

170-WP-016-001

Catalogue Interoperability Protocol (CIP) - Technical Note on Guide

White Paper

White Paper--Not intended for formal review or
government approval.

August 1997

Prepared Under Contract NAS5-60000

RESPONSIBLE ENGINEER

<u>Janet F. Hylton /s/</u>	<u>8/15/97</u>
Janet F. Hylton	Date
EOSDIS Core System Project	

SUBMITTED BY

<u>George Percivall /s/</u>	<u>8/22/97</u>
George Percivall	Date
EOSDIS Core System Project	

Hughes Information Technology Systems
Upper Marlboro, Maryland

This page intentionally left blank.

Contents

1. Introduction

1.1 Purpose.....	1-1
1.2 Organization.....	1-1
1.3 Review and Approval.....	1-1

2. Motivations

3. Definition of Guide

3.1 Guide Data.....	3-1
3.2 Guide Services.....	3-1

4. Proposed Solutions

4.1 Guide Collection Structure.....	4-1
4.2 BIB-1	4-1
4.3 Guide Descriptor	4-1

5. Suggested Approach to Support Data and Services

5.1 Objects.....	5-1
5.2 Attributes(ARS for Guide Descriptor Object).....	5-1
5.3 Descriptor Relationships	5-3
5.4 Guide Services.....	5-3
5.4.1 Guide and Collection Relationship Model	5-3
5.4.2 Collection and Guide Search Services	5-4
5.4.3 Guide Descriptor Search Services.....	5-4
5.4.4 Guide Document Search Services.....	5-6
5.4.5 Present Services.....	5-6

6. Scenario

7. Recommendations

7.1 Summary of Changes	7-1
7.2 Impact on Existing Documentation.....	7-1
7.2.1 URD.....	7-1
7.2.2 CIP Specification.....	7-1
7.2.3 SDD.....	7-2

List of Figures

5-1. Guide and Collection Relationship Model.....	5-4
6-1. Guide Search Scenario.....	6-1

List of Tables

5-1. Guide Descriptor Attributes	5-1
5-2. Guide Descriptor Use Attributes	5-5

Appendix A.

1. Introduction

1.1 Purpose

The purpose of this paper is to present the recommended approach for incorporating Guide Data within the CIP framework. To accomplish this goal several alternatives are presented followed by a detailed explanation of the recommended alternative.

1.2 Organization

This paper describes the recommended approach for supporting Guide information within ICS, to include metadata and documents. The paper begins with a brief discussion of the motivating factors that forced an examination of the existing approach as specified in the Release B Specification. Section three provides a definition of guide data and services. Four presents the various proposed solutions that were examined during this analysis. Five identifies and describes a suggested approach to support the data and services identified in section three. Section six presents an operational scenario for searching Guide Descriptor Data. Section seven identifies the recommended changes to the existing CIP documentation. Lastly, section eight captures the current guide requirements as stated in the URD, the estimated number of guide documents that the ICS is expected to describe and a detailed explanation of a proposed solution.

1.3 Review and Approval

This document was prepared as part of the ECS Contract between NASA-GSFC-ESDIS and Hughes as described in ECS Engineering Support Directive #25, ECS Extensions Support.

Questions regarding technical information contained within this Paper should be addressed to the following ECS and/or GSFC contacts:

- ECS Contacts
 - Janet F. Hylton, (301) 925-0466, janet@eos.hitc.com
- GSFC Contacts
 - Yonsook Enloe, (301) 286-0794, yonsook.enloe@gsfc.nasa.gov

Questions concerning distribution or control of this document should be addressed to:

Data Management Office
The ECS Project Office
Hughes Information Technology Systems
1616 McCormick Drive
Upper Marlboro, Maryland 20774-5372

This page intentionally left blank.

2. Motivations

The Release B Specification employs the use of a Guide Collection Structure which in theory mirrors the Product Collection Structure. This structure offers the EO Product data the facility to aggregate potentially large volumes of information into logical EO collections to support quicker access to both local and remote information . Unlike Product information the guide is not burdened with large volumes of data (reference Appendix 8), that may require aggregation to support quicker access. Furthermore, after careful examination of existing Guide Systems coupled with numerous discussions with the EO Guide Author Community it was thought that this approach should be re-examined by the PTT.

Therefore, this Technical Note (TN) begins with the identification and definition of Guide Data and Services. It then explores the various proposed solution that have been suggested to the PTT Core Engineering Team. Of the proposed solutions, this TN, then develops a single solution, the recommended solution, by providing technical information from which the readers of this TN can adequately evaluate the recommendation. The technical information explores the data, services, and integration with the existing collections design. The TN concludes with recommended changes to the various CIP Documents.

This page intentionally left blank.

3. Definition of Guide

3.1 Guide Data

Guides are textual documents that are developed to assist the user in understanding the Earth Observation environment. Guide Documents provide additional detailed information about a specific mission, sensor, or organization. The most frequently covered guide topics include:

Data Centers - provides details about the center that holds the data.

Project/Campaign - describes projects and associated intensive field campaigns coordinated to collect data for a focused study.

Source/Platform - describes the source that held the instrument and sensors during data collection. Source is intended to communicate the data collection environment which includes satellite, aircraft, buoys, ships, and ground station platforms. It also includes humans in the case of hand-held sensors or human observations, paper or electronic surveys in the case of questionnaires, and computers in model analysis.

Sensor/Instrument - describes the instrument and its component sensor (s) that actually collected the data. This includes the human eye as a sensor, as well as, paper or electronic questionnaires used to collect data. It also includes computer model analyses that generate data.

Dataset - describes the dataset, collection procedures, algorithms, and processing data.

In addition to Guide Documents there are Guide Metadata which serves to characterize the related document. This metadata is used by the system to identify potential documents of interest. Each Guide Document in ICS will be characterized in the form of Guide Metadata. It is both the Guide Document and the Guide Metadata that are the focus of this Technical Note.

3.2 Guide Services

Guide services include both search and present services of the metadata and document content. A guide service can be characterized by the following:

- the guide data can be searched through pre-defined metadata
- guide data can also be searched by free text searches.
- the guide data can be presented as metadata or as a textual document.
- Guide Documents can be retrieved through a known URL

This page intentionally left blank.

4. Proposed Solutions

In support of the data and services identified above the following solutions were proposed to the PTT Core Engineering Team.

4.1 Guide Collection Structure

The details regarding this recommendation can be found in Appendix 8 of this document. The following addresses two key points which caused concern for adopting this approach.

Organizing guide data into collection hierarchies is currently not supported within the EO Science Data Provider Community.

The overhead (approximately 20 required collection attributes) coupled with the complexity associated with constructing and maintaining a collection structure for Guide Documents did not suggest that the collection structure was an adequate solution for the Guide Data.

4.2 BIB-1

The BIB-1 approach was recommended to the PTT Core Engineering Team as a potential solution for the Guide. Although this solution is consistent with traditional approaches for providing a standard mechanism for accessing and retrieving documentary information it is believed that it will not satisfy the requirements. Both NASA and ESA have identified a need to establish a bi-directional association between Guides and Collections. With the BIB-1 solution it is not possible to enforce an association between local guide information and a CIP Collection. Also, the BIB approach is not specific to the Earth Observation Data.

4.3 Guide Descriptor

The preferred approach is the Guide Descriptor. This approach, which is described in detail in the following sections will not impose an artificial aggregation of information. It also supports bi-directional associations. Additionally, the CIP Guide Descriptor includes attributes which are specific to Earth Observation data such as Temporal Keyword and Spatial Keyword

This page intentionally left blank.

5. Suggested Approach to Support Data and Services

The following presents an approach for incorporating the Guide Data and Services within ICS. The approach recommends the Guide Descriptor solution. The Guide Data are described in terms of an object, attributes and relationships. The services are described as search and present.

5.1 Objects

The guide metadata is described using the Guide Descriptor data object. Guide Descriptors serve as a framework for capturing the metadata associated with each Guide Document. This framework is then used by the system to identify the data that are available to the prescribed services.

5.2 Attributes(ARS for Guide Descriptor Object)

The Guide Descriptor data object contains attributes which serve to provide, in a consistent fashion, the information specific to each metadata entry for a Guide Document. For example, it is required that each Guide Document metadata entry (Guide Descriptor), will contain an abstract attribute and value. The following table identifies the Guide Descriptor attributes and attribute characteristics. An “X” in the column indicates that the attribute specified in col 1 possesses the characteristic. An attempt was also made to map the CIP Guide Attributes to the current BIB-1 Attributes, the results of which are expressed in the last column.

Table 5-1. Guide Descriptor Attributes (1 of 2)

Guide Descriptor Attributes	Use	Mandatory	Compound	BIB-1
Item Descriptor ID	X	X		Local Control Number (12)
Item Descriptor Name	X	X		Title (4)
(Purpose)				
Abstract	X	X		Abstract (62)
(Version ID)				
(External Citation Publication)				
Author	X	X	X	Author (1003
Person Name	X	X		Author Name Personal (1004)
(Job Position)				
Organization Name	X	X		Author Name Corporate (1005)
(Address)			X	
Street Address				

Table 5-1. Guide Descriptor Attributes (2 of 2)

Guide Descriptor Attributes	Use	Mandatory	Compound	BIB-1
City				
State				
Postal Code				
Country				
Email Address				
Telephone Number				
Fax Number				
Guide			X	
Document Type	X	X		
Document Variant	X	X	X	
Language	X	X		Code Language (54)
Document Name	X	X		Title (4)
(Publication)	X	X	X	
Publication Date	X	X		Date of Publication (31)
Publication Place				Place of Publication (59)
0(Document Service Options)n			X	
Group ID	X			
ItemWord Size				
0(Document Delivery Options)n			X	
ItemByteSize				
DocumentFormat				
DocumentCompression				
Document Data	X	X	X	
[ItemData				
Document Pointer]				
Keywords			X	
(Spatial Keywords)	X			
(Temporal Keywords)	X			
General Keywords	X	X		
Revision			X	
Creation Date	X	X		Date of Acquisition (32)
(Update Frequency)				
Revision Date	X	X		Date Time Last Modified (1012)
Related Item Descriptor	X		X	
Item Descriptor ID	X			
Relation Description				

5.3 Descriptor Relationships

The relationship between the Guide Descriptors and Collection Objects is established through the “Related Item Descriptor” compound attribute in either the Collection Descriptors or the Guide Descriptors. This relationship provides a non searchable entity association which infers that the Related Item Descriptor’s Item Descriptor ID is returned in the result set rather than the entity itself. This Item Descriptor ID can then be used by the user to target a specific collection or Guide Descriptor. The following relationships are supported:

Guide Descriptor related to Collection

Collection Descriptor related to Guide Descriptor

Guide Descriptor related to Guide Descriptor

Additionally, the Document Pointer allows a relationship to exist between the Guide Descriptor and the Guide Document.

The Item Data allows the instance of the document to be reflected in the metadata.

5.4 Guide Services

The Guide Services allow the Guide Descriptor data to be searched and presented by ICS. The following presents each of the services and a brief description of their recommended capabilities. Additionally, a Guide and Collection Relationship Model which illustrates the relationships between the collection structure and Guide Descriptor is also presented.

5.4.1 Guide and Collection Relationship Model

This model illustrates how the Guide Descriptor concept is integrated within the CIP collections framework. It is important to note that it will be possible to access the guide information without first searching the collection structure and determining the related Guide Descriptors. As can be seen below, the Guide Descriptors may relate to collections or to other Guide Descriptors. The later may be required if it is necessary to relate a Platform Guide to an Instrument Guide. Additionally, collections may relate to local or remote Guide Descriptors and Guide Descriptors may relate to local or remote collections or other Guide Descriptors.

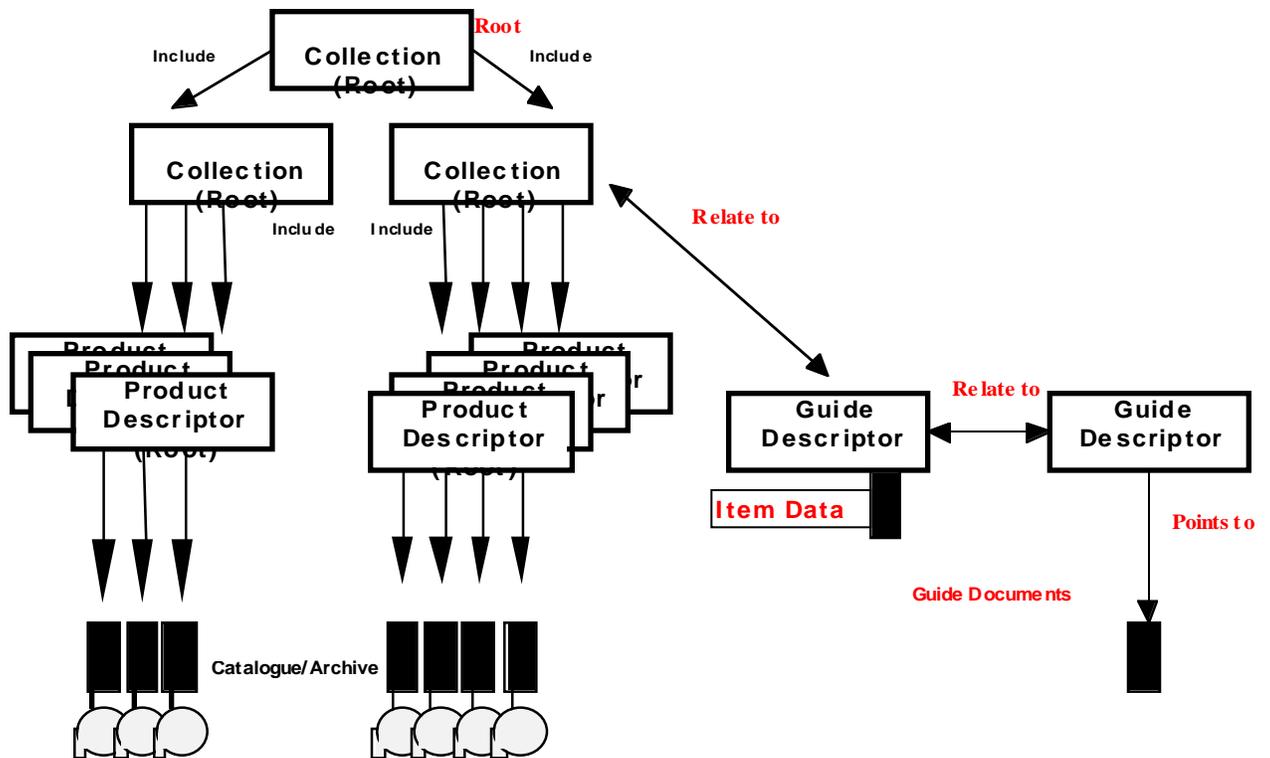


Figure 5-1. Guide and Collection Relationship Model

5.4.2 Collection and Guide Search Services

With this solution it will be possible to relate collections to guides and present the guide related Item Descriptor ID. However, searching Guide Descriptor data or Guide Document data from within a collection search is currently not supported.

5.4.3 Guide Descriptor Search Services

Guide Descriptor Search Services are services which operate on the guide metadata. There are three types of Guide Descriptor services; Attribute Value Pair, Free Text Searches, and Attribute Value Pair and Free Text Search combinations. The later service is a combination of metadata and document. The following discusses each of these services.

Attribute Value Pair(Metadata). This type of search searches the metadata attributes of the Guide Descriptors for a specific value. The selected descriptors' attributes must contain this value as specified in the search query. Any of the use attributes described in the above Guide Descriptor can be searched as attribute value pairs. Typically, the Keyword, Document Type, General Keyword, and, or, Item Descriptor ID attributes will be used most frequently in this type of search request.

Because there are various types of Guide Documents; sensor, platform, etc., coupled with the specific value associated with the document type; i.e. Document Type=Sensor and Sensor=AVHRR; it will be necessary to introduce an additional mandatory attribute to capture the specifics of the document type. **It is recommended, to minimize the impact on the Release B Specification, that the general keyword be used in the Guide Descriptor to reflect this information. Perhaps we should introduce a new attribute....any thoughts??**

Free Text Searches(Metadata). This type of search implies that the value specified in the search query for a specific attribute is contained in the value of the Guide Descriptors' attribute. The query specified value does not need to match the entire contents of the attributes value however the query specified value must be contained within the attributes value. Free Text Searches will most likely occur on the Purpose and Abstract metadata attributes.

Attribute Value Pairs and Free Text(Metadata and Document) This search is a combination of the above. For Example, Document Type=Sensor and **General Keyword**=AVHRR and Free Text (document)= Five channels. Only those Guide Descriptors with Sensor as a Document Type, and AVHRR as a **General Keyword**, will be evaluated further to determine if the free text phrase, "Five channels" is contained in the document.

The following table identifies the Guide Descriptor Use Attributes, their structure and relations that will be supported:

Table 5-2. Guide Descriptor Use Attributes (1 of 2)

Guide Descriptor Use Attributes	Structure	Relations
Item Descriptor ID	bib-1 101 Name (normalized) Int'lString	LT,LTE,EQ,GTE,GT,NE,AM
Item Descriptor Name	bib-1 101 Name (normalized)Int'lString	LT,LTE,EQ,GTE,GT,NE,AM
Abstract	bib-1 108 Int'lString	LT,LTE,EQ,GTE,GT,NE,AM
Author	bib-1 101 Name (normalized)Int'lString	LT,LTE,EQ,GTE,GT,NE,AM
Person Name	bib-1 101 Name (normalized)Int'lString	LT,LTE,EQ,GTE,GT,NE,AM
Organization Name	bib-1 101 Name (normalized)Int'lString	LT,LTE,EQ,GTE,GT,NE,AM
Document Type	bib-1 101 Enum	LT,LTE,EQ,GTE,GT,NE,AM
Document Variant	CIP-6 Compound	EQ,NE,AM

Table 5-2. Guide Descriptor Use Attributes (2 of 2)

Guide Descriptor Use Attributes	Structure	Relations
Language	bib-1 101 String	LT,LTE,EQ,GTE,GT,NE,AM
Document Name	bib-1 101 String	LT,LTE,EQ,GTE,GT,NE,AM
Publication	CIP-6 Compound	EQ,NE,AM
Publication Date	CIP-1 Date	LT,LTE,EQ,GTE,GT,NE,AM
Group ID	bib-1 101 Enum	LT,LTE,EQ,GTE,GT,NE,AM
Spatial Keywords	bib-16 WordList Int'lString	LT,LTE,EQ,GTE,GT,NE,AM
Temporal Keywords	bib-16 WordList Int'lString	LT,LTE,EQ,GTE,GT,NE,AM
General Keywords	bib-16 WordListInt'lString	LT,LTE,EQ,GTE,GT,NE,AM
Creation Date	CIP-1 Date	LT,LTE,EQ,GTE,GT,NE,AM
Revision Date	CIP-1 Date	LT,LTE,EQ,GTE,GT,NE,AM
Related Item Descriptor	CIP-6 Compound	EQ,NE,AM
Item Descriptor ID	bib-1 101 Name (normalized)Int'lString	LT,LTE,EQ,GTE,GT,NE,AM

5.4.4 Guide Document Search Services

Free Text Searches(Document Text). This type of search implies that the value specified in the search query must be contained in the selected Guide Document(s).

5.4.5 Present Services

The present services allows both Guide Metadata, and the Guide Document, to be returned to the requester as a result of the search request. There are three types of present services.; brief, full and the Guide Document.

5.4.5.1 Brief

The user may request that a brief set of elements be returned as a result of the search query in which case the following will be presented:

Item Descriptor ID

Item Descriptor Name

Purpose

Abstract

5.4.5.2 Related Item Descriptor

Item DescriptorID

Relation Description

5.4.5.3 Full

The user may request that a full set of elements be returned as a result of the search query in which case all of the Guide Descriptor attributes will be presented.

5.4.5.4 Document

If the Guide Document is requested to be presented then the textual document will be returned to the requester.

This page intentionally left blank.

6. Scenario

The following graphically illustrates an operational scenario for searching Guide Descriptor data.

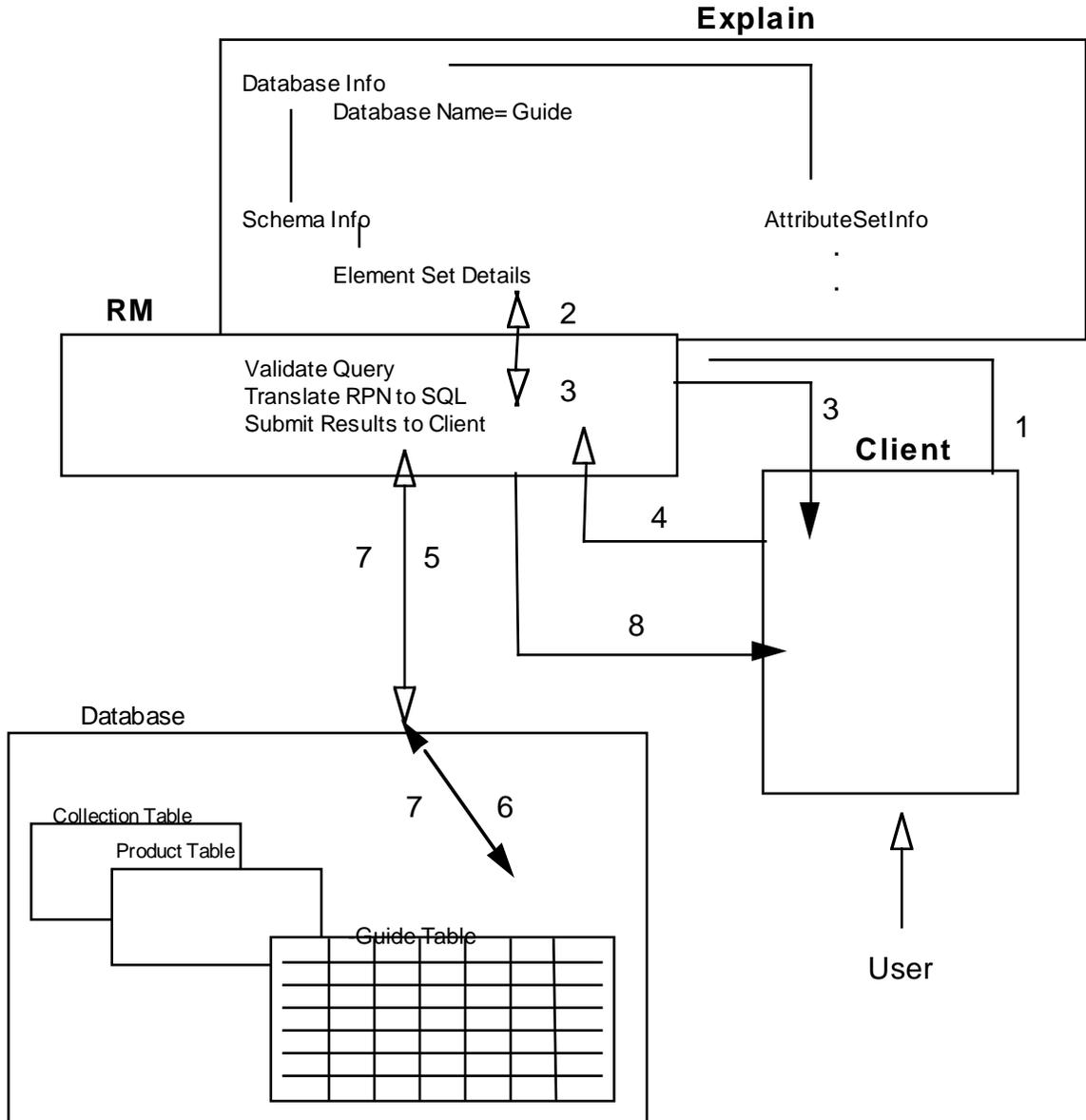


Figure 6-1. Guide Search Scenario

- Step 1 Client requests information from RM
- Step 2 RM requests information from Explain
- Step 3 Information in Explain provided to the Client by the RM
- Step 4 User selects Guide search and provides search qualifiers
- Step 5 RM Validates the query and Translates the RPN to SQL, and submits the query to the database.
- Step 6 The query targets the Guide Table.
- Step 7 The results are sent back to the RM.
- Step 8 The RM presents the results to the User via the Client

Additional scenario to be provided by Marc Gorman , ESA.

7. Recommendations

The following identifies the changes that would need to occur to accommodate the recommendations set forth in this TN.

7.1 Summary of Changes

Add one (1) new requirement.

Modify three (3) existing requirements

Eliminate Guide Collection Descriptor

Modify Guide Descriptor

7.2 Impact on Existing Documentation

7.2.1 URD

Re-word UR ID 4 to the following:

ICS shall support the following relationships;

Guide Descriptor related to Collection Descriptor

Guide Descriptor related to Guide Descriptor

Collection Descriptor related to Guide Descriptor

Modify URD to reflect bi-directional association and Search by URL

Modify URD to reflect free text search of Guide Document should be possible in Rel B

Modify URD to reflect free text search of Guide Documents and Metadata possible in Rel B

7.2.2 CIP Specification

Eliminate Guide Collection Descriptor

Add additional attributes to the Guide Descriptor's ARS

Add Document(text) as a type of element set to element set details in Explain

Add Guide to the list of valid database names.

7.2.3 SDD

Add Guide Descriptor Information to the data architecture section of the SDD.

Modify Functional Framework in Section 4 to include references to Guide.

Modify Bandwidth in Section 5 to reflect Guide.

Appendix A.

A.1 Guide Requirements (source: CIP URD, dtd 12 Mar 1997)

UR Id : 4

Source : DPRS TN [R1, Section 4.1.5] / URD 0.1 Review [RID 96]

Priority : 1

Need : B

Qualifier : RP

The CIP shall support collections of Guide Descriptors.

Note : The retrieval of Guide data can already be performed from many data providers via standard HTTP/WWW access methods, but this UR recognises the importance of integration of Guide within the ICS domain.

UR Id : 6 Source : DPRS TN [R1, Section 4.1.5] Priority : 2 Need : C Qualifier : RP The CIP should support collections comprised of both collection descriptors and product and/or guide descriptors at the same level and in a hierarchy.

UR Id : 16 Source : DPRS TN [R1, Section 4.1.5] Priority : 1 Need : B Qualifier : RP The CIP shall support a guide service such that lists of guide descriptors can be searched. Note : The lists of guide descriptors may also be hierarchical.

UR Id : 17 Source : DPRS TN [R1, Section 4.1.5] Priority : 1 Need : A Qualifier : RP The CIP shall support the retrieval of guide data using guide descriptors. Note Guide data is not usually subject to an order process as it is usually freely retrievable documentary or bibliographic information designed to help the user gain knowledge about a data provider's services and/or data products.

UR Id : 41 Source : PTT Tokyo / CIP-BS Team Priority : 1 Need : B Qualifier : RP The value of product or guide item descriptor identifier attributes should be URLs. Note : URNs will be considered for CIP-C.

UR Id : 80 Source : DPRS TN [R1, Section 4.1.5.7] Priority : 1 Need : B Qualifier : RP The CIP shall be able to process a list of zero or more returned guide descriptors as a result of a guide search query.

UR Id : 81 Source : DPRS TN [R1, Section 4.1.5.8] Priority : 1 Need : B Qualifier : RP The CIP shall support the retrieval of guide data from information obtained as a result of guide search queries.

UR Id : 82 Source : DPRS TN [R1, Section 4.1.5.8] Priority : 2 Need : B Qualifier : RP The CIP shall support the specification of the format and delivery method of the guide data to be retrieved, if these options are available. **Note :** Examples of guide data format include the document format (text, word processor format, PDF, etc.) or media format, the delivery may include physical transfer (e.g. by FTP), direct synchronous display, transfer by e-mail, HTML, printing etc.

UR Id : 83 Source : DPRS TN [R1, Section 4.1.5.7] Priority : 2 Need : B Qualifier : RP The CIP shall support those guide descriptors that have no external reference to guide data (i.e. the guide text may be an attribute value of the guide descriptor). In this case, any further request for the guide data should be ignored.

UR Id : 84 Source : OTN [R31] / CIP-BS Team Priority : 2 Need : B Qualifier : RP The CIP shall support the retrieval of general price information. **Note 1 :** The scope of the general price information provided (e.g. general pricing policy) depends on the item descriptor to which this guide data is related, i.e. it could be a product, a collection, or could even be related to the whole data provider (e.g. if

The CIP shall enable the directing of a search query to one or more ICS domain sites. **Note :** Searches are targeted at collections, which can span EO catalogue sites. 3.1.1.5.1 Guide Search Queries The interrogation of a guide catalogue service, although conceptually similar to that of a collection has subtle differences which are important enough to highlight in separate specific requirements. An appropriate analogy for guide is that of a bibliographical database where guide descriptor = bibliographical reference (title, author, publisher: keywords) and the guide data is abstract text or a full document. Further, bibliographical databases can be searched either on the parameters in the reference or as full text and there is also the possibility of hyperlink navigation. The requirements in this section are at a high level treating guide data as objects within the CIP domain. More specific requirements would effectively describe bibliographical information retrieval systems (and also perhaps Z39.50 [R17]) which would not be appropriate in this URD. Currently the services of guide search and retrieval are often handled using the common HTTP/WWW protocols; this has been found very successful for this task and therefore interoperability between guide and the rest of the catalogue services may be achieved through mechanisms to link the CIP and HTTP/WWW protocols. The first set of requirements in this section refers to the searching of the guide descriptors only, i.e. not the data itself, which is given by UR 85.

UR Id : 77 Source : DPRS TN [R1, Section 4.1.5.7] Priority : 1 Need : B Qualifier : RP

The CIP shall support a guide search query via the specification of guide descriptor attribute values.

Note : In general, the keywords used to specify the search query would be the value of an identified attribute such as ‘GUIDE NAME’.

UR Id : 78 Source : DPRS TN [R1, Section 4.1.5.7] Priority : 2 Need : B Qualifier : RP

The CIP should support a guide search query via free text, where the text is assumed to be part of one or more specified guide descriptor attributes.

UR Id : 79 Source : DPRS TN [R1, Section 4.1.5.7] Priority : 2 Need : B Qualifier : RP

The CIP should support a guide search query via an attribute value/free text combination.

UR Id : 85 Source : DPRS Team URD 2.1 Comment Priority : 2 Need : C Qualifier : RP

The CIP should support an optional keyword/free text search of the guide data itself.

Note : In this case, the CIP must handle a keyword or free text search directly of the guide data. This can either be managed by the Retrieval Manager software or passed on to the catalogue server software which may support the searching. The sub requirements of UR 86 map onto those of the 'Describe Guide' section of the DPRS TN [R1] used to return the data definitions of guide entries.

UR Id : 86 Source : DPRS TN [R1, Section 4.1.5.6] Priority : 1 Need : B Qualifier : RP

The CIP shall support the access and retrieval of the attributes of guide descriptors.

UR Id : 87 Source : DPRS TN [R1, Section 4.1.5.6] Priority : 1 Need : B Qualifier : RP

The CIP shall support the retrieval of a guide descriptor by specifying a guide identifier.

UR Id : 88 Source : DPRS TN [R1, Section 4.1.5.6] Priority : 1 Need : B Qualifier : RP

The CIP shall support the specification of a subset of guide descriptor attributes to be returned in the guide search query results.

UR Id : 89 Source : DPRS TN [R1, Section 4.1.5.6] Priority : 1 Need : B Qualifier : RP

The CIP shall support the specification of those semantic attributes of the guide descriptor attributes to be returned in the search query results.

UR Id : 90 Source : DPRS TN [R1, Section 4.1.5.6] CIP-BS Team URD 2.1 Comment Priority : 1 Need : B Qualifier : RP

The CIP shall support the default of returning a minimum subset of guide descriptor attributes if no preference is specified in the specified search query.

**UR Id : 91 Source : DPRS TN [R1, Section 4.1.5.6] URD 2.1 Comment Priority : 1 Need : C
Qualifier : RP**

The CIP shall support the navigation of guide hyperlinks via http.

A.2 Estimated Volumes

This section of the appendix provides estimates of the current (V0 System) and projected (ECS) number of Guide Documents. Additionally the number of V0 processed queries by type of Guide Document is also presented in this section. MUIS is unknown at this time.

A.2-1 Number of Guide Documents

The following table reflects the estimated total number of anticipated Guide Documents for each known proposed, and existing, Guide System. Column one contains the types of Guide Documents, columns 2 through 4 identifies the estimated number of Guide Documents per Guide System. For Example V0 has 694 Data Set Guides.

Table A-1. Number of Guide Documents

Guide Document	V0	ECS	MUIS
Data Set	694	230	
Data Center		0	
Source		3	
Sensor		5	
Campaign		3	

A.2.2 Total Query

The following presents the five types of Guide Documents that are currently supported in the V0 system and the number of queries, to date, that have been processed by the system, for the Guide Document type.

Table A-2. Number of V0 Queries

Guide Document	Total Queries
Data Center	8
Data Set	583
Source	186
Sensor	192
Campaign	45

A.3 Glossary

This appendix provides terms used in this TN.

Guide Document	Guides are textual documents that are developed to assist the user in understanding the Earth Observation environment. Guide Documents provide additional detailed information about a specific mission, sensor, or organization.
Guide Metadata	are characteristics about the Guide Document. This would include information about the author, publication, keywords, type and name of the document. The intent of the metadata is to provide sufficient information about the Guide Document to aid the metadata search service in locating Guide Documents of interest.
Guide Descriptor	is the collection of metadata used to describe the Guide document.
Guide Data	Guide Data as referenced in this TN refers to both the metadata and the Guide Documents.

A.4 References

1. Catalogue Interoperability Protocol (CIP) User Requirements Document, CEOS/WGISS/PTT/ICS-URD-B, Issue 2.2, 12. March 1997, CEOS PTT.
2. Catalogue Interoperability Protocol (CIP) Specification - Release B, CEOS/WGISS/PTT/CIP-B, Issue 2.2, 12. March 1997, CEOS PTT.
3. ICS System Design Document, CEOS/WGISS/PTT/SDD, Issue 1.2, March 1997, CEOS PTT.
4. Guide Document Creation Manual, IMS, NASA, January, 1997.
5. Guide Technical Note (Draft), Nill, Christiane, October 1996.
6. Guide Subsystem Design Document, EOSDIS IMS, NASA Code505, Version 0.1 October 1993.

A.5 Notes on Guide provided by Logica, authored by Christiane Nill

References

'Implementing Z39.50 in a multi-national and multi-lingual environment', Makx Dekkers, Pica - centrum voor bibliotheekautomatisierung, Leiden, The Netherlands
'Application Profile for the Government Information Locator Service (GILS)', Eliot Christian, September 24, 1996
'ZSTARTS Profile: Z39.50 Profile for the Stanford Protocol for Internet Search and Retrieval'. Draft 3, September 25, 1996

Definition of Guide:

Guide data:

Guide data are textual documents that are aimed to help the user to understand the Earth Observation environment, specific EO disciplines and to provide background context to a specific mission, sensor or organisation. The most frequently covered guide areas include:

- past, current and future EO missions
- sensors types and their characteristics
- product types and their potential application domains
- availability and suitability of EO data sets
- data provider details and services
- data processing methods and details
- algorithms
- data origin information (e.g. in-situ measurement)
- past and future international EO activities and projects (e.g. CEOS)
- general prize information (e.g. services, products)

Even if guide services have similar features to bibliographical systems, they have peculiarities on their own. A guide service can be characterised by the following:

- the guide data is usually freely retrievable documentary or bibliographic information designed to enhance the user's understanding as well as attracting new users.
- the guide data can be available in different languages
- the guide data is described, analogously to the product, by a guide descriptor. The user can locate guide data of their interest by searching the associated guide descriptors (e.g. keywords, author, title).
- guide data can also be searched by free text search in the documents themselves. This features distinguishes them from product data, since currently products can only be located by searching their descriptors, but not by searching ('data mining') the data itself.
- the guide data includes multi-media features, such as hyperlinks to other documents or encapsulation of images. Especially, the provision of hyperlink features, which facilitate user navigation, has been proven valuable.
- the guide data can be retrieved by different retrieval methods: returned always in the result of a search or retrieved only upon request in a number of different formats.

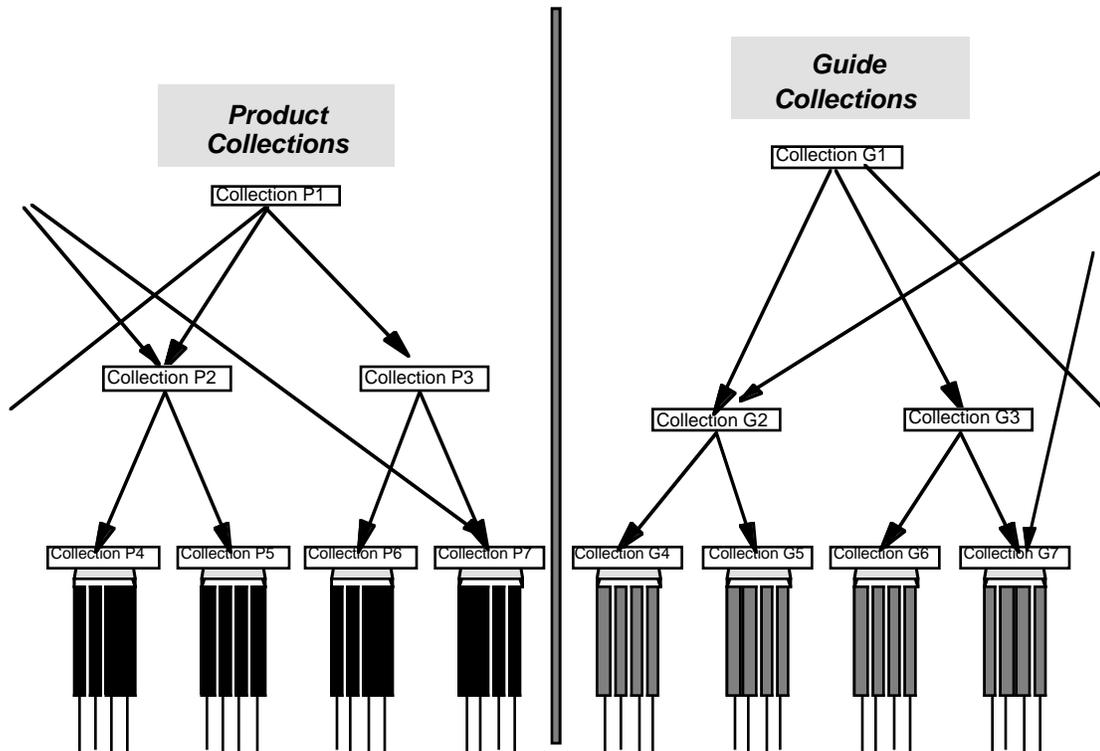
Guide Collections

Analogously to the product descriptors, guide descriptors can be grouped together according to specific criteria into guide collections. Conceptually, the same rules as for product collections apply (e.g. navigation, commonality, identifiers, etc.). Differences between product collections/guide collections include:

- guide collection attributes slightly differ from the product collection attributes. But all administrative collection attributes remain unchanged independently from the item type of the terminal members.
- only provider theme/user theme collections exist, provider archive collections are not really available. This is because all guide descriptor will have different attributes, there is no single database where all guide descriptors will have the same structure.

Relation Guide/Product

Guide and product search/retrieval can occur independently from each other (2 completely decoupled systems) or together. Figure 1-1 highlights that for one data provider site, the product collections can be connected with product collections of other data providers (remote product collections) and that the guide collections can include and can be included by remote guide collections.



FigureA-1. Guide/Product Collections

In order to achieve connectivity between the two independent collection trees the concept of **RelatedItem** is introduced.

RelatedItems

In Figure 1-2 the RelatedItem concept is graphically presented. The line in the figure describe the collection and its members (parent/child), versus the dashed lines denote that a relation between the items exist other than child/parent.

The following product/guide relationship scenarios are supported by the **RelatedItem** concept:

- a user locates a product collection of interest. The user would like to obtain more detailed information about the sensor type of the collection. The collection descriptor includes references to any related guide collections or single guide descriptors discussing the sensor characteristics (e.g. P2 => G1 or P3 => guide descriptor in G7).
- a user identifies a product descriptor. Included in the product descriptor is a reference to a related publication of a 'fire detection application', where the product has been used (product descriptor in P6 => guide descriptor in G4).
- a user investigates by studying guide documents the product type that he would need The guide descriptor includes references to the related product type collections (e.g. G4 => P7).

The RelatedItem concept provides the means where any kind of relation between different item descriptors can be described and links between them can be provided. The item descriptors can be of the same type (e.g. a relation between a product collection with another product collection) or of different types (e.g. a relation between a guide descriptor and a product collection descriptor). Related items do not necessarily have to be locally, but could also be remote to each other. The concept is general enough to support also new item types, e.g. software, advertisement, which are expected to be introduced in later releases of the CIP.

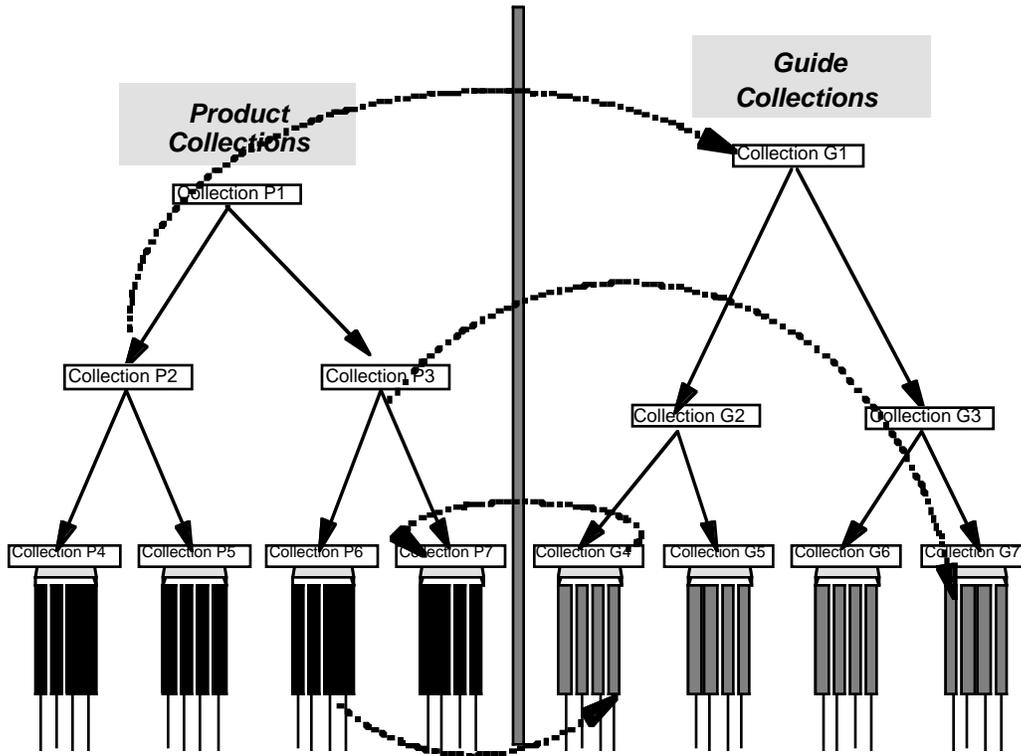


Figure A-2. RelatedItems

Ideally product collections and guide collections are organised according to the same criteria (e.g. Keywords).

Guide Descriptor Schema:

Guide =

administrative elements - applicable to all item descriptors#####

ItemDescriptorId		+	
(Authoritative)		+	
Revision		+	/** refers to document descriptor, as defined in CIP-B***/
(Review)		+	/** of descriptor ***/
Version	Id	+	/** of descriptor as defined in CIP-Spec. ***/
Purpose		+	

document related elements

Title		+	/** bib-1 ***/
0{Abstract}1		+	/** of document. ***/
PublicationDate		+	/** refers to document, bib-1***/
PublicationPlace		+	/** refers to document, bib-1 ***/
1{ Author }n		+	/** author + contact details ***/
1{LanguageOfDocument}n	+		/** the language in which the document is available ***/
Document		+	/** pointer or document itself ***/
DocumentType		+	/** controlled list of terms, e.g. user manual, algorithm description, etc.)
0 { ReferencePaper } n		+	/** as defined in CIP-B ***/
(NumberOfWords)	+		/** integer , ZSTARTS***/
(NumberOfBytes)	+		/** integer, ZSTARTS ***/
Keywords		+	/** as defined for collections ***/
0{ServiceOptions } n	+		/** see CIP-B Spec, note that for product and documents different delivery media/formats apply***/
0{ CrossReference } n		+	/** hyperlinks within doc, from GILS***/
0{ RelatedItems }n		+	/** e.g. Related product

descriptors,

guide descriptors, collection descriptors, which are found to be relevant ***/

Author =

Contact

Document =

[DocumentData | DocumentPointer]

CrossReference =

Title		+
RelationshipDescription		+
CrossReferenceLinkage		

CrossReferenceLinkage =
 LinkageType + /** URLs within document **/
 document, /** data content type, e.g.
 Linkage image **/
 /** URL **/

RelatedItems=
 ItemDescriptorId +
 RelationshipDescription

Changes to Collection Schema

Not applicable Collection attributes

The following collection attributes do not really apply for guide collections, they are product specific:

- Browse
- SpatialCoverage
- Data Resolution
- ProcessingLevel
- ArchivingCentreId
- ProcessingCentre
- DataOriginator
- QACollectionStatistics
- DeliveredAlgorithmPackage
- Investigator

Changes to existing collection attributes

- It is not clear what the difference between Investigator/Technical is. For guide collections, only one contact point for questions is necessary. I propose that Investigator/Technical are merged into Technical, and is made repeatable.
- I propose that within Contact the 'OrganisationName' is made mandatory.
- Add in ServiceOptions structure an optional DeliveryContact (Contact)
- Include 'LanguageOfDocument' in ReferencePaper.

I propose to structure the collection attributes in three categories:

- identification (applicable for all collections)
- administrative (applicable for all collections)
- attributes dependent on terminal collection item type member.

Additional attributes for guide collection descriptor:

- PublicationPlace
- LanguageOfDocument

Proposed changes to Spec.:

- 'Keywords' - introduce two different compounds: controlled/uncontrolled keywords
- 'RelatedCollections' - the concept should be broadened into 'RelatedItems', since between any kind of item relationships can exist.
- "document-text" as new structure attribute (bib-1, 106) add text for Truncation Attributes, table A.7 updates ?

Discussion points:

- all item descriptors, in particular of heritage systems (e.g. product descriptors) might be in their native language. We need to analyse if Z39.50 'Variant' can be used to cover this. In general the multi-language environment needs to be better considered in the CIP.
 - Negotiation of the language to be used during initialization (e.g. what character set should be used: 7-bit ASCII)
 - dependent on the language (country) sort criteria are likely to change for specific attributes. If the sort service is set-up this needs consideration.
- CIP includes different date/time representations than bib-1, ZSTARTS, etc. In the case CIP would be required to be compatible with these profiles, we run into problems.
- For guide documents, when full text searches are applied, etc. ranking and sorting of the results will be required.
 - Currently, we do not foresee for CIP-B the sort operations.
 - Ranking will be an essential feature for evaluating the results from a document free text search.
- Since for guide descriptors not really provider archive collections exist, the mandate that 'all product descriptors must be physically contained in at least one provider archive collection and therefore all archive collections must have a corresponding attribute definition table' is not really applicable. How shall we handle this?