;Ø-5	@1-3	S	
Relationship unspecified		ØØ			
Related project is predecessor		Ø1			
Related project is successor		Ø2			
Subject relationship		Ø3			
Secondary or parallel related project		Ø4			
Primary related project		Ø5			
Record number			@1		
Details in free form			@2		
Searchable subject relationship			@3		
Number of Project or Contract	R38	ØØ	@1,2	S	
Project number			@1		
Contract number			@2		
Discipline	R39	ØØ	@1,2-..8/	S	S
Source			@1		
Discipline			@2-..8/		
Name of Contact Person	R4Ø	ØØ	@1-3	S	S
Name			@1		
Function			@2		
Address			@3		

DATA ELEMENT MATRIX : RESEARCH PROJECTS AND INSTITUTIONS

NAME OF DATA FIELD/ELEMENT	FIELD TAG	INDI-CATORS	SUBFIELD IDENTI-FIERS	STATUS	
				PROJ.	INST.
Geographical Area	R41	ØØ	@1-4	S	S
Geographical name			@1		
Code			@2		
Source of geographical name			@3		
Source of code			@4		
Cooperating Institution	R42	ØØ	@1-6	S	S
Name			@1		
Address			@2		
Country code			@3		
Region code			@4		
Abbreviation or acronym			@5		
Date of establishment			@6		
Note(s)	R99	ØØ	@1	S	S

DATA ELEMENT MATRIX : RESEARCH PROJECTS AND INSTITUTIONS

NOTES

E = essential (mandatory)
S = supplementary (optional)

No entry at the field level as a whole against a data element in the "status" column means that the data element in question is not relevant for the description of either a project or an institution.

For essential data elements, the status of supplementary sub-elements in individual subfields has been separately indicated whenever practicable to enter and not immediately obvious. This has not been done for supplementary data elements because of the ambiguity that might be created in this way with regard to sub-elements not specifically marked "S" (mark "E" being excluded in principle).

1/ Indicator position 1 may be used to link one or more of the names of researchers in R13 to one or more affiliations recorded in R15. For details see page 2.R15.1.

2/ Indicator position 1 may be used to link the affiliation in field R15 to one or more of the names of researchers entered in field R13. For detailed specifications see page 2.R15.1.

3/ This data element is used only when the name and address of the organization of a person associated with a project or institution is not the same as that of the institution or project described in the record.

4/ In principle subfields other than 1 and indicators are left undefined to allow for maximum flexibility in implementation. For details see page 2.R19.1.

5/ Essential only if no classification code(s) in field R20 or abstract in field R21 are entered.

6/ Essential only if no keyword(s) in field R19 or an abstract in field R21 are entered.

7/ Essential only if no keyword(s) in field R19 or classification code(s) in field R20 are entered.

8/ In principle subfields other than 1 and indicators are left undefined to allow for maximum flexibility in implementation. For details see page 2.R39.1.

PART 2: DATA ELEMENT DEFINITIONS

Part 2 provides detailed definitions of data elements, arranged in alphanumeric order of data field codes.

Each data element is defined in terms of:

(a) a brief summary of the essential features (Field definition);

(b) a detailed description of the data element (Data description),

(c) Examples, wherever necessary and appropriate.

General conventions

The following conventions are applicable to all fields:

(a) Indicators
Indicator positions 1 and 2 are reserved for the uses indicated in the Manual. Where they are not so used, they are entered as zeros. If either or both of the indicator positions is used, the value zero always has the meaning "not specified" (see, for example, field R11).

In the Manual, only two indicator positions are shown. In a specific implementation, one or more extra indicator positions could be inserted, if required, after indicator position 2 and before the first subfield identifier.

(b) Subfield identifiers
As defined in Part 3, a subfield identifier consists of the ISO character IS_1 and one other symbol (usually a numeric digit). For the purposes of illustration, the IS_1 code is represented by the symbol "@". Expressions of the form "subfield 1", "subfield 2" are used to designate "the subfield introduced by the identifier @1", "@2" and so on.

(c) Field separators
The field separator character IS_2 is omitted in all examples, but should be understood as being present in the record as the character immediately following the end of the data string shown in any example.

(d) Character coding
No attempt is made in the examples to reproduce the code structures which would be used in the machine record: all data strings are shown as plain text.

(e) Representation of "zero" and "space"
To avoid ambiguity, the symbol "Ø" is used for the number "zero". "Space" or "blank" is represented by "b".

(f) Implementation options

Where a number of user options exist, it has not always been possible to show all alternatives in the set of examples chosen for a particular data field. In such cases, the selection of a particular option does not imply that this is a "preferred" implementation.

(g) "Notes" subfield

The notes subfield (identifier @N) is an optional subfield which may be included in any data field to incorporate additional free form information which the user wishes to associate specifically with the content of the field. This subfield should be used only as a last resort when the information cannot be entered in any defined field. @N is applicable to any field and so has not been included in the definitions of the fields.

R11: PROJECT TITLE

1. Field definition

Tag: R11
Indicators: Position 1: Ø
Position 2: Ø, 1, 2, 3, 4
Subfields: 1: Title
2: Language of title
3: Abbreviated or other title (optional)
Repeatable: Yes, if it is required to enter more than one form of title, for example in a different language.

2. Data description

Field R11 is used for the project title and is mandatory.

Indicators

Indicator position 2 is used to specify what kind of title is entered in subfield 1 as follows:

Ø Exact nature of title not specified

1 Official title (mandatory)

2 Enriched title (official title plus additional information)

3 Transliterated title

4 Translated title (translated by the recorder).

Subfields

1 Title

The official title of the project should be entered. If different from the language(s) of the data base, it should be translated into that (these) language(s) and entered in subfield 1 of one or more repeated fields as required. If the official title of the project is repeated in a number of languages, these may all be entered, each in a repeated field, or, alternatively, the one in the main working language of the project may be entered alone. If the title of the project is in a language using a non-Roman script it should be transliterated according to the UNISIST recommended transliteration schedules (see Appendix C) and entered in subfield 1. If the official title does not convey a correct or meaningful description of the project, additional information should be added to the official title within square brackets, bearing in mind, however, that an abstract of the project can be entered in field R2Ø.

2 Language of title

Subfield 2 is provided to enable a language code to be entered in order to identify the language of the title. This should be entered if it differs from the working language of the project as given in R12, if more than one language is entered in R12, or if titles in different languages are entered in R11. For a set of language codes see Appendix B.

3 Abbreviated or other title

When a project is commonly known under an abbreviated or other variant title, this variant title may be entered as an optional element in subfield 3.

3. Examples

Example 1

The official title of the research project is:
"La rationalisation des systèmes de traitement de l'information documentaire en archéologie".

Contents of field R11:
Ø1@1LaÞrationalisationÞdesÞsystemesÞdeÞtraitementÞdeÞ
l'informationÞdocumentaireÞenÞarchéologie@2FRE

Assuming that the language of the database is English, a translation may be made and entered in a repeated field R11:
Ø4@1RationalizationÞofÞprocessingÞsystemsÞofÞdocumentaryÞ
informationÞinÞtheÞfieldÞofÞarcheology@2ENG

N.B. : The second indicator is 4 as the translation was made by the recorder

Example 2

A joint Belgian/British project has titles in Dutch (code DUT) and English (code ENG):
"Veredeling van nieuwe polyploide variëteiten van grassen en klavers"
"Breeding of new polyploid grasses and clovers".

Contents of repeated fields R11:
Ø1@1VeredelingÞvanÞnieuweÞpolyploideÞvariëteitenÞvanÞ
grassenÞenÞklavers@2DUTØ1@1BreedingÞofÞnewÞpolyploidÞ
grassesÞandÞclovers@2ENG

N.B. : As each title is an official title, both fields have the second indicator set to 1.

ØØ1: RECORD CONTROL NUMBER

1. Field definition

In accordance with ISO 2709, field ØØ1 does not have indicators or subfields. It is not repeatable.

2. Data description

The record control number identifies the record and should be entered according to the requirements of the inputting organization or the (international) information network in which it is participating. The number which may include any alphanumeric character and which may be of any length as required should be unique. This field is mandatory.

3. Example

(This example is taken from the CARIS Manual)

CARIS record number is "NG 4510001"

Contents of field ØØ1:
NGØ451ØØØ1

RØ1: DATE OF INPUT OF RECORD

1. Field definition

Tag: RØ1
Indicators: ØØ
Subfield: 1: Date
Repeatable: No.

2. Data description

The date of input of a record should be entered in this field according to the format of International Standard ISO 2014: Writing of calendar dates in all-numeric form, i.e. YYYYMMDD, where YYYY = year, MM = month, and DD = day. This field is mandatory.

Note that this date is the date of putting the record on file by the information service and not the date of provision of the information, such as the date of completion of a questionnaire by a researcher. This date is only of importance to the information service itself.

3. Example

The record is entered on 1st May 1974.

Contents of field RØ1:
ØØ@119740501

RØ2: LANGUAGE OF RECORD

1. Field definition

Tag: RØ2
Indicators: ØØ
Subfield: 1: Language of record
Repeatable: Yes, if the record is entered in more than one language.

2. Data description

Field RØ2 is intended for the language of the record, that is the language in which the notes etc. in the record are written. "Language" here is different from "language" in field R12 where the working language(s) of the project described in the record are indicated. Normally the language of the record will be the same as the working language of the information system inputting the record. The language should be entered in the form of a three-letter code. A list of these codes is found in Appendix B. This field is mandatory.

3. Example

The language of the record is French.

Contents of field RØ2:
ØØ@1FRE

RØ3: INPUTTING ORGANIZATION

1. Field definition

Tag: RØ3
Indicators: ØØ
Subfields: 1: Inputting organization: code
2: Inputting organization: text
3: Country code
Repeatable: No.

2. Data description

Field RØ3 is intended for entering the name in coded or text form of the organization entering the record. This field is typically of a "housekeeping" nature, intended for use within co-operative (international) information systems. Tapes which are being exchanged between computerized individual systems will be clearly labelled as to their origin and field RØ3 may then be automatically generated in each record during processing at the receiving centre.

This field is optional. Alternatively, this information may be included in the record control number.

Subfields

1 Inputting organization: code

This subfield is used to enter the code which may have been assigned to the inputting organization in a co-operative information system on research in progress. If no such code exists, the name of the inputting organization is entered in subfield 2.

2 Inputting organization: text

The name of the inputting organization should be entered in the form by which it is usually known. Abbreviations may be used if well known.

3 Country code

The country where the organization is situated should be entered in subfield 3 using the 2-character alphabetic code of ISO 3166: Codes for the representation of names of countries (see Appendix A).

3. Example

The inputting organization is the Library Association of the United Kingdom (abbreviated as LA). The country code for the United Kingdom is GB.

Contents of field RØ3:
ØØ@1LA@3GB

Example 3

A project is officially entitled "An investigation of Late Quaternary continental shelf sediments".

The recorder decides that it is advantageous to state that the investigation is confined to the shelf off the Otago Peninsula, New Zealand, and formulates an enriched title.

Contents of field R11:
Ø2@1AnØinvestigationØofØLateØQuaternaryØcontinentalØshelfØ
sedimentsØ[offØOtagoØPeninsula,ØNewØZealand]@2ENG

Example 4

One of the research programmes of the International Institute of Tropical Agriculture is entitled "Grain Legume Improvement Program". The official title of the programme exists in English and in Spanish. It is also known by the abbreviation "GLIP" in both languages.

Contents of repeated field R11:
Ø1@1GrainØLegumeØImprovementØProgram@2ENG@3GLIPØ1@1
ProgramaØdeØmejoraØdeØlasØlegumbres@2SPA@3GLIP

R12: WORKING LANGUAGE(S) OF PROJECT

1. Field definition

Tag: R12
Indicators: ØØ
Subfield: 1: Language code(s)
Repeatable: No.

2. Data description

Field R12 is used to record the working language(s) of the project. A list of codes is found in Appendix B. Each code should be entered in subfield 1 and separated from the next by comma, space. This field is mandatory.

3. Example

The languages of the project are English and Arabic.

Contents of field R12:
ØØ@1ENG,ßARA

1. Field definition

Tag: R13
Indicators: Position 1: Ø (but see field R15, indicator position 1)
Position 2: Ø, 1, 2, 3, 4
Subfields: 1: Name of person
2: Date of birth
3: Speciality
4: Country of nationality
9: Role
Repeatable: Yes, if more than one name is to be recorded.

2. Data description

This field is used to enter the name of a person involved in a research project. It is mandatory. Each name should be entered in a repeated field.

Indicators

The role of the persons involved in the project should be made known by means of an indicator.

Indicator Position 2

Ø Role not specified
1 Principal researcher
2 Other researcher
3 Director or Project Head
4 Other (role specified in subfield 9)

Subfields

1 Name of person

The name should be entered in subfield 1 in its full form when known. The name should be entered family name (surname) first, followed by personal names (forenames), for example: Browne, John Henry. The family name should be followed by a comma and each forename should be preceded by a space.

If family name and initials only are known, the initials should be entered followed by a full stop in place of a full forename. Initials are not separated from each other by a space, but a space must occur between any initial and a family name or part of a family name (e.g. a prefix).

Titles which are part of the name (Sir, Lord, etc.) should be entered in parentheses following the forenames or initials. Titles or additions to names which are often used in addresses (e.g. Mr, Mrs, Prof., Dr) may be entered in the same way. Titles representing qualifications such as degrees or membership of learned societies, or military honours (e.g. M.A., Ph.D., F.L.A.), should not be entered.

Certain names cause problems because it is not always clear which part of the name is the family name and which the forenames. This category includes double-barrelled names and names containing prefixes such as de, del, de la, van, von, etc. No international standard governing the entry of names in computerized information systems exists. A publication which reflects the treatment favoured by a number of national libraries is: International Federation of Library Associations and Institutions. Names of persons: national usages for entry in catalogues. 3rd ed., London, IFLA International Office for UBC, 1977, 203pp. The UNISIST Reference Manual also contains guidelines on the treatment of personal names, which are reproduced in Appendix E. Given the existing diversity in the treatment of personal names and the lack of an international standard, it is not practical to prescribe precise rules in the present Manual. The UNISIST guidelines may be used but, ultimately, the decision is left to the partners engaged in a particular exchange of information on research in progress.

2 Date of birth

The date of birth of a researcher may be entered in subfield 2, as an optional element, to distinguish him from another scientist with an identical name. The date of birth or the country of nationality of the researcher may be entered in subfield 3 or 4 respectively as an alternative means of distinction. The date of birth should be formatted in accordance with ISO 2014: Writing of calendar dates in all-numeric form, i.e. YYYYMMDD, where YYYY = year, MM = month, and DD = day.

3 Speciality

The speciality of a researcher may be entered in subfield 3, as an optional element, to distinguish him from another scientist with an identical name. The date of birth or the country of nationality of the researcher may be entered in subfield 2 or 4 respectively as an alternative means of distinction.

4 Country of nationality

The country of nationality of a researcher may be entered in subfield 4, as an optional element, to distinguish him from another scientist with an identical name. The date of birth or the speciality of the researcher may be entered in subfield 2 or 3 respectively as an alternative means of distinction.

9 Role

The contents of this subfield should be entered in free form. If this subfield is used the second indicator should be set at 4.

3. Examples

Example 1

The principal researcher is C.D. Van Loon, other researchers are J. Haanstra-Verbeek, and G.M. Tichelaar.

Contents of repeated field R13:

```
Ø1@1Loon,ØC.D.ØVan
Ø2@1Haanstra-Verbeek,ØJ.
Ø2@1Tichelaar,ØG.M.
```

Example 2

The Project Head is Anthony Chapman; the other researchers are Michael Mowat and Frederick Goodridge. Administrative assistant is John Peter Smith.

Contents of repeated field R13:

```
Ø2@1Mowat,ØMichael
Ø2@1Goodridge,ØFrederick
Ø3@1Chapman,ØAnthony
Ø4@1Smith,ØJohnØPeter@9AdministrativeØassistant
```

Example 3

A Spanish agronomist, Dr. A. Garrido Fernandez, is director of a research project. He has a Chilean colleague with an identical surname and initials. Both are well known in the field of plant genetics.

Contents of field R13:

```
Ø3@1GarridoØFernandez,ØA.@4Spain
```

or

```
Ø3@1GarridoØFernandez,ØA.@4ES
```

R14: NAME OF INSTITUTION

1. Field definition

Tag: R14
Indicators: ØØ
Subfields:
1: Name of institution
2: Address of institution
3: Country code (optional)
4: Region code (optional)
5: Abbreviation or acronym (optional)
6: Date of establishment (optional)
7: Working language(s) code(s) (optional)
Repeatable: Yes, if the research is being conducted at more than one institution.

2. Data description

Field R14 is used to enter the name of the institution where research is conducted, the address, country and region code and the acronym or other abbreviated form of the institution name. This field is mandatory.

Subfields

1 Name of institution

Where several levels of the organization exist, the name of the lowest-level unit within the organization that is concerned with the research should be mentioned in any case. The name of the units should be entered from the largest to the smallest. For large and complex organizations, such as some university or government departments, intermediate levels, the inclusion of which does not add significant information to the entry, may be omitted, provided always that the most specific unit is entered and that the entry provides an unambiguous identification of the organization.

In order to facilitate manipulation of the organizational levels by computer, a "+" ("plus" sign) should be inserted instead of a space between each organizational unit.

The following additional conventions apply to this field:

(a) If the name of the institution is not in the Roman alphabet it should be transliterated according to the UNISIST schedules (see Appendix C).

(b) A fuller form of the name than that supplied may be entered if known.

(c) The full name of the institution is mandatory even if an abbreviated form of the name is entered in subfield 5.

(d) The name of the institution should always be entered in its official or working language(s) and, if necessary, may be repeated in other languages. Each different language version of the same institutional name should be placed in a repeated subfield 1. To distinguish between official different language versions and translations by personnel of the data base, the translation should be entered in square brackets.

2 Address of institution

The address or location of the institution should be entered in subfield 2 in the complete form required for postal purposes, ignoring any redundancy which may arise where the place name forms part of the organization (e.g. "University of Cambridge, Cambridge, England"). However, an incomplete address may be entered when no fuller information is available. Useful details such as telephone and telex number, and cable address may be added.

3 Country code

The country where the institution is situated may be entered in subfield 3 using the 2-character alphabetic code of ISO 3166: Codes for the representation of names of countries (see Appendix A).

4 Region code

A code denoting the geographical or geopolitical region in which the country of the institution is situated may be entered in subfield 4.

5 Abbreviation or acronym

In addition to the full name entered in subfield 1, an abbreviation or acronym of the name of the institution may be entered in subfield 5. This should correspond exactly to the full name in subfield 1. If several levels of the organization are cited in subfield 1, the same number of levels should be entered in subfield 5 and each should be separated by "+". If no abbreviation exists for a particular level, that level should be entered in full.

6 Date of establishment

If known, the date when the research institution was established may be entered in subfield 6. The date should be formatted in accordance with the International Standard ISO 2014: Writing of calendar dates in all-numeric form. The day, month and year should be entered in the form YYYYMMDD. Both the day and month may be entered as ØØ if not known precisely or if they are considered as unnecessary information.

Subfield 6 is optional. Consequently, the use of indicator position 2 to indicate the degree of certainty of the date, as in R16, is not explicitly foreseen here. However, parties to an exchange are free to introduce it if they so wish.

In any case, if a date is recorded which is uncertain it should be followed by a question mark. When validation programs require an entry in subfield 6 in all cases and the date of establishment is not known at all, a question mark should be entered.

7 Working language(s) code(s)

The working language(s) of the research institution may be entered in subfield 7 in coded form. A list of codes is found in Appendix B. If more than one code is entered, each should be separated from the next by a comma and a space.

3. Examples

Example 1

The institution where the research is conducted is the Research and Development Section, Property Services Agency (PSA), Department of the Environment. The levels of the organization are entered in descending order. The address is also entered along with the country code and the recognized abbreviated form of the organization.

Contents of field R14:

ØØ@1Department␣of␣the␣Environment+Property␣Services␣Agency+Research␣and␣Development␣Section@24Ø,␣Wellesley␣Road,␣Croydon,␣Surrey,␣Great␣Britain@3GB@5DOE+PSA+Research␣and␣Development␣Section

Example 2

The institution where the research is conducted is the Dyson Perrins Laboratory in the Department of Organic Chemistry at the University of Oxford. As there is only one Dyson Perrins Laboratory, it is permissible to omit Department of Organic Chemistry.

Contents of field R14:

ØØ@1University␣of␣Oxford+Dyson␣Perrins␣Laboratory@2South␣Parks␣Road,␣Oxford@3GB

Example 3

The name and address of a research institute in the Federal Republic of Germany in energy technology is known to be:

"Forschungsinstitut fuer Energietechnik
Hauffstrasse 14
D-7441 Wolfschlugen
Federal Republic of Germany".

The recording service has a policy to include the date of establishment in its records and assumes that this date is 1978.

Contents of field R14:

ØØ@1Forschungsinstitutbfuerbenergietechnik@2Hauffstrassebb
14,bD-7441bWolfschlugen,bFederalbRepublicbofbGermany@3DE@6
19780000?

R15: AFFILIATION OF RESEARCHER

1. Field definition

Tag: R15
Indicators: Position 1: Ø-9
Position 2: Ø
Subfields: 1: Name of institution
2: Address of institution
3: Country code (optional)
4: Region code (optional)
5: Abbreviation or acronym (optional)
6: Date of establishment (optional)
Repeatable: Yes, see indicator position 1.

2. Data description

Field R15 is used to enter the name and address of a single institution to which one or more of the persons associated with the research project are affiliated, when it is not the same institution as that where the research is being conducted. R15 may also be used to record the affiliation of a person associated with the research institution described in the record, when the affiliation is different from the address of the research institution. This field is optional.

Indicators

Indicator position 1

This indicator, the use of which is optional, may be used to link an affiliation to the appropriate personal name in field R13 when a number of personal names require to be linked to different affiliations. This is done by setting the first indicator in the field to a number other than Ø and giving the first indicator of the corresponding field R13: NAME OF RESEARCHER the same value. The numbers 1 to 9 should be used according to the order in which the personal names are entered, the affiliations being entered in the same order. This treatment is possible for any number of names linked to as many as nine affiliations.

Subfields

1 Name of institution

Where several levels of the institution are cited (e.g. laboratory, faculty, university) they should be entered in descending order of scale, from the larger unit to the smaller. For large and complex organizations, such as some university or government departments, discretion may be exercised in omitting intermediate levels, the inclusion of which does not add significant information to the entry, provided always that the most specific unit is cited and that the entry provides an unambiguous identification of the organization.

In order to facilitate manipulation of the organizational levels a "+" ("plus") should be inserted instead of a space between each organizational unit.

The following additional conventions apply to this field:

(a) If the name of the institution is not in the Roman alphabet it should be transliterated according to the UNISIST schedules (see Appendix C).

(b) A fuller form of the name than that supplied may be entered if known.

(c) The full name of the institution is mandatory even if an abbreviated form of the name is entered in subfield 5.

(d) The name of the insitution should always be entered in its official or working language(s) and, if necessary, may be repeated in other languages. Each different language version of the same institutional name should be placed in a repeated subfield 1. To distinguish between official different language versions and translations by personnel of the data base, the translation should be entered in square brackets.

2 Address of institution

The address or location of the institution should be entered in subfield 2 in the complete form required for postal purposes, ignoring any redundancy which may arise where the place name forms part of the name of the organization (e.g. "University of Cambridge, Cambridge, England"). However, an incomplete address may be entered when no fuller information is available. Useful details such as telephone and telex number, and cable address may be added.

3 Country code

The country where the institution is situated may be entered in subfield 3 using the 2-character alphabetic code of ISO 3166: Codes for the representation of names of countries (see Appendix A).

4 Region code

A code denoting the geographical or geopolitical region in which the country of the institution is situated may be entered in subfield 4.

5 Abbreviation or acronym

In addition to the full name entered in subfield 1, an abbreviation or acronym of the name of the institution may be entered in subfield 5. This should correspond exactly to the full name in subfield 1. If several levels of the organization are cited in subfield 1, the same number of levels should be entered in subfield 5 and each should be separated by "+". If no abbreviation exists for a particular level, that level should be entered in full.

6 Date of establishment

If known, the date when the research institution was established may be entered in subfield 6. The date should be formatted in accordance with the International Standard ISO 2014: Writing of calendar dates in all-numeric form. The day, month and year should be entered in the form YYYYMMDD. Both the day and month may be entered as ØØ if not known precisely or if they are considered as unnecessary information.

Subfield 6 is optional. Consequently, the use of indicator position 2 to indicate the degree of certainty of the date, as in R16, is not explicitly foreseen here. However, parties to an exchange are free to introduce it if they so wish.

In any case, if a date is recorded which is uncertain it should be followed by a question mark. When validation programs require an entry in subfield 6 in all cases and the date of establishment is not known at all, a question mark should be entered.

3. Examples

Example 1

Affiliation: "Lubrication Research Laboratory,
Department of Mechanical Engineering,
School of Engineering and
Applied Science,
Columbia University,
New York, NY 10027,
U.S.A."

Contents of field R15:

ØØ@1ColumbiaØUniversity+LubricationØResearchØ
Laboratory@2ColumbiaØUniversity,ØNewØYork,Ø
1ØØ27@3US

Example 2

Affiliation: "Empresa Brasileira de Assistência
Técnica e Extensão Rural (EMBRATER)
Sistema Nacional de Informação
e Documentação Agrícola
Caixa Postal 04-0019
70000 Brasília, D.F.
Brasil"

Contents of field R15:

ØØ@1EmpresaÞBrasileiraÞdeÞAssistênciaÞTécnicaÞeÞ
ExtensãoÞRural@1[BrasilianÞEnterpriseÞforÞTechnicalÞ
AssistanceÞandÞRuralÞExtension]@2CaixaÞPostalÞ
Ø4-ØØ19,Þ7ØØØØÞBrasília,ÞD.F.,ÞBrasil@3BR@5EMBRATER

R16: STARTING DATE OF PROJECT

1. Field definition

Tag: R16
Indicators: ØØ
Subfield: 1: Starting date of project
Repeatable: No.

2. Data description

Field R16 is used to record the starting date of the project. The date should be entered in subfield 1 in accordance with the International Standard ISO 2014: Writing of calendar dates in all-numeric form. The day, month and year should be entered in the form YYYYMMDD. The month may be entered as ØØ if it is not known precisely. The day may always be entered as ØØ. This field is mandatory.

3. Example

The project started on 26 March 1976.

Contents of field R16:
Ø1@119760300

R17: COMPLETION DATE OF PROJECT

1. Field definition

Tag: R17
Indicators: Position 1: Ø
Position 2: Ø, 1, 2
Subfield: 1: Completion date of project
Repeatable: No.

2. Data description

Field R17 is used to record the date of completion of the project. The date should be entered in subfield 1 in ISO standard format. The day, month and year should be entered in the form YYYYMMDD. The month may be entered as ØØ if it is not known precisely. The day may always be entered as ØØ. This field is mandatory.

Indicators

Indicator Position 2

This indicator position is used to indicate the level of certainty of the date in subfield 1. The indicator should be used as follows:

Ø Date unknown

If the completion date is unknown, indicator position 2 should be set at Ø and "Date of completion unknown" entered in subfield 1.

1 Certain date

If the project has come to an end or if the date of its ending is known for certain, indicator position 2 should be set at 1.

2 Projected date

If the project is scheduled to come to an end at a particular date, but the date is only projected or provisional, indicator position 2 should be set at 2.

3. Examples

Example 1

The project finished in December 1979, which is known for certain from a project report published in a journal.

Contents of field R17:
Ø1@119791200

Example 2

The project must end in March 1988 when funding ceases.

Contents of field R17:
Ø1@119880300

Example 3

The project is likely to finish in 1985.

Contents of field R17:
Ø2@119850000

R18: CURRENT STATUS OF PROJECT

1. Field definition

Tag: R18
Indicators: ØØ
Subfields: 1: Current status of project: code
2: Current status of project: note in free form (optional)
Repeatable: No.

2. Data description

Field R18 is intended for entering details on the current status of the research project. Codes from the list below should be entered in subfield 1 and in addition a note of explanation may optionally be entered in subfield 2 in free form. This field is mandatory.

Subfields

1 Current status of project: code

Subfield 1 should contain one of the following codes:

Completed	A
In abeyance	B
Continuing	C
Abandoned	D
Under revision/revised	R

2 Current status of project: note in free form

Subfield 2 is used to note information additional to the code, entered in free form.

3. Examples

Example 1

The project is continuing at the time the record is entered.

Contents of field R18:
ØØ@1C

Example 2

The project is officially in abeyance, but it is unlikely to be revived.

Contents of field R18:
ØØ@1B@2Little␣likelihood␣of␣the␣project␣being␣revived

R19: KEYWORD(S)

1. Field definition

Tag: R19
Indicators: ØØ (see below)
Subfields: 1: Source
2-...: Keyword(s) (see below)
Repeatable: Yes (see below).

2. Data description

Field R19 is used to record one or more words or groups of words which describe the subject content of the research project. Such words denoting subject content are also referred to as index terms, descriptors, keywords, keyword phrases, etc. No particular thesaurus is recommended. Parties to an exchange are left free to determine from which source they will select their keywords, but this source -whether formally published or not - should be mentioned in subfield 1. The use of field R19 is mandatory if no classification code(s) (RØØ) or an abstract (R21) are provided.

Indicators and subfields

In principle the indicators and subfields other than 1 are left undefined and parties to an exchange should agree on defined indicators and subfields according to the requirements of their agreed indexing system for exchange.

Field R19 may be repeated if required for the keyword entry system selected. Field R19 should be repeated, however, if keywords are chosen from more than one source. In that case each different source and the keyword(s) selected from it should be entered in the appropriate subfields of repeated fields R19.

The source should be recorded in free form. Abbreviations may be used if generally known, for example "OECD" in "OECD Macrothesaurus".

Any number of keywords may be entered in one or more subfields starting with subfield 2. Different single or compound keywords should be separately identifiable following conventions agreed by the parties to an exchange.

The obvious reason for the flexibility allowed in this field is the existing diversity in indexing systems used. Consequently, it is recommended to users of the Manual to provide extensive documentation on their particular implementation of this field (see Part 5).

3. Example

An information centre on research in progress uses "free-language" subject terms, i.e. index terms not limited to a set of terms in a thesaurus or similar authority list of terms. The keywords selected to describe the subject aspects of the project being recorded are "cultivation", "turnips", "organic compost".

Contents of field R19:

ØØ@1Uncontrolled¥keywords@2cultivation;¥turnips;¥organic¥ compost

or

ØØ@1Uncontrolled¥keywords@2Cultivation.Turnips.Organic¥ compost

or

ØØ@1Uncontrolled¥keywords@2Cultivation@3Turnips@4Organic¥ compost

RØ: SUBJECT CLASSIFICATION CODE

1. Field definition

Tag: RØ
Indicators: ØØ
Subfields: 1: Name of classification scheme and edition
2: Classification number
Repeatable: Yes, if it is required to apply more than one classification number to a document.

2. Data description

Field RØ is provided to enter a classification number applied to the document in accordance with any classification scheme individual services may wish to use, including "in-house" unpublished schedules and local adaptations of published schemes like the Universal Decimal Classification (UDC), the Dewey Decimal Classification (DC), etc. This field is mandatory if no keyword(s) (R19) or an abstract (R21) are provided.

Subfields

1 Name of classification scheme and edition

The name and edition of the classification scheme applied is entered in subfield 1, in free form. Abbreviations may be used if the name of the scheme is well known in the abbreviated form (e.g. UDC, DC, LC for the Universal Decimal Classification, the Dewey Decimal Classification and the U.S. Library of Congress Classification schemes respectively).

2 Classification number

The number as found in the scheme's schedules is to be entered in subfield 2.

3. Example

The subject of research is the "technology of dam construction in Yugoslavia". The reporting information service is providing classification numbers in accordance with both the FID/Unesco Broad System of ordering (BSO), 3rd revision, 1978, and the Dewey Decimal Classification, 9th abridged edition.

Contents of repeated field RØ:
ØØ@1BSO,Ø3rdØrev.Ø1978@2712,72-Ø26,YU
ØØ@1DCØ9Øabr.@2627.Ø9427

R21: ABSTRACT

1. Field definition

Tag: R21
Indicators: ØØ
Subfields: 1: Abstract
2: Language of abstract (optional)
Repeatable: Yes, if the abstract is given in more than one language.

2. Data description

Field R21 is intended for a short text (abstract) describing a research project or a research institution. It should be entered in the language(s) agreed by the participants in a particular exchange agreement or information system. The use of this field is mandatory if no keyword(s) (R19) or classification number(s) (R2Ø) are provided.

Subfields

1 Abstract

In the abstract for a project should be included the objectives, the methodology used and any applications. The abstract for an institution should include the main field(s) of research and related activities and services.

2 Language of abstract

In some cases when an abstract is provided in more than one language, it may be useful to indicate the language of the abstract given in subfield 1. This should be done in coded form. A set of codes is found in Appendix B.

3. Examples

Example 1

Contents of field R21:

ØØ@1The␣aim␣of␣the␣project␣is␣the␣completion␣of␣the␣biological␣survey␣of␣the␣Flora␣and␣Fauna␣of␣the␣area␣round␣the␣Shackleton␣base␣in␣the␣Antarctic.␣This␣is␣achieved␣by␣on␣the␣spot␣investigation.␣The␣results␣are␣intended␣for␣use␣in␣a␣conservation␣project.

Example 2

A directory of research institutions provides descriptive information on the organizations under three headings: "Aims", "Primary Information" and "Services".

Contents of field R21:

ØØ@1AIMS␣:␣Technical␣and␣economic␣research␣in␣the␣energy␣field.␣Econometric␣models.␣PRIMARY␣INFORMATION␣:␣reports,␣conference␣proceedings.␣SERVICES␣:␣library␣services,␣SDI,␣query␣answering

R31: ACADEMIC DEGREE FOR WHICH RESEARCH IS UNDERTAKEN

1. Field definition

Tag: R31
Indicators: ØØ
Subfield: 1: Type of degree
Repeatable: No.

2. Data description

Field R31 is used to enter a note on the type of degree, if any, for which the research project is being undertaken. An abbreviation, if generally recognized in the language of the data base, may be used. This field is optional.

3. Example

The degree is Doctor of Philosphy (Ph.D.)

Contents of field R31:
ØØ@1Ph.D.

R32: INSTITUTION AWARDING DEGREE

1. Field definition

Tag: R32
Indicators: ØØ
Subfields:
1: Name of institution
2: Address of institution
3: Country code (optional)
4: Region code (optional)
5: Abbreviation or acronym (optional)
Repeatable: No.

2. Data description

Field R32 is used to enter the name of the institution awarding the degree for which the research is being conducted. This field should not be entered if the institution awarding the degree is the same as that where the research is conducted. This field is optional.

Subfields

1 Name of institution

The name of the university, college, etc, should be entered in its official form in subfield 1. It is not necessary to mention the faculty or department. Otherwise, the same rules apply as those for R14: NAME OF INSTITUTION:

(a) If the name of the institution is not in the Roman alphabet it should be transliterated according to the UNISIST schedules (see Appendix C).

(b) A fuller form of the name than that supplied may be entered if known.

(c) The full name of the institution is mandatory even if an abbreviated form of the name is entered in subfield 5.

(d) The name of the institution should always be entered in its official or working language(s) and, if necessary, may be repeated in other languages. Each different language version of the same institutional name should be placed in a repeated subfield 1. To distinguish between official different language versions and translations by personnel of the data base, the translation should be entered in square brackets.

2 Address of institution

The address or location of the institution may be entered in subfield 2 in the complete form required for postal purposes, ignoring any redundancy which may arise where the place name forms part of the organization (e.g. "University of Cambridge, Cambridge, England"). However, an incomplete address may be entered when no fuller information is available. Useful details such as telephone and telex number, and cable address may be added.

3 Country code

The country where the institution is situated may be entered in subfield 3 using the 2-character alphabetic code of ISO 3166: Codes for the representation of names of countries (see Appendix A).

4 Region code

A code denoting the geographical or geopolitical region in which the country of the institution is situated may be entered in subfield 4.

5 Abbreviation or acronym

In addition to the full name entered in subfield 1, an abbreviation or acronym of the name of the institution may be entered in subfield 5.

3. Example

The research is being undertaken for the degree Docteur-Ingénieur at the Université Paris Sud.

Contents of field R32:
ØØ@1Université ØParis ØSud

R33: FINANCIAL SUPPORT

1. Field definition

Tag: R33
Indicators: ØØ
Subfields:
1: Source of support
2: Amount
3: Grant number
4: Abbreviation or acronym
5: Address
6: Country code
7: Region code
Repeatable: Yes, if there is more than one source of support.

2. Data description

Field R33 is used to enter details relating to the financial support of the research project. The field which is optional is divided into seven subfields.

Subfields

1 Source of support

The name of the organization supporting the research project should be entered in subfield 1. Where several levels of the organization exist, it will be sufficient in most cases to enter the highest level only. However, should it be necessary or desirable to enter all or several levels, this should be done following the same conventions as those specified for subfield 1 of R14: NAME OF INSTITUTION. The names of the organizational levels should be entered from the largest to the smallest. For large and complex organizations, such as some university or government departments, intermediate levels, the inclusion of which does not add significant information to the entry, may be omitted, provided always that the most relevant unit is entered and that the entry provides an unambiguous identification of the organization. In order to facilitate manipulation of the organizational levels by computer a "+" ("plus" sign) should be inserted instead of a space between each organizational unit.

The following additional conventions apply to this field:

(a) If the name of the institution is not in the Roman alphabet it should be transliterated according to the UNISIST schedules (see Appendix C).

(b) A fuller form of the name than that supplied may be entered if known.

(c) The full name of the institution is mandatory even if an abbreviated form of the name is entered in subfield 4.

(d) The name of the institution should always be entered in its official or working language(s) and, if necessary, may be repeated in other languages. Each different language version of the same institutional name should be placed in a repeated field 1. To distinguish between official different language versions and translations by personnel of the data base, the translation should be entered in square brackets.

When the project is financed from resources available to the institution where the research is conducted, "Recurrent budget", "Own resources" or a similar expression may be entered.

2 Amount

The amount of the financial support should be entered in subfield 2. Either the total amount over the entire project period should be entered or annual allocations may be entered followed by the year in parentheses. Each entry should be seperated from the next by a semi-colon and a space.

3 Grant number

If the organization which is providing the financial support has allocated an identification number to the grant given to the project being described, this number, if known, may be entered in subfield 3. It should be copied without any alteration to the form in which it is used by the financing institution.

4 Abbreviation or acronym

In addition to the full name entered in subfield 1, an abbreviation or acronym of the name of the supporting institution may be in entered in subfield 4. This should correspond exactly to the full name in subfield 1. If several levels of the organization are cited in subfield 1, the same number of levels should be entered in subfield 4 and each should be separated by "+". If no abbreviation exists for a particular level, that level should be entered in full.

5 Address

The address or location of the institution may be entered in subfield 5 in the complete form required for postal purposes, ignoring any redundancy which may arise where the place name forms part of the name of the organization (e.g. "University of Cambridge, Cambridge, England"). However, an incomplete address may be entered when no fuller information is available. Useful details such as telephone and telex number, and cable address may be added.

6 Country code

The country where the institution is situated may be entered in subfield 6 using the 2-character alphabetic code of ISO 3166: Codes for the representation of names of countries (see Appendix A).

7 Region code

A code denoting the geographical or geopolitical region in which the country of the institution is situated may be entered in subfield 7.

3. Examples

Example 1

The source of support is the United Nations Development Programme, which is contributing the amount of $99,000 over three years. The United Nations Development Programme is well known under the abbreviation "UNDP". The UNDP grant number is CMR/79/415.

Contents of field R33:

```
ØØ@1UnitedÞNationsÞDevelopmentÞProgramme@2$99,ØØØÞ
overÞ3Þyears@3CMR/79/415@4UNDP
```

Example 2

A two-year project is financed on a fifty-fifty basis from the recurrent budget of the Museum Zoologicum Bogoriense (institution where research is conducted), Jalan Juanda 3, Bogor, Indonesia and from a contribution by an interested international institution: Centro Internacional de Agricultura Tropical (CIAT), Apdo. Aéreo 67-13, Cali, Colombia. The total budget is made up as follows:

1981: US$50,000; 1982: US$ 150,000.

The project is being recorded by an English-language data base.

Contents of field R33 (repeated):

```
ØØ@1RecurrentÞbudget@2US$Þ25,ØØØ(1981);ÞUS$Þ75,ØØØ(1982)ØØ@1
CentroÞInternacionalÞdeÞAgriculturaÞTropical@1[International]Þ
CentreÞforÞTropicalÞAgriculture]@2US$Þ25,ØØØ(1981);ÞUS$Þ75,ØØØ
(1982)@4CIAT
```

or

```
ØØ@1OwnÞresources@2US$Þ1ØØ,ØØØØØ@1CentroÞInternacionalÞdeÞ
AgriculturaÞTropical@2US$Þ1ØØ,ØØØ@5Apdo.ÞAéreoÞ67-13,ÞCali,Þ
Colombia
```

(Other variants of recording including more or less of the available information are possible.)

R34: RESOURCES

1. Field definition

Tag: R34
Indicators: ØØ
Subfields: 1: Number of staff
2: Equipment
Repeatable: No.

2. Data description

Field R34 is used to enter the number of staff involved in the research and details of any equipment of special significance used in the project. This field which is optional is divided into two subfields.

Subfields:

1 Number of staff

Details of the number of staff should be entered in this subfield. If it is not already clear from the rest of the record, the staff may be divided according to their roles, for example research officers, administrative staff, the number of each being given. Each type of staff should in this case be separated from the next by a semi-colon and a space.

2 Equipment

Details of any special equipment used by the project should be entered in subfield 2 in free form.

3. Example

A project consists of a project head, eight professional research staff, two laboratory assistants and a secretary. The most important piece of equipment consists of a laser generator.

Contents of field R34:
ØØ@19Ⱥprofessionalⱥstaff;Ⱥ2Ⱥlab.ⱥassistants;Ⱥsecretary@2
laserⱥgenerator.

R35: BUDGET BREAKDOWN

1. Field definition

Tag: R35
Indicators: ØØ
Subfield: 1: Notes on budget breakdown
Repeatable: No.

2. Data description

This field is used to enter in free form details on the composition of the budget. The use of this field is optional.

3. Example

A project has £30,000 for staff salaries, £1,000 for travel and £3,500 for equipment.

Contents of field R35:
ØØ@1Staff␣salaries:£3Ø,ØØØ;␣Travel:£1,ØØØ;␣Equipment:£3,5ØØ

R36: RELATED DOCUMENT(S)

1. Field definition

Tag: R36
Indicators: Position 1: Ø
Position 2: Ø, 1, 2
Subfield: 1: Availability note
Repeatable: No.

2. Data description

This field is optional and is intended to indicate whether any publications or unpublished documents related to the research project are available. Indicator position 2 is used for this purpose as follows:

Indicator Position 2

Ø: No project documentation is available.

1: Related documents are available.

2: The official project report is available.

The information given by means of the above indicator setting may be supplemented by an availability note:

Subfield 1: Availability note

The main objective of subfield 1 is to supply any additional information on available documentation which is considered useful. When the documents are not available from the institution where the research is conducted but from somewhere else the name and address of the organization where the documents can be obtained can be entered. This field can also be used to mention, for example, that a list of project documents is available.

3. Examples

Example 1

The project report is entitled "Final Report on a project to study the strategic developments affecting the long-term energy situation". The author is M. Klee, publisher Johnson of London, date of publication 1980. The document is in English.

Contents of field R36:
Ø2

or

Ø2@1ENG

Example 2

No project documentation is available but an official report is expected to be released in July 1983.

Contents of field R36:
ØØ@1Project␣report␣expected␣to␣be␣released␣in␣July␣1983

or

Ø2@1Planned␣date␣of␣release␣of␣project␣report␣:␣July␣1983

Example 3

A research institution is carrying out a project on domestic solar energy production and storage on behalf of an international petroleum company. No project documentation will be made available to the public.

Contents of field R36:
ØØ

or

ØØ@1Research␣under␣private␣contract

Example 4

Documentation on a research project in China is available, in Chinese, upon request via local Chinese embassies. No information is given whether a final report is or will be available.

Contents of field R36:
Ø1@1Project␣documents␣available,␣in␣Chinese,␣from␣local␣Chinese␣embassies

R37: RELATED RESEARCH PROJECT

1. Field definition

Tag: R37
Indicators: Position 1: Ø
Position 2: Ø, 1, 2, 3, 4, 5
Subfields: 1: Related research project - record number
2: Related research project - details in free form
3: Related research project - searchable subject relationship
Repeatable: Yes, if it is desired to refer to more than one related project

2. Data description

Field R37 is provided for recording any related research project. Preferably, the record number of the related project should be entered in subfield 1, but, alternatively, short details, such as the title of the project, may be entered in free form in subfield 2 if the record number is not known. This field is optional.

Indicators

The second indicator position should be used to denote the relationship between the research project entered in field R37 and the project being described. It should be noted that the meaning attached to the indicator settings refers to the status of the project entered in field R37 and not to the project being described in the main body of the record:

Ø Relationship unspecified
1 The project entered in subfield 1 or 2 is a predecessor of the project described in the main body of the record
2 The project entered in subfield 1 or 2 is a successor of the project described in the main body of the record
3 The project entered in subfield 1 or 2 is related to the project described in the main body of the record in terms of subject matter.
4 The project entered in subfield 1 or 2 is an integral or separate (self-contained) part of the larger (broader) research project or programme described in the main body of the record.
5 The project or programme entered in subfield 1 or 2 contains or coordinates a number of smaller (narrower in scope) research projects, one of which is described in the main body of the record.

Indicators and subfields have been designed to allow for maximum flexibility in implementation: they can be combined in different ways to suit different individual requirements for (searchable) specificity.

Subfields

1 Related research project - record number

When known, the record number of the related research project should be entered in subfield 1 without any alteration or addition.

2 Related research project - details in free form

When the related research project has not been recorded or when the corresponding record number is not known, information should be entered in subfield 2 in free form. Normally this will consist of the title of the related project and the name and address of the institution where it is carried out. Additional information, e.g. details on the nature of the relationship such as subject relationship, may also be given in free form. Subject information in coded or equivalent form should be entered in subfield 3.

3 Related research project - searchable subject relationship

When the scientific discipline applicable to the related project is identical with or otherwise related to that of the project described in the record, the relevant classification number(s) or keyword(s) may be entered in subfield 3. When more than one number or keyword are applicable, each number or term or string of terms should be separated from the next by a semi-colon and a space. Alternatively, subfield 3 may be repeated for each number or term.

If the subject coverage of the related project and the project described in the record are identical, a service may prefer to set indicator position 2 at 3 and not enter anything in subfield 3 depending on the system's objectives and design.

3. Examples

Example 1

The project being recorded is the successor to an earlier project numbered AJ505.

Contents of field R37:
Ø1@1AJ5Ø5

Example 2

The project being recorded "Investigations into single-grain sowing of cereal" is related to one entitled "Single-grain sowing in cereal farming" which has not been entered in the system, and so has no control number. This project is undertaken at the same institute.

Contents of field R37:
ØØ@2Single-grainØsowingØinØcerealØfarmingØ(atØsameØinstitute)

or

Ø3@2Single-grainØsowingØinØcerealØfarmingØ(atØsameØinstitute)

Example 3

A record describes the project "In-depth study of those applying for artificial insemination by donor in order to have a child". Assigned uncontrolled index terms: "artificial insemination"; "medicine"; "law". A recording information service wishes to relate this project to another project entitled: "Medico-legal aspects of human artificial insemination". This project has been assigned the same uncontrolled index terms and the record number 80-H1-01.

Contents of field R37:
Ø3@18Ø-H1-Ø1@3artificialØinsemination;Ømedicine;Ølaw

or

Ø3@18Ø-H1-Ø1

or

Ø4@18Ø-H1-Ø1@3artificialØinsemination;Ømedicine;Ølaw

(in this case the related project is a sub-project of or is collaborating with the project described in the record, i.e. "In-depth study of ...")

R38: NUMBER OF PROJECT OR CONTRACT

1. Field Definition

Tag: R38
Indicators: ØØ
Subfields: 1: Project number
2: Contract number
Repeatable: Yes, if more than one project or contract number is assigned.

2. Data description

Field R38 is used to enter:

(a) the number assigned to the project by the institution where it is being carried out; and/or

(b) the number of the contract under which the research project is conducted.

This field is optional.

Subfields

1 Project number

The project number should always be entered in subfield 1, exactly in the same form, including all punctuation and spaces, as that used by the organization which is responsible for the execution of the project. Care should be taken not to confuse the project number with any other numbers such as the grant number (to be entered in field R33), the contract number (to be entered in subfield 2), or the record control number (to be entered in field ØØ1).

2 Contract number

The contract number should always be entered in subfield 2, exactly in the same form, including all punctuation and spaces, as that used by the organization which commissioned the research work. Care should be taken not to confuse the contract number with any other numbers such as the grant number (to be entered in field R33), the project number (to be entered in subfield 1) or the record control number (to be entered in field ØØ1).

3. Example

A research institution has assigned the number IE/060300/74/0887 to one of its projects. This project is being carried out under a contract, numbered PGI/790061. Some government financial support is also being received under the terms of a grant numbered IE/74/00495.

Contents of field R38:

ØØ@1IE/Ø6Ø3ØØ/74/Ø887@2PGI/79ØØ61

NB: the grant number is not entered here.

R39: DISCIPLINE

1. Field definition

Tag:	R39
Indicators:	ØØ
Subfields:	1: Source 2-...: Discipline (see below)
Repeatable:	Yes, if more than one discipline and/or source for its description are applicable.

2. Data description

Field R39 is used to enter the main scientific discipline or activity with which the project or institution described in the record is concerned.

This information may be entirely or partly identical with or supplementary to that given in R19: KEYWORD(S), R2Ø: SUBJECT CLASSIFICATION CODE and R21: ABSTRACT. Normally, however, the information in R39 will be at a higher generic level than that in fields R19, R2Ø and R21. For example, if the terms "cultivation", "turnips" and "organic compost" are entered in R19, the corresponding entry in R39 could be "horticulture" or "organic horticulture".

The main discipline or activity may be indicated in free form or in accordance with a specific thesaurus, classification system or other type of controlled terminology. Given the existing diversity in subject description, subfields other than 1 and indicators are left undefined and parties to an exchange should agree on defined subfields and indicators according to the requirements of their subject description system.

Subfield 1 is used to indicate, in free form, which classification or indexing system is the source for the entry. If applicable the edition and/or year of publication of the subject identification system should also be given and, if well known, the name of the system may be cited in its abbreviated form.

Field R39 is optional although parties to a particular exchange may decide to make this entry mandatory.

3. Examples

Example 1

A project team is studying the social and economic organization of groups of nomads in the Western Sahara. This project is recorded in a data base using the "Broad System of Ordering" (BSO) classification system. The entry in field R2Ø is "533,36-530,21-023,61". However, on a broader level, this project is concerned with cultural anthropology. An appropriate entry may therefore also be made in field R39: A consistent application of this indexing policy would enable listings of projects and institutions in the broad field of cultural anthropology.

Contents of field R39:
ØØ@1BSO,Ø3rdØrev.Ø1978@2533

Example 2

A project concerned with research in the field of metal fatigue of jet aircraft wings is carried out at an institution specializing in materials testing. A data base specializing in engineering and technology applied research makes the appropriate entries, in uncontrolled indexing terms. In R19, it enters the terms "jet aircraft; wings; metal fatigue; testing".

Contents of field R39:
ØØ@1UncontrolledØindexingØterms@2aircraftØmetalØ
fatigueØtesting

R40: NAME OF CONTACT PERSON

1. Field definition

Tag: R40
Indicators: 00
Subfields: 1: Name of contact person
2: Function of contact person
3: Address
Repeatable: Yes, if there is more than one contact person.

2. Data description

Field R40 is used to enter the name(s) of the person(s) to be contacted for further information on a particular research project or institution. This field is mandatory for records describing research institutions and for records describing research projects when the person to be contacted is different from any of the persons entered in R13, or when it is felt necessary or desirable to specify which person should be contacted from two or more mentioned in R13.

Subfields

1 Name of contact person

The name should be entered in subfield 1 in its full form when known. The name should be entered family name (surname) first, followed by personal names (forenames), for example: Brown, John Henry. The family name should be followed by a comma and each forename should be preceded by a space.

If family name and initials only are known, the initials should be entered followed by a full stop in place of a full forename. Initials are not separated from each other by a space, but a space must occur between any initial and a family name or part of a family name (e.g. a prefix).

Titles which are part of the name (Sir, Lord, etc.) should be entered in parentheses following the forenames or initials. Titles or additions to names which are often used in addresses (e.g. Mr, Mrs, Prof., Dr) may be entered in the same way. Titles representing qualifications such as degrees or membership of learned societies, or military honours (e.g. M.A., Ph.D., F.L.A.), should not be entered.

Certain names cause problems because it is not always clear which part of the name is the family name and which the forenames. This category includes double-barrelled names and names containing prefixes such as de, del, de la, van, von, etc. No international standard governing the entry of names in computerized information systems exists. A publication which reflects the treatment favoured by a number of national libraries is: International Federation of Library Associations and Institutions. Names of persons: national usages for entry in catalogues. 3rd ed., London, IFLA International Office for UBC, 1977, 203pp. The UNISIST

Reference Manual also contains guidelines on the treatment of personal names, which are reproduced in Appendix D. Given the existing diversity in the treatment of personal names and the lack of an international standard, it is not practical to prescribe precise rules in the present Manual. The UNISIST guidelines may be used but, ultimately, the decision is left to the partners engaged in a particular exchange of information on research in progress. However, treatment of personal names should be consistent and adequately documented (see Part 4).

2 Function of contact person

This is an optional subfield which may be used to indicate the particular function, job title, etc. (e.g. Head, Chief Engineer, "Geschaeftsfuehrer") of the person to be contacted. Although not essential this information can be useful for formatting addresses.

If the information available from a contact person is limited or of a very specific nature, this can also be indicated in this field, in parentheses. If this is entered in addition to details on function etc., it should follow that information separated from it by a space.

3 Address

This is an optional subfield which should only be completed when the address of the contact person is different from the address of the institution which is entered in R14. If entered, the address should be given in the complete form required for postal purposes, ignoring any redundancy which may arise where the place name forms part of the name of the organization (e.g. "University of Cambridge, Cambridge, England"). However, an incomplete address may be entered when no fuller information is available. Useful details such as telephone number, telex number, and cable address may be added.

3. Examples

Example 1

The person to be contacted for any type of information on research conducted at a French university institute specializing in economic study of public transport is listed as:

"Madame Nicole Clerc, Bibliothécaire".

Contents of field R40:
ØØ@1Clerc,ØNicole@2Bibliothécaire

or

ØØ@1Clerc,ØNicole(Madame)@2Bibliothécaire

or

ØØ@1Clerc,ØNicole(Mme)

Example 2

An information sheet on an Italian research institution, specializing in applied research on electromechanical constructions and electric systems lists two contact persons, one responsible for corporate planning and another one for technological processes and quality control:

"Dr. Ing. Adelchi Zancan, Direttore Pianificazione, Dr. Ing. Giulio Piazzi, Direttore Tecnologie e Qualità".

Contents of repeated field R40 (in an English-language data base):
ØØ@1Zancan,ÞAdelchi(DrÞIng.)@2DirectorÞCorporateÞ
PlanningØØ@1Piazzi,ÞGiulio(DrÞIng.)@2DirectorÞ
TechnologiesÞandÞQuality

or

ØØ@1Zancan,ÞAdelchi@2(corporateÞplanning)ØØ@1Piazzi,Þ
Giulio@2(technologicalÞprocessesÞandÞqualityÞcontrol)

NB: Additional variants of completing field R40 (repeated) are possible, including mention of the functions in the original language. The two variants shown above convey the same information albeit in a different form.

R41: GEOGRAPHICAL AREA

1. Field definition

Tag:	R41
Indicators:	ØØ
Subfields	1: Geographical name 2: Code 3: Source of geographical name 4: Source of code
Repeatable:	Yes, if it is necessary to cite more than one geographical name.

2. Data description

Field R41 is used to indicate the geographical area(s) - if any - which are covered by or are relevant to the research project or institution described in the record. Note that this information has nothing to do with the place of origin of the record or the geographical location of a project or institution; it exclusively covers the geographical area which is part of or connected with the investigation carried out by a project or which is of particular interest to the activities of a research institution.

Field R41 is optional.

Subfields

1. Geographical name

The name of the geographical area which is or forms part of the subject of investigation of a research project or which is of particular interest to a research institution is entered in subfield 1. A geographical name may refer to any kind of geographical or geopolitical entity such as a country, a political grouping of countries (e.g. an intergovernmental organization like NATO or the Warsaw Treaty Organization), a marine or land area such as the South Western Pacific or the Subtropical Zone, a particular geographical grouping like "Arab Countries", etc.

The full geographical name must always be entered, even if it is also entered in coded form in subfield 2. Exception can be made for geographical entities which are usually referred to in abbreviated form (e.g. NATO, USA).

Since the precise form of geographical names is often determined by the administrative or political framework within which an information system operates, no specific source is recommended. For those who have a free choice in this matter, consultation of the following may be useful:

- ISO 3166-1974. Codes for the representation of names of countries (a maintenance agency frequently issues official amendments to the names and codes)

- Post Office guide. Current edition and supplements. London (UK), H.M.S.O.

- The Times atlas of the world. [Latest] comprehensive edition. London (UK), Times Newspapers, 197. - (The Times Concise Atlas of the World is also published regularly since 1972.)

- Group 3: Geographical Codes, in: Ingrid PRINCE-PERCIBALLI. AGRIS Classification Scheme (FAO-AGRIS-3). Rome (Italy), FAO. Current edition and regular updates available from AGRIS Coordinating Centre, FAO, Rome, Italy.

2 Geographical code

The geographical name may be entered in coded form in subfield 2.

For the names of countries the two-letter alphabetic codes of ISO 3166 are recommended. For all other geographical names, users of the Reference Manual may devise their own codes or adapt any existing coding scheme, including the AGRIS codes referred to above.

3 Source of geographical name

It is recommended to enter the source used for listing geographical names in subfield 3. This is of particular importance in connection with references to geopolitical areas which are often defined in a slightly different way by different organizations. The source must be cited in sufficient detail to make identification possible. In the case of a published source, the edition and/or the year of publication must be mentioned.

If the source for both the name and the code is the same it should be entered only once, in subfield 4.

4 Source of code

When a code has been entered in subfield 2, the source must always be mentioned in subfield 4. The same detail must be given as for the source in subfield 3.

3. Examples

Example 1

A research unit in the legal department of an institute for the study of international relations specializes in comparative studies of legal institutions in English speaking Africa. AGRIS geographical names and codes are used.

Contents of field R41:

ØØ@1AnglophoneÞAfrica@2G1O2@4AGRISÞClassificationÞ
Scheme(Rev.Þ3),ÞRome,ÞFAO

Example 2

A record is made of a research project concerned with growing of high-yielding wheat varieties in Mexico.

Contents of field R41:

ØØ@1Mexico@2MX@4ISO3166-1974

or

ØØ@1Mexico

R42: COOPERATING INSTITUTION

1. Field definition

Tag: R42
Indicators: ØØ
Subfields:
1: Name of institution
2: Address of institution
3: Country code (optional)
4: Region code (optional)
5: Abbreviation or acronym (optional)
6: Date of establishment (optional)
Repeatable: Yes, if there is more than one co-operating institution.

2. Data description

Field R42 is used to enter the name, address etc. of an institution which is co-operating with the institution which is described in field R14. The co-operation may concern a particular research project or may consist of a formal or informal general undertaking at the overall organizational level. The latter case would be reflected in entries of field R42 in records describing research institutions.

This is an optional field.

The specifications for R42 are identical with those for R14.

Subfields

1 Name of institution

Where several levels of the organization exist, the name of the unit within the organization more directly concerned with the co-operation should be mentioned in any case. The name of the units should be entered from the largest to the smallest. For large and complex organizations, such as some university or government departments, intermediate levels, the inclusion of which does not add significant information to the entry, may be omitted, provided always that the most specific unit is entered and that the entry provides an unambiguous identification of the organization.

In order to facilitate manipulation of the organizational levels by computer, a "+" ("plus" sign) should be inserted instead of a space between each organizational unit.

The following additional conventions apply to this field:

(a) If the name of the institution is not in the Roman alphabet it should be transliterated according to the UNISIST schedules (see Appendix C).

(b) A fuller form of the name than that supplied may be entered if known.

(c) The full name of the institution is mandatory even if an abbreviated form of the name is entered in subfield 5.

(d) The name of the institution should always be entered in its official or working language(s) and, if necessary, may be repeated in other languages. Each different language version of the same institutional name should be placed in a repeated subfield 1. To distinguish between official different language versions and translations by personnel of the data base, the translation should be entered in square brackets.

2 Address of institution

The address or location of the institution should be entered in subfield 2 in the complete form required for postal purposes, ignoring any redundancy which may arise where the place name forms part of the organization (e.g. "University of Cambridge, Cambridge, England"). However, an incomplete address may be entered when no fuller information is available. Useful details such as telephone and telex number, and cable address may be added.

3 Country code

The country where the institution is situated may be entered in subfield 3 using the 2-character alphabetic code of ISO 3166: Codes for the representation of names of countries (see Appendix A).

4 Region code

A code denoting the geographical or geopolitical region in which the country of the institution is situated may be entered in subfield 4.

5 Abbreviation or acronym

In addition to the full name entered in subfield 1, an abbreviation or acronym of the name of the institution may be entered in subfield 5. This should correspond exactly to the full name in subfield 1. If several levels of the organization are cited in subfield 1, the same number of levels should be entered in subfield 5 and each should be separated by "+". If no abbreviation exists for a particular level, that level should be entered in full.

6 Date of establishment

If known, the date when the research institution was established may be entered in subfield 6. The date should be formatted in accordance with the International Standard ISO 2014: Writing of calendar dates in all-numeric form. The day, month and year should be entered in the form YYYYMMDD. Both the day and month may be entered as ØØ if not known precisely or if they are considered as unnecessary information.

Subfield 6 is optional. Consequently, the use of indicator position 2 to indicate the degree of certainty of the date, as in R16, is not explicitly foreseen here. However, parties to an exchange are free to introduce it if they so wish.

In any case, if a date is recorded which is uncertain it should be followed by a question mark. When validation programs require an entry in subfield 6 in all cases and the date of establishment is not known at all, a question mark should be entered.

3. Example

An English-language data base on research institutions specializing in energy problems lists a university department in the UK which co-operates on a regular basis with an institution in Costa Rica and in Italy respectively. The name, address, etc. of the two institutions are:

"Escuela de Ciencias Geograficas
Facultad de Ciencias de la Tierra y el Mar
Universidad Nacional
Heredia
Costa Rica"

and

"Istituto di Economia delle Fonti di Energia (IEFE)
Universita Commerciale Luigi Bocconi
Via Sarfatti, 25
20136 Milano
Italia".

The institution in Costa Rica was established towards the end of 1973. The date of foundation of the Italian institute is unknown. The recording data base has a standard entry for this information in both its machine-readable and printed products and translates the name of institutions into English.

Contents of repeated field R42:

ØØ@1Universidad␣Nacional+Escuela␣de␣Sciencias␣Geograficas@1[National␣University+School␣of␣Geographical␣Sciences]@2Heredia,␣Costa␣Rica@3CR@6197300@1Universita␣Commerciale␣Luigi␣Bocconi+Istituto␣di␣Economia␣delle␣Fonti␣di␣Energia@1[University␣for␣Business␣Studies␣Luigi␣Bocconi+Economics␣of␣Energy␣Resources␣Institute]@2Via␣Sarfatti,␣25,␣20136␣Milano,␣Italy@3IT@5IEFE@6?

R99: NOTE(S)

1. Field definition

Tag: R99
Indicators: ØØ
Subfield: 1: Note(s)
Repeatable: Yes.

2. Data description

Field R99 is provided to enter any ancillary data required in the record which cannot appropriately be entered in any of the fields, including their "Notes" subfield, defined in the Manual. This will be the case for data which are relatively informal in nature, or of highly infrequent occurrence.

It must be stressed that, although field R99 has been provided to meet the possibility of an occasional need for the inclusion of ancillary data, its use is recommended only as a last resort. Where an individual service regularly needs to include data elements which are outside the scope of the present Manual, it is recommended that "local" fields tagged from ZØØ to Z99 should be defined for this purpose.

Field R99 may be completed in free form and may be repeated if required.

PART 3: RECORD FORMAT AND RELATED SPECIFICATIONS

CHAPTER 3.1: RECORD FORMAT

3.1.1. Record format: general

The record format in the Manual is to be regarded as a specific implementation of the international standard ISO 2709: Documentation-Format for bibliographic information interchange on magnetic tape, a revised version of which has been published in 1981. The Manual's implementation is based on the revised standard but it should be noted that the previous (1973) edition of ISO 2709 remains a valid implementation of the revised version.

The record structure defined by ISO 2709 will be referred to hereafter as the "ISO record".

The ISO record is divided into three sections: a fixed length label or leader occupying the first 24 characters or bytes, a variable length directory, and data fields of variable length. Although the standard does not mention fixed length fields, such fields may be considered as a special type of variable length fields and, consequently, their use is in conformity with the ISO 2709 record structure. Some aspects of the record structure are described below, but for full details the reader should consult ISO 2709. A diagrammatic representation of the record format is given at the end of Chapter 3.1.

3.1.2 Record format: label or leader

The table below shows the contents of the fixed length label (or leader) at the beginning of each record, as specified in ISO 2709 and as applied in the Reference Manual implementation (an asterisk in the right hand column indicates exact correspondence with the ISO Standard):

Characters (or bytes)	ISO Standard	Reference Manual implementation
0 to 4	Record length	*
5	Record status character (e.g. new, deleting)	Record status character (see notes below; if not used, enter as zero)
6 to 9	Implementation codes	* Reserved for use in bibliographic records
10	Indicator length	Indicator length: minimum 2 bytes for Reference Manual exchange records, but additional indicator positions may be defined by agreement between parties to an exchange (see 3.1.4 below)
11	Identifier length	2 bytes (see 3.1.4 below)
12 to 16	Base address of data	*
17 to 19	For user systems	*
20, 21, 22	Directory map	*
23	For future use	*

Character position 5

indicates the status of a record by means of the following upper case alphabetic codes:

N New record

an entirely new full record.

R Replacement record

substitutes completely a previously issued record.

A Amending record

substitutes part of a previously issued record (users may prefer a replacement record for this purpose)

D Deleting record

deletes an entire record.

I Interim record

a provisional record in anticipation of a definitive full record.

Character positions 6-9

are reserved for "implementation" codes, i.e. codes to be defined at the discretion of individual bibliographic user systems. These character positions should not be used in records describing research projects or institutions and should be entered as zero.

Any character position not used in the record label should be set to zero.

3.1.3 Record format: directory

The directory is a table containing a variable number of fixed length entries, terminated by a field separator code (see 3.1.8).

Each entry corresponds to a specific data field (record identifier, reserved or bibliographic field) in the record, and is divided into four parts (elements):

(1) Tag element: a three character code identifying the content of the data field which corresponds to the directory entry.

(2) Length of datafield element: the number of characters or bytes occupied by the data field which corresponds to the directory entry, including indicators and field separator but excluding the record separator code if the data field is the last field in the record.

(3) Starting character position element: a decimal number giving the position of the first character of the data field which corresponds to the directory entry. The position is computed relative to the base address of the data fields part of the record (i.e. the starting character position of the first data field following the directory), which is zero.

(4) Implementation-defined part element: the implementation-defined part of each directory entry is a new but optional additional element introduced in the revised ISO 2709. Its content is not defined in the standard. In anticipation of possible international guidelines on the use of this fourth element, its use within the Reference Manual context is left optional and undefined.

The length of the directory entries is controlled in character positions 20-22, which, with character position 23 (reserved for future use and hence to be zero filled), make up the directory map. Position 20 determines the length of the "length of data field" element in the directory entry, position 21 determines the number of character positions allocated to hold the starting character position part and position 22 indicates the length of the implementation-defined part. If the latter is not used, position 22 is set to zero.

No part of the directory entry may exceed nine characters in length. Since the tag element is always three characters long, it follows that the maximum total length of any directory entry is thirty characters. However, the revised ISO 2709 also stipulates that "all entries in a directory have the same structure". Within the context of the Reference Manual this is interpreted to mean that, within records as well as within files, the respective lengths of each part of each directory entry and, consequently, the total length of each directory entry, have to be the same. Therefore, the contents of the parts other than that for the tag and the implementation-defined element

will have to be right-justified with zero fill if necessary. Any character position not used in the implementation-defined part will have to be filled with blank or zero but right justification does not apply because the position of the numeric or alphabetic characters to be entered here will be ordered (first, second, etc) in relation to the category of information which is being coded. Obviously zero filling does not apply to the tag element because tags in the Reference Manual are always three-character alpha-numeric.

Where the length of a data field exceeds the largest number (N) which can be stored in the "length" part of the directory entry, two or more successive directory entries are assigned, and the field is treated as if it were divided into a series of parts of length N and a remainder part. Each directory entry referring to a field of this type contains the following elements:

(a) The tag which identifies the field, repeated in all entries.

(b) Length of data field = zero, except in the final directory entry, which contains the length of the remainder part of the data field.

(c) Starting character position of the part to which the directory entry refers.

(d) Length of implementation-defined part = zero, except in the final directory entry, which contains the actual length (contents) of this part if used.

3.1.4 Record format: data fields

In the Reference Manual implementation of the ISO record format, a data field is defined as consisting of:

(a) indicators;
(b) one or more subfields;
(c) a field separator (see introduction to Part 2, page 2.0.1, and 3.1.10 below).

The number of indicators may be varied by individual users to meet their own system requirements. However, the first two indicator positions are reserved for use as defined for each data field in Part 2 of the Reference Manual. Consequently, the minimum length of the indicator in a Reference Manual exchange record is "2", and the reserved indicator positions should not be used for any other purposes. The indicator length is shown in character position 10 of the record leader.

A subfield consists of a subfield identifier followed by a data string which is delimited by either another subfield identifier or a field separator. A subfield identifier, in Reference Manual exchange records, consists of a subfield identifier flag (see introduction to Part 2, page 2.0.1, and 3.1.10 below) and one other character, normally a decimal digit or upper case letter. In Part 2 of the Reference Manual the subfield identifier flag is represented by the symbol "@". The subfield identifier "@N" is reserved for "Notes" subfields.

It will be observed that the Reference Manual implementation of the ISO record format uses only "type 4" fields of the four field alternatives shown in the figure at the end of the standard. An appropriate adaptation of the ISO diagram is given on page 3.1.11.

3.1.5 Record format: tagging scheme

The ISO record format prescribes three-character tags. Early versions of the standard have insisted that tags should be numeric, and this has been the most common implementation practice. The revised standard allows numeric, alphabetic and alpha-numeric tags.

Additionally, ISO 2709 assigns special significance to certain groups of tags as specified below:

Tag ØØ1: record identifier data field. In the Reference Manual implementation, the content of this field is not defined, since the record identifier will vary from one user system to another. The principle of reserving tag ØØ1 for an identifier is to be followed; the specifications for its use are regarded as a matter for agreement between parties to an exchange.

Tags ØØ2-ØØ9: reserved data fields. These are conventionally used to store groups of fixed length data elements which may be required for the processing of the record. They do not carry indicators or subfield identifiers. The use of reserved data fields is not excluded in the Reference Manual but is left undefined. All data elements treated in the Reference Manual are deliberately regarded as variable length, or potentially variable length. In this connection it should be noted that ISO 2709 does not prescribe that data elements in the reserved fields should be fixed length.

Tag assignments in the Manual have been made arbitrarily from base RØØ. It was felt that the allocation of specific tag representations should be unstructured and non-hierarchical, to be consistent with modern "table-oriented" programming methods. This has two benefits: maximum flexibility of assignment, and effectiveness of table-oriented program design. An intellectual structuring of groups of tags assigned to "related" data elements may be useful for some purposes, but this structuring should be reflected in the contents of the tables used to interpret the tags, not in the tag representations themselves. The usefulness of such intellectual groupings is solely for input or output, not for exchange between machine systems.

Users may need more fields than are present in the Reference Manual. Such local fields should be allocated tag notations within the range ZØ1-Z99.

Whereas tags should be listed in ascending alpha-numerical order in the directory, there is no requirement to list fields in the record following the same order.

3.1.6 Linking of records

Certain situations may arise in which it is desirable to cross-reference related research projects. In the present Manual this is made possible by using the repeatable field R37: RELATED RESEARCH PROJECTS, where the nature of the relationship is indicated in the second indicator position and the related project is entered in subfield 1 or 2. The field is repeated for each different related project.

It is assumed that a record exists in the data base of each related project and, if not, that such a record will be created. The normal procedure, therefore, will consist of entering the record number of the related project in subfield 1. Only in those cases when a record describing the related project does not exist, and when for some reason it is impossible or unpracticable to create such a record, should a short description of the related project be entered in subfield 2. This second procedure is to be avoided whenever possible in order to forestall the creation of "nested" or "embedded" records, i.e. sub-records, using the same tags and data elements as the main record, but contained in one data field of the main record. This technique not only may result in unusually long records but also makes access to individual data elements of the embedded records difficult.

3.1.7 Physical tape standards

It should be noted that the assumption is made throughout Part 3 of the Reference Manual that the basic medium for exchange will be nine-track, half-inch magnetic tape recorded at 800 bpi in NRZI mode in an industry-compatible form, complying where applicable with relevant ISO standards. However, in recognition of recent technical developments in relation to formats of commercially available tapes, parties to an exchange are left free to agree on the use of nine-track tape recorded in other modes or at other packing densities.

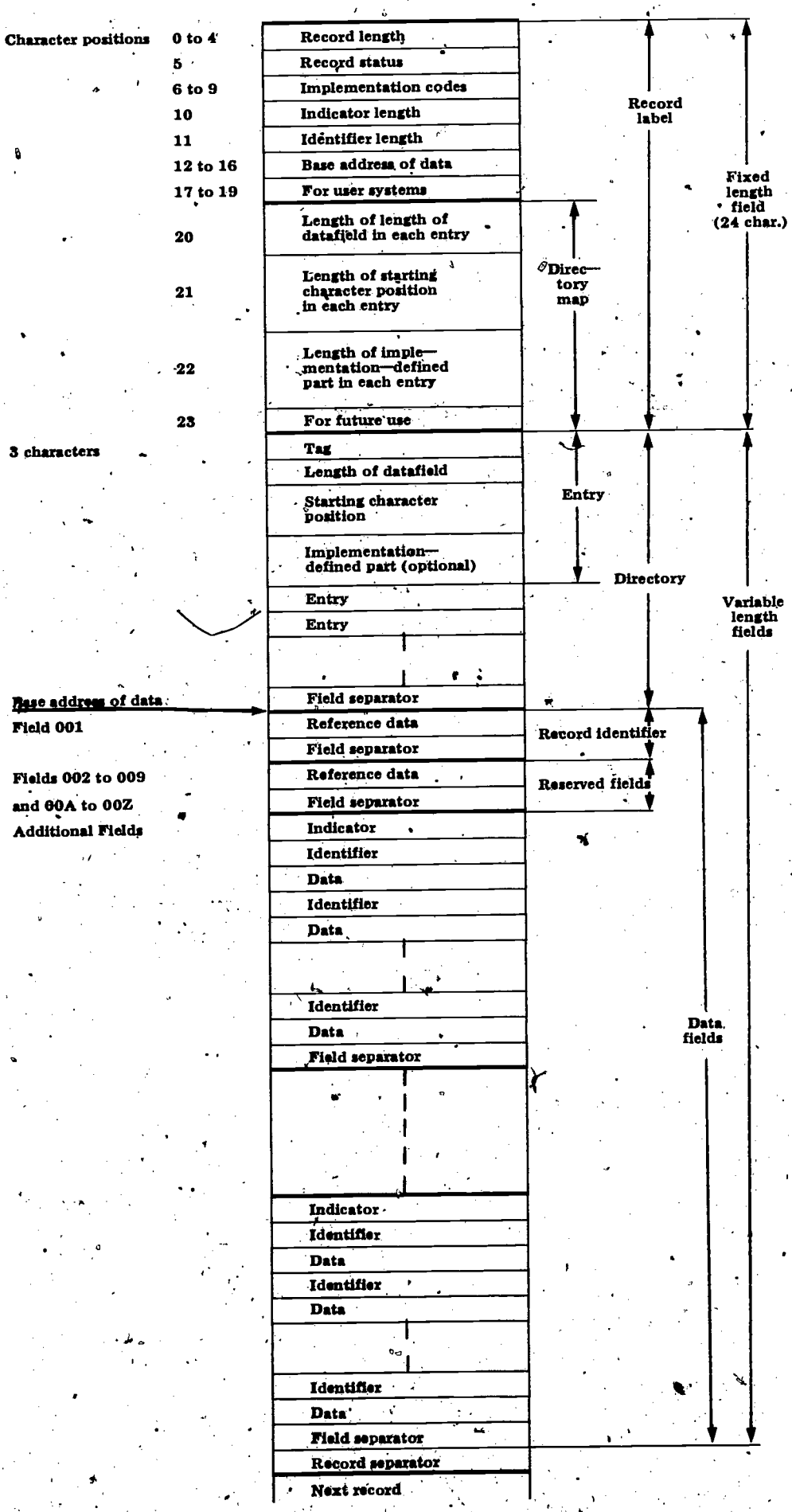
In passing it may be observed that the essential principles and specifications of the Reference Manual are considered to be valid for other physical formats or media, e.g. paper tape, magnetic cassettes, on-line transmission, etc.

3.1.8 Standard separators

The following standard separators or delimiters are used in the ISO record format (ISO 2709), and, therefore, in the Reference Manual exchange format:

Record separator (terminates a complete record)	IS_3 (see Table 1 in Chapter 3.2)
Field separator (terminates a complete data field)	IS_2 "
Subfield identifier flag (introduces a subfield identifier)	IS_1 "

DIAGRAMMATIC REPRESENTATION OF THE REFERENCE MANUAL RECORD STRUCTURE



CHAPTER 3.2 : REPRESENTATION OF CHARACTER SETS

The character set for use in information exchange recommended in conjunction with the Reference Manual is that represented in Table 1 on page 3.2.2. This is a valid implementation of the international standard ISO 646-1973: 7-bit coded character set for information processing interchange. It should be noted that other implementations are possible under the standard but this requires explicit agreement between partners to an exchange. In the absence of any such agreement, users of the Reference Manual are assumed to use the character set given in Table 1 on their exchange tapes.

If more characters and signs than the 128 which can be represented following the conventions of ISO 646 are needed, the code extension techniques described in ISO 2022-1973: Code extension techniques for use with the ISO 7-bit coded character set are to be applied. Correct implementation of the character sets referred to requires consultation of the full text of these ISO standards which are available from national standardization institutions or direct from ISO Central Secretariat, P O Box 56, Geneva 20, Switzerland.

On page 3.2.4 the USSR GOST character set intended for users of the Cyrillic alphabet is represented in Table 2.

TABLE 1: ISO CHARACTER SET ⑤

b_7	0	0	0	0	1	1	1	1
b_6	0	0	1	1	0	0	1	1
b_5	0	1	0	1	0	1	0	1
column / row	0	1	2	3	4	5	6	7
b_4 b_3 b_2 b_1 0 0 0 0	NUL	TC 7 (DLE)	SP	0	@	P	②	p
0 0 0 1	TC 1 (SOH)	DC 1	!	1	A	Q	a	q
0 0 1 0	TC 2 (STX)	DC 2	" ①	2	B	R	b	r
0 0 1 1	TC 3 (ETX)	DC 3	£	3	C	S	c	s
0 1 0 0	TC 4 (EOT)	DC 4	$	4	D	T	d	t
0 1 0 1	TC 5 (ENQ)	TC 8 (NAK)	%	5	E	U	e	u
0 1 1 0	TC 6 (ACK)	TC 9 (SYN)	&	6	F	V	f	v
0 1 1 1	BEL	TC 10 (ETB)	' ①	7	G	W	g	w
1 0 0 0	FE 0 (BS)	CAN	(	8	H	X	h	x
1 0 0 1	FE 1 (HT)	EM	)	9	I	Y	i	y
1 0 1 0	FE 2 (LF)	SUB	*	:	J	Z	j	z
1 0 1 1	FE 3 (VT)	ESC	+	;	K	[	k	③
1 1 0 0	FE 4 (FF)	IS 4 (FS)	, ①	<	L	③	l	③
1 1 0 1	FE 5 (CR)	IS 3 (GS)	–	=	M	]	m	③
1 1 1 0	SO	IS 2 (RS)	.	>	N	^ ①	n	④
1 1 1 1	SI	IS 1 (US)	/	?	O	_	o	DEL

NOTES TO TABLE 1: ISO CHARACTER SET

(1) The graphic characters in positions 2/2, 2/7, 2/12 and 5/14 have respectively the significance of quotation mark, apostrophe, comma and upward arrow head; however, these characters take on the significance of the diacritical signs diaeresis, acute accent, cedilla and circumflex accent when they are preceded or followed by the backspace character in position 0/8.

(2) The symbol in position 6/0 represents grave accent.

(3) Positions 5/12, 7/11, 7/12 and 7/13 are reserved exclusively for specific characters to be explicitly agreed between partners to a particular exchange. These positions are primarily intended for alphabet extensions. If they are not required for that purpose, they may be used for symbols.

(4) Position 7/14 is used for the graphic character overline, the graphic representation of which may vary to represent the tilde or another diacritical sign provided that there is no risk of confusion with another graphic character included in the table.

(5) In the table the columns and rows are identified by numbers written in binary and decimal notations.

Within any one character the bits are identified by b_7, b_6.....b_1, where b_7 is the highest order, or most significant bit, and b_1 is the lowest order, or least significant bit.

Any one position in the table may be identified either by its bit pattern, or by its column and row numbers. For instance, the position containing the digit 1 may be identified:

- by its bit pattern in order of decreasing significance, i.e. 0110001;
- by its column and row numbers, i.e 3/1.

TABLE 2: GOST CHARACTER SET

	0	1	2	3	4	5	6	7
0	ПУС	АР1	Пробел	0	ю	п	Ю	П
1	НЗ	(СУ1)	!	1	а	я	А	Я
2	НТ	(СУ2)	"	2	б	р	Б	Р
3	КТ	(СУ3)	#	3	ц	с	Ц	С
4	КП	СТП	¤	4	д	т	Д	Т
5	КТМ	НЕТ	%	5	е	у	Е	У
6	ДА	СИН	&	6	ф	ж	Ф	Ж
7	ЗВ	КБ	'	7	г	в	Г	В
8	ВШ	АН	(	8	х	ь	Х	Ь
9	ГТ	КН	)	9	и	ы	И	Ы
10	ПС	ЗМ	*	:	й	з	Й	З
11	ВТ	АР2	+	;	к	ш	К	Ш
12	ПФ	РФ	,	<	л	э	Л	Э
13	ВК	РГ	—	=	м	щ	М	Щ
14	ВЫХ	РЗ	.	>	н	ч	Н	Ч
15	ВХ	РЭ	/	?	о	ъ	О	ЗБ

PART 4: A BRIEF GUIDE TO DATA BASE DOCUMENTATION

CHAPTER 4.1: INTRODUCTION

Good documentation is essential for effective use of data bases. Technical details concerning the computer-related characteristics are, of course, necessary. But equally important are the intellectual details of coverage, indexing and abstracting policies, editorial policies and practices, selected choices from among alternatives allowed by the Reference Manual specifications, and extensions for user-defined data fields.

Guidelines are provided for topics which need to be covered in the documentation. There is no intention to prescribe the organization, arrangement, or format of such documentation. The outline of this section can be used as a model, if desired, but this is not essential. Each data base producer will need to determine the purposes which are to be served by the documentation and to design the format, content, organization, depth of treatment, and editorial style accordingly. If desired, the documentation may take the form of an extension to the Reference Manual, identifying any additional rules adopted for the data base, or it may incorporate Reference Manual material where appropriate to provide a self-contained publication. In either case, detailed references to the Reference Manual and to UNIBID should be included for recipients who do not already have the Reference Manual or the UNIBID address readily available.

Data base producers who obtain copyright protection for their documentation should consider granting copying or re-publication permission for those portions of the documentation which describe literature coverage and general editorial characteristics of the data base. Such permission will help ensure completeness and accuracy of data base descriptions included in search guides and similar publications issued by information dissemination centres to their subscribers.

CHAPTER 4.2: DATA CONTENT AND EDITORIAL POLICIES

4.2.1 Introduction

Detailed descriptive information on the data content, and the editorial policies which govern it, is at least as important for the machine-readable data base as it is for companion or related printed publications. This information is often to be found in the introductions to printed products which may be usable for drafting this part of the documentation with little editing. It is strongly recommended that relevant descriptive information be included in its entirety in the data base documentation, rather than referring data base recipients to other published sources which may not be readily available.

4.2.2 Types of Information on Research in Progress Covered

Identify the types of information on research in progress covered by the data base, such as research programmes, projects, institutions, lists of research workers, financing institutions. Provide quantitative data where possible to indicate the approximate numbers of each type for some representative publishing cycle (for example, for a year if the data base uses annual volume numbering). If the data base is partitioned or sequenced by type of research information, provide a description of this arrangement and its relationship to periodically distributed magnetic tapes.

Special attention should be given to any types of information on research in progress which are not described in the Reference Manual, with some indication as to how the Reference Manual specifications have been extended to handle them. This may be by analogy with one of the defined types, citing any fields which have been modified or added to accommodate the additional types, or by a complete description for each additional type similar to what is provided in Part 1 of the Reference Manual.

4.2.3 Selection Criteria

The documentation for data bases with a subject matter, discipline, or mission orientation should include a definition or description of the area of coverage and of the selection criteria used in determining which research projects, institutions, etc will be included. Particular attention should be given to the areas of overlap with related disciplines, so that users of the documentation can make reasonably accurate judgements for inter-disciplinary uses of the data base.

If the data base is a composite file of records obtained from a number of sources, the selection criteria should be specified independently for each source unless they are identical for all. It may also be appropriate for some types of composite data bases to give the approximate number of records from each source. If the distribution form of the data base is partitioned or sequenced by source or other form of selection criteria, a description should be provided.

4.2.4 Indexing and Abstracting Policies

A general description should be provided of indexing, classification, or other means of subject access provided in the data base. The information should include, as appropriate, the authority sources for all indexing or classification schemes, quantitative data on the frequence of occurrence (density) of assigned terms or codes, the representation form of the index or classification entries in the data base (e.g. text, numeric codes, etc.), and any additional characteristics of the scheme, such as weights, roles, importance levels, and so forth. The information provided for each type of indexing need not duplicate the detailed descriptions provided for individual fields (see Chapter 4.5); rather, it should provide a general overview of the subject access techniques used and how they relate to each other. Sufficiently detailed information should be provided for printed, microform, or machine-readable versions of authority sources (texts, thesauri, user guides, etc.), so that purchase orders or other acquisition procedures can be initiated on the basis of the information given.

For data bases which include abstracts or other types of narrative annotation of content, the following information should be provided, as appropriate: conventions used in preparing the abstracts, the language(s) that can occur, abbreviations used, any special conventions employed for graphics, mathematical symbols or other characters not available in the basic character set, the source(s) of abstracts (e.g. whether provided directly by the reporting institution or research worker or prepared by the data base producer), and any other special features which would affect retrieval based on the abstract text or interpretation of the abstract when displayed as part of an annotated directory. If abstracts or annotations of content are included with only some of the records, criteria for this difference should be explained.

4.2.5 Currency Data

General information should be included concerning the elapsed time between the publication dates of the information on research in progress selected and the appearance of its corresponding record in the data base. If possible, figures should be given separately for the different types of information or sources if the differences are significant.

4.2.6 Related Publications

Include in the documentation descriptions of all printed or microform publications available from the data base producer, or other sources, which are related to the data base. This would include companion publications, such as directories, which incorporate some or all of the information recorded in the data base; thesauri, word guides, authority listings, and index code listings; derivative products, such as standard SDI services; and educational or user aid materials. The information provided for each such publication should include a brief description of the work, its publication schedule, its relationship to the data base, and sufficient availability information for acquisition purposes.

4.2.7 Educational Support

If the data base producer, or some other organization, provides consulting, advisory, or educational services related to the data base, such resources and services should be described with sufficient information for their acquisition or use. Such educational support could include periodic seminars and workshops, as well as individual technical support services, educational publications, and visual aid materials. While it may not be feasible or desirable to include an actual schedule of educational training programmes in the documentation, information on how to obtain the schedule should be included, perhaps as an order form.

4.2.8 Document Delivery Services

Those data base producers that provide copies of some or all of the documents produced by the projects cited in their data base, or have arrangements with other organizations to do so, should include in the documentation a description of such services, their source(s), and acquisition procedures.

CHAPTER 4.3: GENERAL CHARACTERISTICS OF THE DATA BASE

4.3.1 Introduction

Data bases vary widely in the design conventions which are adopted and especially in the policies and practices used for handling related research projects and institutions. Information provided on the logical and physical structure of the data base, as well as specific conventions or practices common to the entire data base, greatly simplifies the use of a data base by its recipients.

4.3.2 Distribution Service Content and Schedule

Describe the content of the magnetic tapes distributed, identifying the number of files per issue or other distribution period and the general content of each file. Also, describe the sequence of records in all files containing research or related data (e.g. a data dictionary file) and explain any assumptions related to the sequencing which can or should be taken into account in processing the file.

Describe the schedule(s) for delivery of the distribution service magnetic tapes, such as weekly, monthly, quarterly, and so forth, in as much detail as possible without necessitating frequent revisions to the documentation. If more than one distribution service plan is available for a given data base, all can be described in the documentation with an indication that the choice is the recipient's option. If retrospective or archival collections are available, describe their contents and periods of coverage and the units in which they can be obtained (e.g. volumes or years). If addenda records are issued separately or if corrected data base records are available, describe such services, their distribution frequencies, and the procedures for their procurement.

4.3.3 Definition of a Logical Record

Include in the documentation an explicit definition of what constitutes a single logical record. This can usually be done by relating the contents of a logical record to the corresponding projects or institutions which it describes. Particular attention should be given to those cases where more than one research project or institution is described in the same logical record. Examples of such cases would include a listing of the research projects forming part of a research programme, an enumeration of research institutions co-operating with or subordinated to a particular institution, etc.

Whenever these or similar types of related projects or institutions are handled in the same logical record, information should be provided as to the conventions used. This may be by means of notes, added entries, detailed subfields in user-defined fields, subrecords, or some other technique. When related projects or institutions are handled through the use of the linking field R37 between separate records, information should be included as to the conditions under which the linking technique is employed (i.e. the types of related projects or institutions which are linked) and any conventions as to the relative sequence of such linked records on the exchange tape.

4.3.4 Language and Character Set

Describe the language or languages used in the records of the data base, including policies regarding translation or transliteration. Data bases specifically designed to accommodate two or more languages, for example by means of the inclusion of abstracts in different languages for each record, should highlight these practices and provide guidance on selection criteria, if necessary, for those recipients that may wish to use only one of the data base languages.

Identify the character set(s) used in the records and the specific transliteration standards or conventions used for languages not handled by the selected character sets. Transliteration tables must be included in the documentation for any character set conversions which are not international (ISO) standards, are not given in the Reference Manual, or which have been modified from either the ISO or Reference Manual standards. Conventions for handling characters, symbols, or typographic conventions not described in official standards must also be explained.

4.3.5 Lease/Licence Agreement Provisions

The documentation for data bases which are distributed under lease, licence, or other contractual agreements should include a description of any restrictions on the use of the data base, general pricing information such as provisions for royalties (though not necessarily specific costs), requirements for display of copyright statements or other indications of proprietary interests, and similar information. If desired, specimen agreements can be included. Such procurement documents are frequently issued and retained by business offices, and the information contained in the agreement may not be readily available to the technical personnel actually using the data base.

CHAPTER 4.4: TECHNICAL SPECIFICATIONS

4.4.1 Technical Documentation

Data bases are created and copied for distribution on many different kinds of computers throughout the world, and they are subsequently processed by recipients on a similar variety of computers. The technical specifications should clearly identify the make and model of computer used for generating the distribution magnetic tapes, including the operating or executive system. Having this information available can often provide insight into data processing problems associated with reading and processing the magnetic tape services which would otherwise be intractable. Also, the control or command language and acronyms used with one make of computer are frequently different from those used with other makes, and this difference between computers can be further complicated by language differences among data processing personnel using the magnetic tape services. In preparing the technical documentation for a data base, it is advisable to write out technical terms in full, supplemented by acronyms or codes if desired, and to follow as far as possible the data processing vocabulary given in ISO 2382: Data processing vocabulary. All deviations from ISO Standards and from the standard conventions for the computer systems on which the magnetic tape was generated should be clearly identified and described in detail.

4.4.2 Internal Magnetic Tape Labels

Identify the national or international standard followed for constructing internal labels for magnetic tape, if the tapes have internal labels. Otherwise, explicitly state that the tapes are not labelled. Describe in detail the content of the producer-defined data fields in the labels, such as volume-serial identification and the data set name.

4.4.3 External Magnetic Tape Labels

All magnetic tapes issued as part of a data base producer's distribution service should have external (paper) labels securely affixed to each physical tape reel containing the following information:

- Data base producer's name
- Data base name
- Type of tape (e.g. regular, update, reissue, etc)
- Volume number and/or Year
- Issue number, or other issue identification
- International Standard Number for the data base
- Recording density
- Number of recorded tracks
- Number of individual reel (Reel ... of ...)
- Block Size (Physical record size)
- Logical record size
- Character set
- Number of records
- Internal volume serial number (if labelled)
- Data set name (if labelled)
- Date of tape (usually date of creation).

Other information, such as a customer identifier or copyright statement is at the discretion of the tape producer. The documentation itself should contain a sample of the external labels used by the data base producer and instructions for the interpretation and use of any data or codes which are not self-explanatory.

4.4.4 Recording Density

All recording densities in which the distribution version of the data base is available should be identified, with an indication of the procedures to be followed to change from one density to another (e.g. when recipients' computer tape drives are upgraded or replaced). If other recording densities are available on a special basis, include this information in the documentation as well.

4.4.5 Number of Recorded Tracks

Identify the number of data tracks used in recording the magnetic tapes (e.g. 7 tr. or 9 tr.). If more than one form is available, indicate which combinations with recording densities are available. If the mapping of data bits into recording tracks does not correspond to ISO 962: Information processing - Implementation of the 7-bit coded character set and its 7-bit and 8-bit extensions on 9-track 12,7mm (0.5 inch) magnetic tape, a detailed explanation of the recording convention must be provided.

4.4.6 Block Size, Logical Record Size, and Spanning Techniques

Identify the physical block size(s) and logical record size(s) in which the distribution tape service is available. If logical records can span physical record and/or block boundaries, describe the technique used, either by reference to the appropriate international standard or with detailed specifications and examples.

This part of the documentation should also specify whether the procedure for entering fields the length of which exceeds the largest number which can be stored in the "length" part of the directory, described in 3.1.3, has been followed.

4.4.7 Character Coding Scheme

Identify the character coding scheme used in representing the data on the magnetic tape, either by reference to the appropriate ISO standard(s) or by including a table specifying the bit patterns and their corresponding graphic characters. If extended character sets are used, in accordance with, for example, the methods described in ISO 2022-1973: Code extension techniques for use with the ISO 7-bit coded character set, provide the reference table of bit patterns or hexadecimal codes and their corresponding graphics.

4.4.8 Magnetic Tape Replacement

Describe the policies and procedures for obtaining replacements of magnetic tapes damaged in shipment, not received, unprocessable, or containing errors. Specify in detail any information required from the recipient for tapes found to be unprocessable or to contain errors. If a form is to be completed to report the problem or to obtain a replacement, include a specimen in the documentation.

4.4.9 Technical Assistance

Data base producers that provide technical advise or assistance in the use of their magnetic tape services should include the name of the person or office to be contacted, the telephone and telex numbers as appropriate, the address to which problem reports are to be sent or from which requested materials are to be obtained, and a description of the types of assistance available (or not provided).

CHAPTER 4.5: DATA FIELD DESCRIPTIONS

4.5.1 Introduction

This section of the documentation should describe the data content of the data base on a field by field basis. Descriptions for fields which are already defined in the Reference Manual can be limited to identifying the selected alternatives, where choices are allowed, and to explaining any deviations or extensions. Descriptions for user-defined fields (the Z series of fields) should follow the same outline and format as is used in Part 2 of the Reference Manual.

4.5.2 Data Field/Element Matrix

Include in the documentation a matrix or table identifying the data elements and corresponding fields, which can occur in the description of research projects and institutions as implemented in the data base. The data elements/fields should include not only those defined in the Reference Manual which also occur in the data base, but also any other ones for which the UNISIST specifications have been extended. The codes used in the matrix or table may simply indicate possible presence or absence, or may provide further information as appropriate to the data base (e.g. E: essential (mandatory), i.e. required to be present; S: supplementary (optional), i.e. may be present).

4.5.3 Reference Manual Field Descriptions

Include in this section of the documentation one entry for each of the Reference Manual fields which can occur in the data base. For each entry, identify the subfields which are actually used (as opposed to those which are available for use), the values which each of the two defined indicator positions can have and whether the field is repeatable. If additional indicator positions are defined for a particular data base implementation, define the codes and their meanings for each, in the same format as used in the Reference Manual. Similarly, if additional subfields have been defined, provide descriptions of their use, supplemented with examples where appropriate. For those subfields defined in the Reference Manual which permit variations according to the producer's editorial policies and practices, provide a description of the data content for each, including examples and citations to relevant authority sources where appropriate.

If the design chosen for the data base documentation is a self-contained document, rather than a supplement or companion to the Reference Manual, a section should be included with each field description stating its agreement with the Reference Manual specifications or identifying significant differences, as described in the preceding paragraphs. This will allow recipients using several different data bases in the Reference Manual format to identify quickly those characteristics of the data base which make it unique or different.

All data bases have the potential for being used retrospectively, for producing special-purpose listings on request, for preparing merged or composite data bases, and for similar applications. When used for such purposes, it is essential to know what changes, if any, have occurred in the data content or its representation over the years of the data base, and mention of these should be made in the documentation. This would include such information as when new data elements (fields or subfields) were introduced to the data base, changes in the meanings of codes, discontinuation of fields or subfields, changes in transliteration policies, adoption of national or international standards which affected the content or representation of data, etc.

4.5.4 User-defined Field Descriptions

This section of the documentation should include one entry for each of the fields added by a data base producer as a user-defined field (i.e. a Z-series field). The recommended format is that of the field descriptions in Part 2 of the Reference Manual: Field Definition, Data Description, and Examples. In addition, information regarding changes over time should be incorporated as it relates to the particular data base.

CHAPTER 4.6.1: EXAMPLES

4.6.1 Representative records

The documentation should include examples of representative records from the data base. It is particularly important that examples be included to illustrate special cases and the handling of related projects, which may require inclusion of records exemplifying the use of field R37 as the method for linking records.

When there is a companion or related printed publication, the most effective and useful examples are those which show both the printed (published) form and a representation of the corresponding data base record(s), formatted in accordance with the Reference Manual specifications for preparation of machine-readable input as given in Part 2. Annotations can be used to identify data elements and to point out differences between the two forms.

Appendix A: Country Codes

ISO 3166-1981: Codes for the representation of names of countries provides a two-letter and a three-letter alphabetic code as well as a three-digit numerical code for representing the names of countries, dependencies and other areas of special geopolitical interest for purposes of international exchange.

The Reference Manual recommends the use of the two-letter codes at all times.

The list below has been copied from the standard. Users wishing to be kept informed about amendments should register for this purpose with the

ISO 3166 Maintenance Agency Secretariat
c/o DIN (Deutsches Institut fuer Normung)
Burggrafenstrasse 4-10
Postfach 1107
D-1000 Berlin 30
Telephone (international): 49 30 26 01 36 2
Telegrams: DEUTSCHNORMEN BERLIN
Telex: 184 273 din d

International organizations may register with:

ISO Central Secretariat
1, rue de Varembé
Case postale 56
CH-1211 Geneve 20
Switzerland
Telephone (international): 4122 34 12 40
Telegrams: Isorganiz
Telex: 23 887 ISO CH

SECTION ONE
ALPHABETICAL LIST OF ENTITIES AND CODES IN ENGLISH

SECTION UN
LISTE ALPHABETIQUE DES ENTITES EN ANGLAIS ET CODES

ENTITY (short name in English) Official name in English	Alpha-2 code*	Remarks
1	2	3
AFGHANISTAN Democratic Republic of Afghanistan	AF	
ALBANIA People's Socialist Republic of Albania	AL	
ALGERIA People's Democratic Republic of Algeria	DZ	
AMERICAN SAMOA	AS	
ANDORRA	AD	
ANGOLA People's Republic of Angola	AO	
ANTARCTICA	AQ	The territory south of 60° south latitude
ANTIGUA	AG	Includes Barbuda, and Redonda
ARGENTINA Argentine Republic	AR	
AUSTRALIA Commonwealth of Australia	AU	Includes Lord Howe Island, Macquarie Island, Ashmore, Cartier
AUSTRIA Republic of Austria	AT	
BAHAMAS Commonwealth of the Bahamas	BS	Turks and Caicos Islands not included

*Changes in previous short names or codes from ISO 3166-1974 are shown in column 3. The previous codes are now part of the reserved code list which is available on request from the Maintenance Secretariat.

1	2	3
BAHRAIN State of Bahrain	BH	
BANGLADESH People's Republic of Bangladesh	BD	
BARBADOS	BB	
BELGIUM Kingdom of Belgium	BE	
BELIZE	BZ	
BENIN People's Republic of Benin	BJ	Previous entry: Dahomey DY
BERMUDA	BM	
BHUTAN Kingdom of Bhutan	BT	
BOLIVIA Republic of Bolivia	BO	
BOTSWANA Republic of Botswana	BW	
BOUVET ISLAND	BV	Also called Bouvetøya
BRAZIL Federative Republic of Brazil	BR	Includes Rocas, Fernando de Noronha Archipelago, Trinidade, Ilhas Martim Vaz, and São Pedro e São Paulo
BRITISH INDIAN OCEAN TERRITORY	IO	Chagos Archipelago
BRITISH VIRGIN ISLANDS	VG	Include Anegada, Jost Van Dyke, Tortola, and Virgin Gorda
BRUNEI	BN	
BULGARIA People's Republic of Bulgaria	BG	
BURMA Socialist Republic of the Union of Burma	BU	

1	2	3
BURUNDI Republic of Burundi	BI	
BYELORUSSIAN SSR Byelorussian Soviet Socialist Republic	BY	
CAMEROON, United Republic of	CM	
CANADA	CA	
CANTON AND ENDERBURY ISLANDS	CT	
CAPE VERDE Republic of Cape Verde	CV	Previous entry: Cape Verde Islands. Includes Boa Vista, Brava, Fogo, Maio, Sal, Santo Antão, São Nicolau, São Tiago, and São Vicente
CAYMAN ISLANDS	KY	Include Grand Cayman, Cayman Brac, and Little Cayman
CENTRAL AFRICAN REPUBLIC	CF	
CHAD Republic of Chad	TD	
CHILE Republic of Chile	CL	Includes Easter Island, Juan Fernández Islands, San Felix, and Sala y Gomez
CHINA People's Republic of China	CN	See also Taiwan, Province of China
CHRISTMAS ISLAND	CX	Australian Christmas Island
COCOS (KEELING) ISLANDS	CC	
COLOMBIA Republic of Colombia	CO	Includes San Andrés y Providencia, Malpelo Islands, Roncador Bank, Serrana Bank, and Serranilla Bank
COMOROS Federal and Islamic Republic of Comoros	KM	Previous entry: Comoro Islands. Includes Anjouan, Grande Comore, Moheli, and other islands
CONGO People's Republic of the Congo	CG	

1	2	3
COOK ISLANDS	CK	
COSTA RICA Republic of Costa Rica	CR	Includes Cocos Islands
CUBA Republic of Cuba	CU	
CYPRUS Republic of Cyprus	CY	
CZECHOSLOVAKIA Czechoslovak Socialist Republic	CS	
DENMARK Kingdom of Denmark	DK	
DJIBOUTI Republic of Djibouti	DJ	Previous entry: French Afars and Issas AI
DOMINICA Commonwealth of Dominica	DM	
DOMINICAN REPUBLIC	DO	
DRONNING MAUD LAND	NQ	Part of Antarctica
EAST TIMOR*	TP	Previous entry: Portuguese Timor. Includes the exclave of Oe-Cussi
ECUADOR Republic of Ecuador	EC	Includes Galapagos Islands (Archipelago de Colon)
EGYPT Arab Republic of Egypt	EG	
EL SALVADOR Republic of El Salvador	SV	
EQUATORIAL GUINEA Republic of Equatorial Guinea	GQ	Includes Rio Muni, Macias Nguema Biyogo, Annobon, Corisco, and Elobey
ETHIOPIA	ET	
FAEROE ISLANDS	FO	

* Provisional change of name

1	2	3
FALKLAND ISLANDS (MALVINAS)	FK	Include West Falkland, and East Falkland, South Georgia, and South Sandwich Islands
FIJI	FJ	Include Viti Levu, Vanua Levu, and Rotuma Islands
FINLAND Republic of Finland	FI	
FRANCE French Republic	FR	
FRENCH GUIANA	GF	
FRENCH POLYNESIA	PF	Includes Society Archipelago, Tuamotu Archipelago, Marquezas Islands, Tubuai Islands, Gambier Islands, Austral Islands, and Clipperton Island.
GABON Gabonese Republic	GA	
GAMBIA Republic of the Gambia	GM	
GERMAN DEMOCRATIC REPUBLIC	DD	
GERMANY, FEDERAL REPUBLIC OF	DE	
GHANA Republic of Ghana	GH	
GIBRALTAR	GI	
GREECE Hellenic Republic	GR	Includes Aegean Islands, Ionian Islands, Dodecanese Islands, Crete, and Mount Athos autonomous area
GREENLAND	GL	
GRENADA	GD	Includes Southern Grenadine Islands

1	2	3
GUADELOUPE	GP	Includes Grande Terre, Basse Terre, Marie Galante, Les Saintes, Iles de la Petite Terre, Désirade, Saint-Barthélémy and Northern St. Martin
GUAM	GU	
GUATEMALA Republic of Guatamela	GT	
GUINEA Revolutionary People's Republic of Guinea	GN	
GUINEA-BISSAU Republic of Guinea-Bissau	GW	
GUYANA Republic of Guyana	GY	
HAITI Republic of Haiti	HT	
HEARD AND MC DONALD ISLANDS	HM	
HONDURAS Republic of Honduras	HN	Includes Swan Islands
HONG KONG	HK	Also called Hisiangkang and Xianggang
HUNGARY Hungarian People's Republic	HU	
ICELAND Republic of Iceland	IS	
INDIA Republic of India	IN	Includes Amindivis, Laccadives, Minicoy, Andaman Islands, Nicobar Islands, and Sikkim
INDONESIA Republic of Indonesia	ID	
IRAN Islamic Republic of Iran	IR	
IRAQ Republic of Iraq	IQ	

1	2	3
IRELAND	IE	
ISRAEL State of Israel	IL	
ITALY Italian Republic	IT	
IVORY COAST Republic of the Ivory Coast	CI	
JAMAICA	JM	Includes Morant Cays, and Pedro Cays
JAPAN	JP	
JOHNSTON ISLAND	JT	Johnston Island is the main island in Johnston atoll
JORDAN Hashemite Kingdom of Jordan	JO	
KAMPUCHEA, DEMOCRATIC	KH	Previous entry: Khmer Republic
KENYA Republic of Kenya	KE	
KIRIBATI	KI	Previous entry: Gilbert Islands GE. Part of the former Gilbert and Ellice Islands.* Includes Fanning Island, Washington Island and Christmas Island in the Line Islands, Ocean Island, Phoenix Islands (Birnie, Gardner, Hull, Mc Kean, Phoenix, Sydney; see separate entry for Canton and Enderbury).
KOREA, DEMOCRATIC PEOPLE'S REPUBLIC OF	KP	
KUWAIT State of Kuwait	KW	
LAO PEOPLE'S DEMOCRATIC REPUBLIC	LA	Previous entry: Laos

* See Tuvalu

RM/RPI

May 1982

1	2	3
LEBANON Lebanese Republic	LB	
LESOTHO Kingdom of Lesotho	LS	
LIBERIA Republic of Liberia	LR	
LIBYAN ARAB JAMAHIRIYA Socialist People's Libyan Arab Jamahiriya	LY	Previous entry: Libya
LIECHTENSTEIN Principality of Liechtenstein	LI	
LUXEMBOURG Grand Duchy of Luxembourg	LU	
MACAU	MO	Previous entry: Macao. Also called Ao-men
MADAGASCAR Democratic Republic of Madagascar	MG	
MALAWI Republic of Malawi	MW	
MALAYSIA	MY	Includes Peninsular Malaysia, Sabah, and Sarawak
MALDIVES Republic of Maldives	MV	
MALI Republic of Mali	ML	
MALTA Republic of Malta	MT	
MARTINIQUE	MQ	
MAURITANIA Islamic Republic of Mauritania	MR	
MAURITIUS	MU	Includes Rodrigues, Agalega Island, and Cargados Carajos

1	2	3
MEXICO United Mexican States	MX	
MIDWAY ISLANDS	MI	
MONACO Principality of Monaco	MC	
MONGOLIA Mongolian People's Republic	MN	
MONTSERRAT	MS	
MOROCCO Kingdom of Morocco	MA	
MOZAMBIQUE People's Republic of Mozambique	MZ	
NAMIBIA	NA	
NAURU Republic of Nauru	NR	
NEPAL Kingdom of Nepal	NP	
NETHERLANDS Kingdom of the Netherlands	NL	
NETHERLANDS ANTILLES	AN	Include Aruba, Bonaire, Curaçao, Saba, St. Eustatius, and Southern St. Martin
NEUTRAL ZONE	NT	Territory between Saudi Arabia and Iraq
NEW CALEDONIA	NC	Includes Isle of Pines, Loyalty Islands, Huon Islands, Belep Archipelago, Chesterfield Islands, and Walpole
NEW ZEALAND	NZ	Includes Antipodes Islands, Auckland Islands, Bounty Islands, Campbell Island, Kermadec Islands, Chatham Islands, and Snares Islands

1	2	3
NICARAGUA Republic of Nicaragua	NI	
NIGER Republic of the Niger	NE	
NIGERIA Federal Republic of Nigeria	NG	
NIUE	NU	Previous entry: Niue Island
NORFOLK ISLAND	NF	
NORWAY Kingdom of Norway	NO	
OMAN Sultanate of Oman	OM	Includes Kuria Muria Islands
PACIFIC ISLANDS (trust territory)	PC	Include Carolines, Mariana Islands (except Guam), and Marshall Islands
PAKISTAN Islamic Republic of Pakistan	PK	
PANAMA Republic of Panama	PA	
PAPUA NEW GUINEA	PG	Includes Bismarck Archipelago, Louisiade Archipelago, Admiralty Islands, d'Entre-casteaux Islands, Northern Solomon Islands (Bougainville, Buka, etc.), Trobriand Islands, New Britain, New Ireland, Woodlark, and associated islands
PARAGUAY Republic of Paraguay	PY	
PERU Republic of Peru	PE	
PHILIPPINES Republic of the Philippines	PH	

1	2	3
PITCAIRN ISLAND	PN	Includes Henderson, Ducie, and Oeno
POLAND Polish People's Republic	PL	
PORTUGAL Portuguese Republic	PT	
PUERTO RICO	PR	
QATAR State of Qatar	QA	
REUNION	RE	Includes Ile Europa, Bassas da India, Juan de Nova, Iles Glorieuses, and Ilé Tromelin
ROMANIA Socialist Republic of Romania	RO	
RWANDA Rwandese Republic	RW	
ST. HELENA	SH	Includes Ascension, Gough Island, Inaccessible, Nightingale Islands, and Tristan da Cunha
ST. KITTS-NEVIS-ANGUILLA	KN	Also called St. Christopher-Nevis -Anguilla
SAINT LUCIA	LC	Previous entry: St. Lucia
ST. PIERRE AND MIQUELON	PM	
SAINT VINCENT AND THE GRENADINES	VC	Previous entry: St. Vincentm Includes Northern Grenadine Islands
SAMOA Independent State of Western Samoa	WS	Previous entry: Western Samoa
SAN MARINO Republic of San Marino	SM	

1	2	3
SAO TOME AND PRINCIPE Democratic Republic of Sao Tome and Principe	ST	
SAUDI ARABIA Kingdom of Saudi Arabia	SA	
SENEGAL Republic of Senegal	SN	
SEYCHELLES Republic of Seychelles	SC	Include Alphonse, Bijoutier, St. Francois Islands, St. Pierre Islet, Cosmoledo Islands, Amirantes, Aldabra, Farquhar, and Desroches
SIERRA LEONE Republic of Sierra Leone	SL	
SINGAPORE Republic of Singapore	SG	
SOLOMON ISLANDS	SB	Previous entry: British Solomon Islands. Include Southern Solomon islands, primarily Guadalcanal, Malaita, San Cristobal, Santa Isabel, Choiseul
SOMALIA Somali Democratic Republic	SO	
SOUTH AFRICA Republic of South Africa	ZA	Includes Walvis Bay, Marion Islands, and Prince Edward Islands
SPAIN Spanish State	ES	
SRI LANKA Democratic Socialist Republic of Sri Lanka	LK	
SUDAN Democratic Republic of the Sudan	SD	
SURINAME Republic of Suriname	SR	

1	2	3
SVALBARD AND JAN MAYEN ISLANDS	SJ	Include Bear Island (Bjørnøya)
SWAZILAND Kingdom of Swaziland	SZ	
SWEDEN Kingdom of Sweden	SE	
SWITZERLAND Swiss Confederation	CH	
SYRIAN ARAB REPUBLIC	SY	Previous entry: Syria
TAIWAN, PROVINCE OF CHINA	TW	Includes Penghu (Pescadores) Islands
THAILAND Kingdom of Thailand	TH	
TOGO Togolese Republic	TG	
TOKELAU	TK	Previous entry: Tokelau Islands
TONGA Kingdom of Tonga	TO	
TRINIDAD AND TOBAGO Republic of Trinidad and Tobago	TT	
TUNISIA Republic of Tunisia	TN	
TURKEY Republic of Turkey	TR	
TURKS AND CAICOS ISLANDS	TC	
TUVALU	TV	New name for the Ellice component of the former Gilbert and Ellice Islands.* Includes Funafuti, Nanumanga, Nui, Nanomea, Nurakita, Niutao, Nukufetau, Nukulaelae, and Vaitupu
UGANDA Republic of Uganda	UG	

* See Kiribati

RM/RPI

May 1982

1	2	3
UKRAINIAN SSR Ukranian Soviet Socialist Republic	UA	
UNITED ARAB EMIRATES	AE	
UNITED KINGDOM United Kingdom of Great Britain and Northern Ireland	GB	Includes Orkney, Shetland Islands, Channel Islands, and Isle of Man
UNITED STATES United States of America	US	
UNITED STATES MISCELLANEOUS PACIFIC ISLANDS	PU	Include Kingman Reef, Baker Islands, Howland Islands, Jarvis Islands, and Palmyra Islands
UNITED STATES VIRGIN ISLANDS	VI	
UPPER VOLTA Republic of the Upper Volta	HV	
URUGUAY Eastern Republic of Uruguay	UY	
USSR Union of Soviet Socialist Republics	SU	
VANUATU	VU	Previous entry: New Hebrides NH
VATICAN CITY STATE (HOLY SEE)	VA	
VENEZUELA Republic of Venezuela	VE	
VIET NAM Socialist Republic of Viet Nam	VN	Previous entries: Democratic Republic of Viet Nam VD and Republic of Viet Nam
WAKE ISLAND	WK	
WALLIS AND FUTUNA ISLANDS	WF	Includes Iles de Horn, Ile Uvea, and Ile Alofi
WESTERN SAHARA*	EH	Previous entry: Spanish Sahara
YEMEN Yemen Arab Republic	YE	

* Provisional change of name

RM/RPI

May 1982

1	2	3
YEMEN, DEMOCRATIC People's Democratic Republic of Yemen	YD	Includes Islands of Kamaran, Perim, and Socotra
YUGOSLAVIA Socialist Federal Republic of Yugoslavia	YU	
ZAIRE Republic of Zaire	ZR	
ZAMBIA Republic of Zambia	ZM	
ZIMBABWE	ZW	Previous entry: Southern Rhodesia RH

Appendix B: Language Codes

An ISO standard for language codes is in the course of preparation to replace an existing recommendation (ISO/R639) which is now considered to be incomplete and unsuitable for use in machine systems.

Until such time as this is available, parties to an exchange may use the UNISIST language codes employed by the ISDS (International Serials Data System) which are reproduced here.

Language	Code
Acholi	ACH
Acoli see Acholi	
Afrinili	AFH
Afrikaans	AFR
Afro-Asiatic (Other)	AFA
Ainu see Miscellaneous	
Akan Group see Niger-Congo (Other)	
Akkadian	AKK
Albanian	ALB
Aleut	ALE
Algonquin	ALG
Aljamia	AJM
Amarinya see Amharic	
Amharic	AMH
Ancient Greek see Greek, Classical	
Ancient Hebrew see Hebrew	
Anglo-Norman see Romance (Other)	
Anglo-Saxon (ca. 600-1100)	ANG
Annamese see Vietnamese	
Anzanite see Elamite	
Apache	APA
Arabic	ARA
Aramaic	ARC
Arapahoe	ARP
Araucanian	ARN
Arawak	ARW
Armenian	ARM
Armoric see Breton	
Ashanti see Niger-Congo (Other)	
Assamese	ASM
Assyro-Babylonian see Akkadian	
Avar	AVA
Avaric see Avar	
Avesta	AVE
Avestan see Avesta	
Aymara	AYM
Azerbaijani	AZE
Azeri see Azerbaijani	
Aztec see Nahuatl	
Baltic (Other)	BAT
Baluchi	BAL
Bamana see Bambara	
Bambara	BAM
Bantu see Niger-Congo (Other)	
Bashkir	BAK
Basque	BAQ
Bedja see Beja	
Beja	BEJ
Belorussian	BEL
Bemba	BEM
Bengali	BEN
Berber Group	BER

Language	Code
Bihari	BIH
Biluchi see Baluchi	
Bishari see Beja	
Blackfoot	BLA
Bohemian see Czech	
Breton	BRE
Bulgarian	BUL
Bulgarian, Old see Church Slavic	
Burmese	BUR
Bushman see Sub-Saharan African (Other)	
Byelorussian see Belorussian	
Caddo	CAD
Cambodian	CAM
Canarese see Kannada	
Carib	CAR
Castillian see Spanish	
Catalan	CAT
Caucasian (Other)	CAU
Celtic Group	CEL
Central American Indian(Other)	CAI
Cewa see Chewa	
Chaldean see Aramaic	
Chamorro see Malayo-Polynesian (Other)	
Chechen	CHE
Cherokee	CHR
Chewa	CEW
Cheyenne	CHY
Chibcha	CHB
ChiChewa see Chewa	
Chinese	CHI
Chinook	CHN
Chippewa see Ojibwa	
Choctaw	CHO
Chorti see Mayan	
Church Slavic	CHU
Chuvash	CHV
CiNyanja see Nyanja	
Classical Greek see Greek, Classical	
Coptic	COP
Cornish	COR
Cree	CRE
Creek see Muskogee	
Creoles and Pidgins	CRP
Croatian see Serbo-Crotian (Roman)	
Cushitic (Other)	CUS
Czech	CZE
Dakota	DAK
Danish	DAN
Dano-Norwegian see Norwegian	
Delaware	DEL
Denca see Dinka	
Devanagari (script) see Sanskrit	
Dinka	DIN
Dravidian (Other)	DRA

Language	Code	Language	Code
Duala	DUA	Ga	GAA
Dutch	DUT	Gaelic (Irish) see Irish	
Dutch, Middle (ca. 1050-1350)	DUM	Gaelic (Scots)	GAE
Efik	EFI	Galla	GAL
Egyptian	EGY	Ganda see Luganda	
Elamite	ELX	Ge'ez see Ethiopic	
English	ENG	Georgian	GEO
English, Middle (ca 1100-1400)	ENM	German	GER
English, Old see Anglo-Saxon		German,Middle High(ca 1050-1850)	GMH
Erse see Irish		German, Old High (ca. 750-1050)	GOH
Eskimo	ESK	Germanic (Other)	GEM
Eskimoan see Eskimo		Gondi	GON
Esperanto	ESP	Gothic	GOT
Estonian	EST	Greek, Biblical see Greek Classical	
Ethiopic	ETH	Greek, Classical	GRC
Ewe	EWE	Greek, Modern	GRE
Fang	FAN	Guarani	GUA
Faroese	FAR	Guerze see Kpelle	
Farsi see Persian, Modern		Gujarati	GUJ
Finnish	FIN	Hausa	HAU
Finno-Ugrian (Other)	FIU	Hawaiian	HAW
Flemish see Dutch		Hebrew	HEB
Fon	FON	Herero	HER
French	FRE	Hindi	HIN
French, Middle (ca. 1400-1600)	FRM	Hindustani (Arabic) see Urdu	
French, Old (ca. 842-1400)	FRO	Hindustani (Nagari) see Hindi	
Frisian	FRI	Hottentot see Sub-Saharan African (Other)	

Language	Code
Hungarian	HUN
Hupa	HUP
Iai see Malayo-Polynesian (Other)	
Icelandic	ICE
Ilocano	ILO
Indic (Other)	INC
Indo-European (Other)	INE
Indonesian	IND
Interlingua	INT
Iranian (Other)	IRA
Irish	IRI
Iroquois	IRO
Isi-Kosa see Xhosa	
Italian	ITA
Japanese (Use for related Japanese languages and dialects)	JPN
Javanese	JAV
Javanese, Old see Malayo-Polynesian (Other)	
Judaeo-Arabic	JRB
Judaeo-German see Yiddish	
Judaeo-Persian	JPR
Judaeo-Spanish see Yiddish	
Kachin	KAC
Kafir see Xhosa	
Kamba	KAM
Kanarese see Kannada	
Kannada	KAN
Kanuri	KAU
Karakalpak	KAA
Karen	KAR
Kashmiri	KAS
Kawi see Malayo-Polynesian (Other)	
Kazakh	KAZ
Kechua see Quechua	
Kewa see Papuan-Australian (Other)	
Khmer see Cambodian	
Khotanese	KHO
Kikuyu	KIK
KiMbundu see Mbundu	
Kinyarwanda	KIN
Kirghiz	KIR
Kirundi see Rundi	
Kongo	KON
Korean (Use for related Korean languages and dialects)	KOR
Kpelle	KPE
Kru	KRO
Kurdish	KUR
Kurukh	KRU
Ladin see Romansh	
Ladino	LAD
Lahnda	LAH
Lallans see Germanic (Other)	
Lamba	LAM
Landsmaal see Norwegian	

Language	Code	Language	Code
Languedoc see Provençal		Mayan	MYN
Laotian	LAO	Mbundu	UMB
Lapp	LAP	Mende	MEN
Latin	LAT	Micmac	MIC
Latvian	LAV	Middle English see English, Middle	
Lettish, see Latvian		Middle French see French, Middle	
Lithuanian	LIT	Middle High German see German, Middle High	
Lolo	LOL	Middle Persian see Pahlavi	
Lowland Scots see Germanic (Other)		Middle Scots see Germanic (Other)	
Luba	LUB	Milanese see Italian Germanic (Other)	
Luganda	LUG	Miscellaneous	MIS
Luiseno	LUI	Modern Hebrew see Hebrew	
Macedonian	MAC	Mohawk	MOH
Madagascan see Malagasy		Moldavian	MOL
Magyar see Hungarian		Mole see Mossi	
Malagasy	MLA	Mongo see Lolo	
Malay	MAY	Mongol	MON
Malayalam	MAL	Mongolian see Mongol	
Malayo-Polynesian (Other)	MAP	More see Mossi	
Maltese	MLT	Mossi	MOS
Mandingo	MAN	Multilingual	MUL
Manobo	MNO	Muskogee	MUS
Manx see Celtic Group		Nahuatl	NAH
Maori	MAO	Nandi see Sub-Saharan African (Other)	
Marathi	MAR	Navaho	NAV
Masai	MAS	Nepali	NEP
Mashona see Shona		Netherlandic see Dutch	

Language	Code
Newari	NEW
Nez Perce see North American Indian (Other)	
Nguna see Malayo-Polynesian (Other)	
Niger-Congo (Other)	NIC
North American Indian (Other)	NAI
Northern Sotho	NSO
Norwegian	NOR
Nubian	NUB
Nyamwezi	NYM
Nyanga see Nyanja	
Nyanja	NYA
Nyoro Group	NYO
Occitan see Provençal	
Ojibwa	OJI
Old Bulgarian see Church Slavic	
Old Church Slavonic see Church Slavic	
Old English see Anglo-Saxon	
Old French see French, Old	
Old High German see German, Old High	
Old Javanese see Malayo-Polynesian (Other)	
Old Persian see Persian, Old	
Old Russian see Slavic (Other)	
Old Swedish see Germanic (Other)	
Oriya	ORI
Osage	OSA
Osmanli see Ottoman Turkish	
Ossetic	OSS
Ostyak see Selkup	
Oto see Otomi	
Otomi	OTO
Ottoman Turkish (Arabic Script)	OTA
Pahari	PAH
Pahlavi	PAL
Pali	PLI
Panjabi	PAN
Panjabi (Western) see Lahnda	
Papuan-Australian (Other)	PAA
Pasato see Pushto	
Pehlevi see Pahlavi	
Pennsylvania German see German	
Persian, Middle see Pahlavi	
Persian, Modern	PER
Persian, Old (ca. 600 B.C. - 400 B.C.)	PEO
Pidgin English see Creoles and Pidgins	
Pilipino see Tagalog	
Polish	POL
Polyglot see Multilingual	
Portuguese	POR
Prakrit	PRA
Provençal	PRO
Punjabi see Panjabi	
Pusato	PUS
Quechua	QUE

Language	Code
Rajasthani	RAJ
Rhaeto-Romance see Romansh	
Riksmaal see Norwegian	
Romance (Other)	ROA
Romanian	RUM
Romansh	ROH
Romany	ROM
Rumanian see Romanian	
Rumansh see Romansh	
Rundi	RUN
Russian	RUS
Russian, Old see Slavic (Other)	
Saka see Khotanese	
Samaritan	SAM
Samoyed see Selkup	
Sandawe	SAD
Sango	SAG
Sanskrit	SAN
Scots Gaelic see Gaelic (Other)	
SeChuana see Tswana	
Selkup	SEL
Semitic (Other)	SEM
Sephardic see Ladino	
Serbian see Serbo-Croatian (Cyrillic)	
Serbo-Croatian (Cyrillic)	SCC
Serbo-Croatian (Roman)	SCR
Serer	SRR
SeSotho Group see Southern Sotho	
Sesuto see Southern Sotho	
Shan	SHN
Shona	SHO
Siamese see Thai	
Sidamo	SID
Sindhi	SND
Singhalese	SNH
Sino-Tibetan (Other)	SIT
Slavic (Other)	SLA
Slovak (Other)	SLO
Slovene	SLV
Sogdian	SOG
Somali	SOM
Songhai	SON
Sorbian languages see Wendic	
Sorbic see Wendic	
Sotho, Northern see Northern Sotho	
Sotho, Southern see Southern Sotho	
South American Indian (Other)	SAI
Southern Sotho	SSO
Spanish	SPA
Sub-Saharan African (Other)	SSA
Sudanic Group see Niger-Congo (Other)	
Sukuma	SUK
Sumerian	SUX
Sundanese see Malayo-Polynesian (Other)	

Language	Code
Sur-Silvan see Romansh	
Susian see Elamite	
Susu	SUS
Swahili	SWA
Swedish	SWE
Swedish, Old see Germanic (Other)	
Syriac	SYR
Tadzhik see Tajik	
Tagalog	TAG
Tai see Thai	
Tajik	TAJ
Tamil	TAM
Tatar	TAR
Tchetchen see Chechen	
Telugu	TEL
Temne	TEM
Tereno	TER
Thai	THA
Tibetan	TIB
Tigre	TIG
Tigrinya	TIR
Tongan see Malayo-Polynesian (Other)	
Tsimshian	TSI
Tswana	TSW
Turkish	TUR
Turkmen	TUK
Turko-Tataric (Other)	TUT

Language	Code
Twi	TWI
Ugaritic	UGA
Uigur	UIG
Ukrainian	UKR
Umbundu see Mbundu	
Undetermined	UND
Urdu	URD
Uzbek	UZB
Vietnamese	VIE
Vote see Votish	
Votian see Votish	
Votic see Votish	
Votish	VOT
Walamo	WAL
Washo	WAS
Welsh	WEL
Wendic	WEN
Wendish see Wendic	
Wolof	WOL
Xhosa	XHO
Xosa see Xhosa	
Yao	YAO
Yiddish	YID
Yoruba	YOR
Zapotec	ZAP
Zenaga	ZEN
Zulu	ZUL
Zuni	ZUN

Appendix C: Transliteration schemes

The policy of the Reference Manual is to adopt ISO standards where available and suitable.

For transliteration of Cyrillic characters, the Reference Manual recommends ISO/R9: International system for the transliteration of Slavic Cyrillic characters. However, since this contains a number of alternatives, one of these alternatives which avoids the use of diacritics has been incorporated in the scheme with a few additional alterations, in section C.1. This is referred to throughout the Reference Manual as the UNISIST recommended transliteration schedules.

Section C.2 covers transcription of languages using the Roman alphabet with diacritics.

In both cases the objective is to represent the required character set within the limitations of a basic Roman alphabet, comprising letters a-z, without diacritics, so that it can readily be processed in machine-readable form by those systems that do not have diacritics available. Avoiding diacritics has meant that the ability to convert back unambiguously from the transliterated form to the original alphabet has been lost.

C.1 Transliteration of Cyrillic characters

These tables give a full non-reversible transliteration scheme for the Cyrillic alphabet and its variants.

Letter number	Cyrillic char.				Used in						Proposed UNISIST transliter.
	printed		manuscript		Russian	Ukrainian	Belorus.	Serbian	Macedon.	Bulgar.	
1	а	А			X	X	X	X	X	X	a
2	б	Б			X	X	X	X	X	X	b
3	в	В			X	X	X	X	X	X	v
4	г	Г			X	X	X	X	X	X)	g /
5	ґ	Ґ				X					gh
6	д	Д			X	X	X	X	X	X	d

Letter number	Cyrillic char. printed		Cyrillic char. manuscript		Used in Russian	Ukrain.	Belorus.	Serbian	Macedo.	Bulgar.	Proposed UNISIST translit.
7	ђ	Ђ						X			dj
8	ѓ	Ѓ							X		g
9	е(ё)	Е(Ё)	е (ё)	Є (Ё)	X	X	X	X	X	X	e
10	є	Є				X					je
11	ж	Ж			X	X	X	X	X	X	zh
12	з	З			X	X	X	X	X	X	z
13	ѕ	Ѕ							X		dz
14	и	И			X	X		X	X	X	i
15	і	І			X	X	X				yi
16	ї	Ї				X					yi
17	ј	Ј						X	X		j
18	й	Й			X	X	X			X	j
19	к	К			X	X	X	X	X	X	k
20	л	Л			X	X	X	X	X	X	l
21	љ	Љ						X	X		lj
22	м	М			X	X	X	X	X	X	m
23	н	Н			X	X	X	X	X	X	n
24	њ	Њ						X	X		nj
25	о	О			X	X	X	X	X	X	o
26	п	П			X	X	X	X	X	X	p
27	р	Р			X	X	X	X	X	X	r
28	с	С			X	X	X	X	X	X	s
29	т	Т	т, т, т		X	X	X	X	X	X	t

Letter number	Cyrillic char. printed		manuscript		Used in Russian	Ukrain.	Belorus.	Serbian	Macedo.	Bulgar.	Proposed UNISIST translit.
30	ѣ	ѣ						X			cj
31	ќ	ќ							X		k
32	у	у			X	X	X	X	X	X	u
33	ў	ў					X				w
34	ф	ф			X	X	X	X	X	X	f
35	х	х			X	X	X	X	X	X	kh
36	ц	ц			X	X	X	X	X	X	ts
37	ч	ч			X	X	X	X	X	X	ch
38	џ	џ						X	X		dzh
39	ш	ш	ш, ш		X	X	X	X	X	X	sh
40	щ	щ			X	X				X	shch
41	ъ	ъ			X	X	X			X	"
42	ы	ы			X		X				y
43	ь	ь			X	X	X			X	'
44	ѣ	ѣ			X		X			X	'
45	э	э			X		X				eh
46	ю	ю			X	X	X			X	yu
47	я	я			X	X	X			X	ya
48	ѧ	ѧ								X	'

C.2 Transciption of diacritics in languages using the Roman alphabet

The following digraphs should be used if it is desired to avoid diacritics.

German	ä	to be represented by	ae
	ö	" " " "	oe
	ü	" " " "	ue
Scandinavian languages	å	to be represented by	aa
	ø	" " " "	oe
	æ	" " " "	ae

In French, Spanish and Italian diacritics may be ignored.

Appendix D: Elements in a Personal Name

Because of its presumed marginal usefulness in the environment of description of research projects and institutions, no attempt has been made to rewrite Appendix D which has been copied from Appendix E in the Reference Manual for Machine-Readable Bibliographic Descriptions, second revised edition, 1981. Many users of the present Manual may have no need for any or most of these detailed specifications. Those who do may find them useful in spite of the fact that the text is geared towards bibiliographic applications.

Appendix D: Elements in a Personal Name

The conventions described in this appendix apply equally to any of subfields 1 to 6 in personal name fields, except as otherwise noted.

The elements in an individual name may be defined as follows:

'Key' name or names	'K'
Forename and/or initials	'F'
Suffix	'S'
Title	'T'

All names are to be entered in the following form:

K,ØF,ØSØ(T)

Commas are used to separate the 'key' names (surnames) from the forename and/or initials, and to separate the forenames from any suffix (such as 'Jr', 'III'). A title, if required, is entered in parentheses at the end of the name. For example:

'Rutherford (Lord)'
'Rutherford, James D.,Jr'
'Rutherford, J.D.'

Surname prefixes are considered to be part of the key name.

'Key' names

The 'key name' element (K) corresponds to the surname or family name in a Western name. The term 'key name' is used rather than 'surname', however, since there may be occasions when it is not clear that the content of this element really represents a surname in the Western sense. (Also, it is envisaged that there may be an exact correspondence between 'K' elements and entry points, or 'keys', in a printed author index). There may be more than one 'K' element if the surname is a compound one (e.g. 'Martinez Moreno'), or in the case of certain oriental names where there is real doubt about which component is the surname.

The 'K' element is always an essential element except in some names consisting only of a religious title and forename(s) (e.g. 'Sister Mary Hilda').

Forename and/or initials

The 'F' element is an essential element unless the fullest available form of the name comprises only a surname and a title or unless all components of the name are treated as a key name.

If one or more forenames are given in full, the first (or second if the individual is generally known by the second forename) may be retained and all others reduced to the initial(s). Initials should each be followed by a full stop.

If the fullest form of the name on the original gives only initials for the forenames, the first forename may be entered as an initial, or may be spelled out in full if this information is readily and unambiguously available from existing reference works (previous indexes, directories, biographical dictionaries, etc).

If a forename appears in abbreviated form (e.g. 'Chr.', 'Jas'), the abbreviation may be retained and entered in the 'F' element:

Example 1

Authorship as shown on the document:

"DR F. GROSS und TH. BECK"

Contents of repeated personal name fields:

First author: Ø1@1Gross,ØF.

Second author: Ø1@1Beck,ØTh.

If a hyphenated forename is reduced to initials, the initial letters of both parts are to be retained, linked by a hyphen (e.g. 'Jean-Paul' gives 'J.-P.').

Suffixes

The 'S' element is used to enter "suffixes" such as 'Jr', 'II', etc. Any such suffix is to be retained as an essential element. Some examples are given below:

English:	Jr, Sr, II, III
Spanish:	hijo, nieto
Portuguese:	filho, neto, sobrinho
Hungarian:	ifj., id.
Russian:	ml.

Example 2

Authorship as shown on the document:

"BY F.S. HARRIS, JR, The Aerospace Corporation, P.O. Box 95085, Los Angeles, Calif. 90045"

Contents of personal name field:

Ø1@1Harris,ØF.S.,ØJr

Suffixes representing titles, or professional or academic qualifications, are not normally entered.

Titles and qualifications

The 'T' element may be used in a few special circumstances to enter a title which forms part of a person's name. In general, however, titles are omitted from names entered in bibliographic descriptions. Detailed rules are suggested as follows:

Academic, professional, religious or military titles preceding the name (such as 'Dr', 'Ing.', 'Rev.', 'General', etc.), and titles or qualifications following the name, are omitted from bibliographic descriptions:

Example 3

Authorship as shown on the document:

"Ing. STEFANIA BAICU"

Contents of personal name field:

Ø1@1Baicu,ØStefania
or
Ø1@1Baicu,ØS.

'Mr', 'M.', 'Mrs', 'Mme', and their equivalents in other languages are normally omitted. 'Mrs', 'Mme', etc., may be retained for married women authors when only the husband's forenames or initials are given in the original, e.g. 'Mrs John J. Doe':

Example 4

Authorship as shown on the document:

"Note de MM. JEAN-MARC DESRUMAUX, JEAN-MICHEL ROUVAEN et Mme CLAUDE MORIAMEZ, présentée par M. René Lucas"

Contents of repeated personal name fields:

First author: Ø1@1Desrumaux,ØJean-Marc
or
Ø1@1Desrumaux,ØJ.-M.

Second author: Ø1@1Rouvaen,ØJean-Michel
or
Ø1@1Rouvaen,ØJ.-M.

Third author: Ø1@1Moriamez,ØClaudeØ(Mme)
or
Ø1@1Moriamez,ØC.Ø(Mme)

The title 'Mme' is included since the name given is that of the husband (but this particular example could be ambiguous: 'Claude' in French is both masculine and feminine). Note also the contraction of hyphenated forenames, and the fact that the person cited as 'presenting' the paper is not included as an author.

'Miss', 'Mlle', 'Ms' and their equivalents in other languages are omitted unless only the surname is given:

Example 5

Authorship as shown on the document:

"Note de Mlle EDITH DEVIN et M. ROBERT LOCQUENEUX, présentée par M. Louis de Broglie"

Contents of repeated personal name fields:

First author: Ø1@1Devin,ØEdith
or
Ø1@1Devin,ØE.

Second author: Ø1@1Locqueneux,ØRobert
or
Ø1@1Locqueneux,ØR.

Terms which indicate affiliation with religious orders (e.g. Sister, Brother) are not retained unless only the forename(s) are given:

Example 6

Authorship as given on the document:

"Sister Helen Therese Nyberg, O.P."

Contents of personal name field:

Ø1@1Nyberg,ØHelenØT.
or
Ø1@1Nyberg,ØH.T.

Honorific titles are normally omitted, but may be retained if they constitute an indispensable part of the name:

Example 7

Authorship as given on the document:

"LORD TODD"

Contents of personal name field:

Ø1@1ToddØ(Lord)

Surname prefixes

All surname prefixes are retained in personal author names. A prefix and the name to which it is affixed are together regarded as forming a single 'key' name. Examples of frequently used prefixes are:

van	la	lo	van der
von	della	du	vander
de	le	des	van de
da	del	de la	

See note below on "Special symbols used in personal names", and examples. Note also that the initial letters of these prefixes may appear in both upper and lower case.

Compound surnames

Compound surnames are the rule for most Spanish and Portuguese authors, and are occasionally found among almost all nationalities.

If the surname is a compound containing a hyphen, e.g. 'Litvak-Gorskaya', the whole compound name should be entered as a single 'key' name.

If it is apparent that the surname is a compound which is not hyphenated, both names should be entered as 'key' names, e.g. 'J. Hunter Dunn'. If in doubt, enter only the final element as a 'key' name and treat the first element as a forename.

Names that indicate marital status

In certain languages a married woman author's name is the same as her husband's with the addition of one or more letters, or a different word-ending. For example, in Hungarian the suffix 'ne' may be applied to either a forename or a surname. Names of this kind should be entered exactly as they appear on the document without modification, and in accordance with the rules previously defined:

Example 8

Authorship as shown on the document:

"GYORGY KAROLYNE, Dr"

Contents of personal name field:

Ø1@1Karolyne,ÞGyorgy

or

Ø1@1Karolyne,ÞG.

(Note also omission of academic title)

Names where 'forename' and 'surname' are not readily identifiable

In practice, particularly with oriental names, there may be many cases where it is not possible to determine with assurance which of two or three names is really the surname. In this event, it is recommended that two or more elements may be treated as 'key' names, entered in the sequence given on the document, and used to generate cross-references in author indexes, if desired.

Examples: 'Teh Fu Yen', 'Krishna Mohana Rao' (or 'Mohana Rao, Krishna')

Guidelines for the treatment of such names are given in: Names of persons, national usages for entry in catalogues, compiled by the IFLA International Office for UBC. Third ed. London, IFLA International Office for UBC, 1977.

Spelling

Individual author names are to be entered in the vernacular, as they appear on the original document, except:

(a) If transliteration from a non-Roman alphabet to Roman alphabet is required, ISO, or by default, UNISIST transliteration schedules recommended in Appendix C are to be used.

(b) If an 'established form' of the name is known to the originator of the bibliographic description, and if this form differs from what has been derived from the original, then the 'established form' may be entered in subfield 2.

This is particularly likely to arise where a non-Russian name is transliterated into Cyrillic for publication in a Russian journal, and is subsequently re-transliterated to the Roman alphabet.

In all cases, the name as given on the document (transliterated if necessary) should be regarded as the primary form for entry in a bibliographic description, since the use of the 'established form' depends on prior knowledge which may not be accessible to all users of a bibliographic data base. Subfield 1 should always carry the name as derived from the document.

Example 9

Authorship as shown on the document:

"St. BOYADJIEW"

In this case a known alternative (and preferred) transliteration exists:

"Boyadzhiev"

Contents of personal name field:

Ø1@1Boyadjiew,ØSt.@2Boyadzhiev,ØSt.

Special symbols used in personal names

Two special symbols may occur in personal names as entered in accordance with UNISIST recommendations. They have been introduced in order to make it possible to deal with certain problems which arise in the production of author indexes and other listings when an author has a complex surname or one which includes prefixes or abbreviations. Their use is in no way mandatory, but they have been defined in such a way that it will be possible for services which exchange bibliographic records to leave all options open for the recipient of an exchange tape to apply whatever conventions he may wish in deriving author indexes from the machine file.

The two symbols are '=' and '+'. Both are to be regarded as 'space' for purposes of display and search matching.

= The connective '=' is intended to be used to link a prefix to the name to which it is affixed and to indicate that the following character is the beginning of a 'strong' component of the name, i.e. one which may (depending on the policy of the individual service) be used as a key for creating an index entry or cross-reference.

Examples: 'Teilhard de=Chardin'
'von=Dorrien'

\+ The connective '+' is intended to be used to link components of a compound surname and to indicate that the following character is the beginning of a 'weak' component of the name, i.e. one which should never be used as a key for creating an index entry or cross-reference.

Examples: 'Gonzales+G.,R.'
'Asin+y Cabrera, M.D.'
'van+der=Avoird,A.'

The remaining examples illustrate the various possibilities which arise when dealing with compound names, and names involving prefixes. Although the Reference Manual leaves certain options open, it would be expected that any individual service, or the parties to an exchange of bibliographic data, would adopt a single coherent policy across the whole of their data base.

Example 10

Authorship as shown on the document:

"AD VAN DER AVOIRD"

Contents of personal name field:

Ø1@1van+der=Avoird,ÞAd
or
Ø1@1vanÞderÞAvoird,ÞAd
or
Ø1@1van+der=Avoird,ÞA.
or
Ø1@1vanÞderÞAvoird,ÞA.

Example 11

Authorship as shown on the document:

"Note de MM. MICHEL BRUNEL et FRANCOIS DE BERGEVIN, transmise par M. Louis Néel".

Contents of repeated personal name fields:

First author: Ø1@1Brunel,ÞMichel

Second author: Ø1@1de=Bergevin,ÞFrancois

or

Ø1@1deÞBergevin,ÞFrancois

Alternatively:

First author: Ø1@1Brunel,ÞM.

Second author: Ø1@1de=Bergevin,ÞF.

or

Ø1@1deÞBergevin,ÞF.

Example 12

Authorship as shown on the document:

"DEREK J. DE SOLLA PRICE"

Contents of repeated personal name fields:

Ø1@1de=SollaÞPrice,ÞDerekÞJ.

or

Ø1@1deÞSollaÞPrice,ÞDerekÞJ.

or

Ø1@1de=SollaÞPrice,ÞD.J.

or

Ø1@1deÞSollaÞPrice,ÞD.J.

Example 13

Authorship as shown on the document:

"LUIS RIVERA OYOLA and R.A. LEE"

Contents of repeated personal name fields:

First author: Ø1@1RiveraÞOyola,ÞLuis

or

Ø1@1RiveraÞOyola,ÞL.

Second author: Ø1@1Lee,ÞR.A.

BIBLIOGRAPHY

As this Manual is in a loose-leaf updatable form, it has been decided to make references to the relevant items in the text reasonably complete for identification purposes. Items referred to are additionally collected together here with a fuller bibliographic description.

Dewey, M.	Dewey Decimal Classification and Relative Index. 19th ed. New York, Forest Press, 1978. 3 vols.
Dierickx, H.	Development of a Manual for Standard "Bibliographic" Description of Research Projects - A Study on Recommended Action by Harold Dierickx and Alan Hopkinson. London, UNIBID, December 1978. 16p. (UNIBID/RM/R/78/01; Unesco Contract No. 695074).
Dierickx, H.	Reference Manual for Machine-Readable Bibliographic Descriptions, Second revised edition. Compiled and edited by Harold Dierickx and Alan Hopkinson. Paris, Unesco, 1981, vp. (PGI-81/WS/22).
ECSSID	ECSSID WG2. Workshop June 19-20, 1980 in Bonn. Aide mémoire. 2p.
Fédération Internationale de Documentation (FID)	B.S.O.: Broad System of Ordering Schedule and Index, 3rd rev. The Hague, F.I.D.; Paris, Unesco, 1978. xiv, 102, 82p. ISBN 92-66-00564-9.
Fédération Internationale de Documentation (FID)	Universal Decimal Classification. 1907-. Note: Versions and editions exist in different languages and are maintained and published in many cases by national standards organizations.
H.M.S.O.	Post Office guide. Current edition and supplements. London (UK), H.M.S.O., vp.
International Federation of Library Associations and Institutions	Names of persons: national usages for entry in catalogues, 3rd ed. London, IFLA International Office for UBC, 1977. x, 193p. ISBN 0-903043-10-6.
International Organization for Standardization	See separate section at the end of the bibliography.

OECD/EADI — Expert Meeting on the Standardization of Registers for Development Information and Documentation. Paris, December 3-4, 1979. 19p.

Prince-Perciballi, I. — AGRIS Classification Scheme. Rome (Italy), FAO (FAO-AGRIS-3). (AGROVOC, a thesaurus of agricultural terms, has been published by FAO in 1981).

SAELEN, K.T. — An Evaluation of the UNISIST Reference Manual for Machine-Readable Description of Research Projects and Institutions. Bergen, August 1980. 6p. (Paper prepared for Working Group 2 of ECSSID: European Cooperation in Social Science Information and Documentation).

Times Newspapers — The Times atlas of the world. [Latest] comprepensive edition. London (UK), Times Newspapers. 197?-.

UNESCO — Meeting of Experts on Information on Research in Progress. Paris, Unesco House, Room XII, 1-4 September 1980. Paris, 27 May 1981. 20p. (PGI-80/CONF.611.5).

UNESCO/EADI — Technical Meeting on an Exchange Format for Information on Research in Progress. Room III, Unesco House, Paris, 9 and 10 June 1980. Minutes. Paris, November 1980, 12p. (PGI/ORI/80/1.4).

International Standards

International standards are published by the International Organization for Standardization (ISO), Geneva. They are generally available from the member bodies of ISO, i.e. the national standards organizations in each country, or direct from ISO Central Secretariat. Standards marked with an asterisk are reproduced in the following handbook available from either the member bodies, the ISO Central Secretariat, or Unesco, Division of the General Information Programme:

Information transfer: handbook on international standards governing information transfer, 1st ed., Geneva, ISO, 1977. (ISO Standards Handbook, 1).

*ISO R9 - 1968	: International system for the transliteration of Slavic Cyrillic characters. 8p.
*ISO 646 - 1975	: 7-bit coded character set for information processing interchange. 12p.
*ISO 962 - 1974	: Information processing - Implementation of the 7-bit coded character set and its 7-bit and 8-bit extensions on 9-track 12.7mm (1/2 inch) magnetic tape. 3p.
ISO 1001 - 1979	: Magnetic tape labelling and file structure for information interchange. 29p.
ISO 2014 - 1976	: Writing of calendar dates in all-numeric form. 1p.
*ISO 2022 - 1973	: Code extension techniques for use with the ISO 7-bit coded character set. 23p.
ISO 2382 - 1974-	: Data processing vocabulary (being issued in 19 parts).
ISO 2709 - 1981	: Documentation - Format for bibliographic information interchange on magnetic tape. 4p.
*ISO 3166 - 1981	: Codes for the representation of names of countries. 22p. Amendments issued frequently.

National standard

GOST 13052 - 74 : Computers and data transmission equipment - 7 bit codes for information interchange. Moscow, GOST, 1974 (in Russian). Available through national standards organizations.

Addresses

The following addresses may be useful for obtaining items in the bibliography:

FID
P O Box 90462
2509 LK The Hague
The Netherlands

International Centre for the Registration of Serials
20 Rue Bachaumont
75002 Paris
France

IFLA International Office for UBC
c/o British Library Reference Division
Great Russell Street
London WC1B 3DG
UK

ISO Central Secretariat
Case Postale 56
1211 Geneva 20
Switzerland

Unesco
Division of the General Information Programme
7 Place de Fontenoy
75007 Paris
France