



Intelligence Community and Department of Defense Content Discovery & Retrieval Integrated Product Team

SOAP Interface Encoding Specification for Content Discovery and Retrieval Brokered Search

Version 2.0

6 September 2013

Distribution Notice:

This document has been approved for Public Release and is available for use without restriction.

Table of Contents

Chapter 1 - Introduction	1
1.1 - Purpose	1
1.2 - Scope	1
1.3 - Background	1
1.4 - Enterprise Need	2
1.5 - Audience and Applicability	3
1.6 - Conventions	3
1.7 - Dependencies	4
1.8 - Conformance	4
1.9 - Version Policies	4
1.10 - Namespaces	4
1.11 - Security	5
Chapter 2 - Development Guidance	6
2.1 - Brokered Search	6
2.1.1 - Stateless/Stateful Search	6
2.1.2 - Broker to Broker Search	6
2.1.3 - Heterogeneous Sources Invocation	7
2.1.4 - Results Aggregation and Paging	8
2.2 - Source Identification	9
2.2.1 - Source Description	9
2.2.2 - Source Routing	9
2.2.2.1 - Static routing	9
2.2.2.2 - Smart routing by Query Inspection (Non-Normative)	9
Chapter 3 - Definitions, Interfaces & Constraints	10
3.1 - Brokered Search Function	10
3.1.1 - Preconditions	10
3.1.2 - Input	10
3.1.2.1 - Header	10
3.1.2.2 - Body	11
3.1.2.3 - Example	12
3.1.3 - Output	13
3.1.3.1 - Header	13
3.1.3.2 - Body	14
3.1.3.3 - Example	16
3.1.4 - Post-conditions	18
3.1.5 - Fault Conditions	18
3.1.5.1 - Fault Message Example	20
3.2 - Subsequent Brokered Search Function	21
3.2.1 - Preconditions	21
3.2.2 - Input	22
3.2.2.1 - Header	22
3.2.2.2 - Body	22
3.2.2.3 - Example	24
3.2.3 - Output	24
3.2.4 - Post-conditions	25
3.2.5 - Fault Conditions	25

Chapter 4 - CDR Conformance Validation	26
Appendix A - Feature Summary	27
Appendix B - Change History	28
Appendix C - Mapping to Specification Framework	29
C.1 - Brokered Search Function	29
C.2 - Source Identification Function	29
Appendix D - Changes from Prior Version	31
Appendix E - Acronyms	34
Appendix F - Bibliography	40
Appendix G - Points of Contact	43
Appendix H - IC CIO Approval Memo	44

List of Figures

Figure 1 - CDR Architecture Model	2
---	---

List of Tables

Table 1 - Dependencies	4
Table 2 - Referenced XML Namespaces	4
Table 3 - Header Elements for Brokered Search Requests	10
Table 4 - Attributes of /cdrs:SearchRequest for initial search request	11
Table 5 - Required Header Elements for Search Response	13
Table 6 - Result Set Elements within Atom Feed	14
Table 7 - cdrb:sourceStatus in the Response	15
Table 8 - SOAP Fault Conditions	19
Table 9 - Header Elements for Subsequent Search Requests	22
Table 10 - Subsequent Requests	22
Table 11 - Revision History	28
Table 12 - Specification Framework Input Variables	29
Table 13 - Specification Framework Output Variables	29
Table 14 - Specification Framework Input Variables	29
Table 15 - Specification Framework Output Variables	30
Table 16 - Summary of Changes from V1.1	31
Table 17 - Acronyms	34

Chapter 1 - Introduction

1.1. Purpose

The **Brokered Search Component**, as defined by the IC/DoD Content Discovery and Retrieval (CDR) Specification Framework (CDR-SF),^[6] serves as the primary mechanism to facilitate the distribution of search requests to applicable/relevant sources ¹ and aggregate the results returned from different sources. The identification of appropriate sources is an additional capability towards source description and source routing.

This specification defines requirements and provides guidelines for the realization of the CDR Brokered Search Component in a web service using the SOAP ² messaging protocol, hereafter termed a broker in this document. The content of this specification describes a broker's behavior, interface and other aspects in detail, providing enough information for Brokered Search providers and consumers to create and use CDR-conformant brokers.

1.2. Scope

This specification defines the SOAP interface between the consumer and the broker to support Brokered Search activities. In particular, the broker's behavior and interfaces are specified to enable stateless/stateful search, synchronous/asynchronous response, broker to broker search, source identification, heterogeneous sources invocation (CDR REST/SOAP conformant sources and non-CDR conformant sources), results aggregation and results paging.

The following capabilities, while related with broker, are out of the scope of the specification:

- Mediation/translation processes that may be implemented by the broker
- Adding a new source that the broker can use as a federation target
- The process for the consumer to obtain the descriptions of the sources
- Routing to unknown sources that are not recognized by the broker
- De-duplicating the same results returned from multiple sources.

1.3. Background

This specification is a part of the set of specifications that define the concrete, implementation-specific guidance for the services defined under the auspices of the Content Discovery & Retrieval (CDR) Integrated Project Team (IPT). The CDR Architecture prescribes an abstract-to-concrete model for the development of architecture elements and guidance for content discovery and retrieval. Each layer or tier of the model is intended to provide key aspects of the overall guidance to achieve the goals and objectives for joint DoD/IC content discovery and retrieval. The following graphic, discussed in detail within the CDR Reference Architecture (CDR-RA),^[2] illustrates this model.

¹The source may be a content collection component or another Broker Search component. It acts as the federation target defined in CDR-SF^[6] for Brokered Search Component.

²SOAP is a protocol used by web services in the exchange of structured information.^[16]

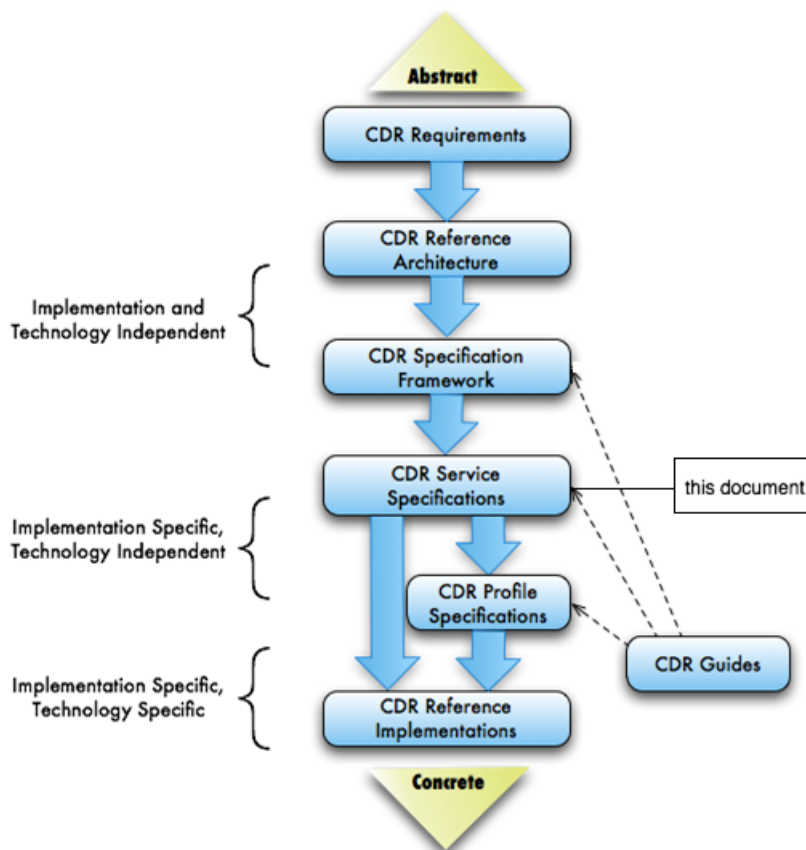


Figure 1 : CDR Architecture Model

As illustrated in Figure 1, the CDR Specification Framework^[6] derives from the CDR Reference Architecture^[2] and describes behavior in terms of the capabilities, components, and usage patterns defined in the CDR-RA.^[2] Multiple CDR Service Specifications are derived from the CDR-SF,^[6] with separate specifications associated with the components of the architecture (e.g., Brokered Search) and, for each service, separate specifications to address Representational State Transfer (REST) and SOAP implementations.

This document is a specification for implementing the CDR Brokered Search Component using the SOAP^[16] standard. It is intended to parallel the corresponding REST specification, the IC/DoD RESTful Interface Encoding Specification for CDR Brokered Search (CDR-RB),^[3] as closely as possible, to minimize the difficulties in interoperating. Additional CDR Guides, Profile Specifications, or Reference Implementations may provide additional guidance on implementing this specification in a particular context.

1.4. Enterprise Need

Enterprise needs and requirements for this specification can be found in the following policies and implementation guidance.

- IC Information Technology Enterprise (IC ITE)

- Intelligence Community Information Technology Enterprise (IC ITE) Increment 1 Implementation Plan^[8]
- 500 Series:
 - Intelligence Community Directive (ICD) 501, Discovery and Dissemination or Retrieval of Information within the IC^[10]
 - Joint IC/DoD Memorandum, IC and DoD Commitment to an Interoperable Service-Based Environment (13 Jul 07)^[15]

1.5. Audience and Applicability

Data Encoding Specifications (DESSs) are primarily intended to be used by those developing tools and services to create, modify, store, exchange, search, display, or further process the type of data being described.

Within the IC, the conditions of use and applicability of this technical specification are defined outside of this technical specification. IC Standard (ICS) 500-20, Intelligence Community Enterprise Standards Compliance, defines the IC Enterprise Standards Baseline (IC ESB) and the applicability of such to an IC element. The IC ESB defines the compliance requirements associated with each version of a technical specification. Each version will be individually registered in the IC ESB. The IC ESB will define, among other things, the location(s) of the relevant artifacts, prescriptive status, and validity period, all of which characterize the version and its utility. Additional applicability and guidance may be defined in separate IC policy guidance.

Within the DoD, the conditions of use and applicability of this technical specifications are defined outside of this technical specification, and can be located within the DoD Information Technology Standards Profile Registry (DISR).

1.6. Conventions

The key words "MUST," "MUST NOT," "REQUIRED," "SHALL," "SHALL NOT," "SHOULD," "SHOULD NOT," "RECOMMENDED," "MAY," and "OPTIONAL" in this specification are to be interpreted as described in the IETF RFC 2119.^[13] When these words are not capitalized, they are meant in their natural-language sense.

When describing concrete XML schemas and example XML documents, this specification uses XPath as the notational convention. Each member of an XML schema is described using an XPath notation (e.g., /x:RootElement/x:ChildElement/@Attribute). The use of {any} indicates the presence of an element wildcard (<xs:any/>). The use of @{any} indicates the presence of an attribute wildcard (<xs:anyAttribute/>).

Examples in this text are distinguished by a blue border. These are meant to be illustrative and only one way that the described syntax can be used.

```
<atom:entry>
  <atom:title>This is an example.</atom:title>
</atom:entry>
```

1.7. Dependencies

This technical specification depends on the additional technical specifications or additional documentation listed in the following table. The documents listed below are referenced in this encoding specification, and are normative or informative as indicated in the dependencies table.

Table 1 - Dependencies

Name	Dependency Description
SOAP Interface Encoding Specification for Content Discovery and Retrieval Search(CDR-SS) ^[7]	CDR-SS ^[7] specifies the CDR Search interface for SOAP. This version of Brokered Search extends the CDR SOAP Search Specification V3.0 (Normative).

1.8. Conformance

For an implementation to conform to this Brokered Search specification, it MUST adhere to all REQUIRED aspects of the specification. A SOAP based broker extends the interface of CDR SOAP Search Service so it MUST conform to CDR SOAP Search Specification.^[7]

1.9. Version Policies

The CDR specifications have been developed and are periodically refreshed to incorporate evolving needs and experience. Work is scheduled to meet the most pressing needs for expanding scope and maintaining consistency across the CDR specification family. Whole number versioning, e.g. v1.0 to v2.0, is used for all new versions.

1.10. Namespaces

Namespaces referenced in this document and the prefixes used to represent them are listed in the following table.

Table 2 - Referenced XML Namespaces

Prefix	URI	Description
soap	http://www.w3.org/2003/05/soap-envelope	W3C SOAP Version
wsa	http://www.w3.org/2005/08/addressing	WS-Addressing Definition
wsaw	http://www.w3.org/2006/05/addressing/wsdl	WS-Addressing WSDL Binding
cdrs	urn:cdr:search:3.0	CDR Search Specification

cdrb	urn:cdr:broker:2.0	CDR Brokered Search Specification
cdrd	urn:cdr:describe:1.0	CDR Describe Specification

1.11. Security

This specification does not directly address security concerns. It will be necessary for any implementation of this specification to address security concerns relevant to the systems with which they interact and the corresponding governance bodies. Several aspects of Brokered search, to include returning only the results for which the requesting entity is authorized, should be addressed in the detailed security plan of an implementation, but are out of scope for this document.

Chapter 2 - Development Guidance

2.1. Brokered Search

The primary function of a broker is to support federated search by distributing a consumer's search request to one or more sources and returning the aggregated search results from the sources. The broker extends CDR Search Service and defines additional request/response parameters as specified in Section 3. The broker extensions supports the capabilities as described in this section.

2.1.1. Stateless/Stateful Search

A broker MAY process all search requests as a new search and no prior state/cache is associated with a search. Such a stateless search MAY provide a simple interface to a consumer. For a stateless search, the broker MUST return the results to the consumer synchronously, i.e. the broker MUST wait until the completion of federation results processing to respond to the consumer.

A broker MAY persist the information associated with an initial search request and provide stateful interaction in order to improve responsiveness and avoid redundant requests to the sources. For a stateful search, the broker response MUST contain the result set identification and the broker MAY support asynchronous response, i.e. the broker MAY return intermediate status information with an empty result to the consumer before the search results are sent. The interim response MAY contain the status for all participating sources to provide diagnostic information to the consumer. The status information could indicate items such as the overall progress for each identified source (e.g. waiting, processing, etc.) or whether results are being added to the set of search results (i.e. number of retrieved results is increasing).

Upon receiving the interim response, a consumer may continue to request the initial search results using the result set identifier. The broker MAY respond with status information or send the results back when it is ready.

Subsequent requests may be directed to the broker to access an identifiable subset of search results associated with the result set identifier such as a new page or results from a particular source. The broker MAY respond with results as requested or with information on the progress in creating the set of search results.

2.1.2. Broker to Broker Search

A broker MAY act as a consumer and MAY distribute a search request to another broker, i.e. the latter broker may act as a source of the former broker. The second broker may pass the request to a third broker. As search propagation goes on, the search request may reach brokers across various networks. A loop may be created if the same broker receives the same search request again via different search propagation paths. Special consideration must be taken to prevent the potential loop situation; hence, the processing of a broker to broker search request SHOULD include sufficient routing processing in addition to search request processing.

To detect and avoid looping, this specification defines a path parameter to track the identity of brokers through which a search request passes. Each broker MUST include its identifier in the

path parameter, as defined in Section 3. A consumer is considered as broker type if the path parameter is found in its search request. When a consumer does not provide the path, it is assumed to be the search request initiator.

Each time a search request is sent from one broker to another, the path parameter **MUST** be updated by the sending broker to include the broker identifier of the sender. To avoid a loop situation, the receiving broker **MUST** check if its own broker identifier is already included in the path parameter value. If the broker's identifier is already in the path and the broker has available a mechanism to determine whether continuing the broker process is warranted, then the broker **MUST** use and **MUST** follow the results of this mechanism to determine whether to continue to distribute the search request. If the broker does not use this mechanism or the results are ambiguous, then the broker **MUST** terminate the search request as in the circumstance above where no mechanism is available. A broker **SHOULD** have documented the mechanism by which it would decide to continue the processing; however, the mechanism by which a broker would decide whether to proceed or terminate the search request is outside the scope of this specification.

Note that any broker may terminate what it recognizes as a loop, even if the received path information is not complete or accurate.

The broker identifier **MUST** be unique across the brokers. The examples of possible types of identifier include UUID, IC-ID^[9], URI, a fully qualified domain name (FQDN), a distinguished name (DN) as used in Lightweight Directory Access Protocol (LDAP), or names in unknown namespaces.

2.1.3. Heterogeneous Sources Invocation

A broker **MUST** have a specific set of sources it is prepared to invoke. Each source is assigned an identifier and the source identifier **MUST** be unique within the broker. The source identifier is used to link the routing information, the search result, or status information with a source.

A broker **MAY** distribute the search to a number of different types of sources and brokers. For brevity, the following terms will be used:

- REST source type: a source conforming to CDR-RS^[4]
- SOAP source type: a source conforming to CDR-SS^[7]
- REST broker type: a source conforming to CDR-RB^[3]
- SOAP broker type: a source conforming to this specification
- broker type: any broker whether a REST or SOAP broker type or a broker not conforming to CDR specifications
- non-broker type : any source that is not broker type

A broker receives a search request and uses it to assemble the search requests to send to each identified source:

- For a source of SOAP type or SOAP broker type, the broker **SHOULD** distribute a search request that includes the parameters in the original search request, with adjustment when applicable.
- For a source of REST type or REST broker type, a broker **SHOULD** map the original search request to a CDR REST Search request, with adjustment when applicable.

- For a non-CDR conformant source, a broker SHOULD create the search request based on the original request and the syntax supported by the source. This assumes that the broker has specialized knowledge of specific non-CDR proprietary interfaces.
- In addition, for a source of broker type, the broker MUST include its broker identifier in the path and pass it to the targeted broker.

A consumer may specify the expected number of results to view or a timeout value that indicates to the broker how long the requester is willing to wait for slow-to-respond sources.

- The broker MAY determine how to allocate the maximum results number across its sources. For example, it MAY request the same number of results from each source, or it MAY vary the number of results requested from the sources. Not all sources allow a consumer to dictate the maximum number of results, so this parameter MAY simply serve as a suggestion indicating the consumers desired maximum number of results.
- When passing a timeout value to a source, the broker MAY reduce the value to ensure adequate performance, such as accounting for time needed for its internal processing of results.

2.1.4. Results Aggregation and Paging

When the search results become available, the broker MAY wait for responses from all invoked sources, aggregate the entire results set and send the aggregated results to the consumer. This tends to be simpler to implement and has the benefit of providing better sorting capability.

The broker MAY aggregate on the partial result set using consumer specified criteria such as timeout or the number of results to be retrieved to reduce the amount of storage required of the broker for the subsequent search requests. A broker MAY drop the search results received after the requested timeout has expired or after the requested number of results is retrieved. The result set would remain static for the subsequent requests and the continuity of search results is guaranteed when paging back and forth across subsets of search results.

A broker MAY return results to the consumer as soon as sufficient number of results are retrieved, e.g. the requested number of results in the first page is retrieved, which may provide reduced response time and better scalability across disparate environments.

Each result returned in a result set MUST indicate the source from which it originated. For a source of broker type, the search path MAY be included. This is similar to the path maintained for broker to broker search (Section 2.1.2) but it indicates the return path of results rather than the outbound path of the search request. A broker MAY provide relevance scores for individual search results with respect to the particular search with which it is identified. Relevance scores generated by a given source can be interpreted as relative relevance of results.

When a paging request is received, i.e., a subsequent search request with result set identifier and *cdrs:startIndex* or *cdrs:startPage* specified, the broker MAY traverse the search results and create next page results based on the information from the original search request and result inspection. Various approaches are possible for the broker to select the results in the results page. For example, the broker MAY distribute the number of selected results equally between all sources or the broker MAY use the relevancy score to select the results.

2.2. Source Identification

2.2.1. Source Description

The broker MAY have an internal list of known sources and MAY provide the consumer with a means to select sources from that list. If the means to select sources is supported, then the broker SHOULD support a means to retrieve descriptions of the known sources.

A broker that implements the means to retrieve source description MUST provide a list of potential sources and SHOULD include links to the description of the source and MAY include elements of that description.

2.2.2. Source Routing

Two different strategies are discussed for a broker to identify a specific set of sources to which the search request can be routed.

2.2.2.1. Static routing

The consumer may specify routing information to limit the brokered search request distribution to a subset of the identified sources. A *routeTo* parameter is specified in Section 3 that contains a list of source identifiers, to which the search request should be routed.

The use of *routeTo* parameter by a broker is optional. When supported, the *routeTo* parameter MAY not be provided in the search request or its value MAY be empty. In such cases, the broker SHOULD select a default set or select the sources by use of query inspection mechanism as specified below.

2.2.2.2. Smart routing by Query Inspection (Non-Normative)

In addition to static methods already described, the broker may identify the sources by query inspection. The broker MAY select source based on the query submitted by the consumer and the properties of the sources. The properties of the source may be obtained through static configuration or through the dynamic mechanism defined by CDR Describe Specification (CDR-SD).^[5]

Chapter 3 - Definitions, Interfaces & Constraints

This section defines the broker interface and contains the technical descriptions of the functions through which the consumer will interact with the service. Support for input and output parameters for each function as well as the default behaviors is described in the following tables in terms of what is expected of the broker.

3.1. Brokered Search Function

A broker **MUST** implement the brokered search function for the consumer to initiate a new search.

3.1.1. Preconditions

The following preconditions **MUST** be satisfied if the brokered search function is to correctly process input and generate results and post-conditions as described.

1. The requester is authenticated and authorized according to applicable policy requirements for this function.
2. The broker has a set of known sources and can use the search interface of the sources.

3.1.2. Input

The input to this function **MUST** be conformant with the input defined by the IC/DoD CDR Search Specification for SOAP.^[7] The input should be directed to the SOAP Endpoint address identified by the implementer.

3.1.2.1. Header

The header of the SOAP message **MUST** contain the Action element, as defined in WS-Addressing.^[17] The purpose of this element is to convey to the service which behavior to invoke. Additional elements, such as other WS-Addressing elements, **MAY** be added to the SOAP header.

Table 3 - Header Elements for Brokered Search Requests

Element Name and Description	Support
/wsa:Action This element (whose content is of type xs:anyURI) conveys the value of the [action] property and indicates to a web service which operation should be invoked.	MUST be supported by Service. MUST be provided by consumer with a value of urn:cdr:broker:2.0:request

3.1.2.2. Body

As defined in CDR SOAP Search Specification,^[7] the body of the SOAP message **MUST** contain a single `/cdrs:SearchRequest` element. The `/cdrs:SearchRequest` element contains the `/cdrs:Expression`¹ element, which provides the query expression.

An extension to CDR Search Specification has been developed to support the brokered search function. Table 4 shows the additional attributes defined by brokered search Function. The use of the parameters is illustrated in the examples in Section 3.1.2.3.

Table 4 - Attributes of `/cdrs:SearchRequest` for initial search request

Attribute Name and Description	Support
<p><code>/cdrs:SearchRequest/@cdrb:routeTo</code></p> <p>A comma-separated list of known source identifiers, to which the search request should be routed.</p> <p>The ordering of sources in the list SHOULD NOT be interpreted to indicate preference, i.e. it is NOT REQUIRED that the sources be invoked in the order listed.</p> <p>Restrictions: The value of source identifier MUST be a string of printable ASCII characters and MUST NOT contain commas. The use of source identifier is further described in Section 2.1.3.</p> <p>Default: The value is empty by default. Section 2.2.2.1 defines the broker behavior when the value is empty.</p>	MAY be supported by service.
<p><code>/cdrs:SearchRequest/@cdrb:maxResults</code></p> <p>The maximum number of results expected to be retrieved.</p> <p>The use of this parameter is further described in Section 2.1.3 and Section 2.1.4 .</p> <p>Restrictions: The value MUST be a non-negative integer.</p> <p>Default: If the value is not provided by the consumer, the broker SHOULD use its own strategy to limit the maximum number of results (e.g. use of a default number).</p>	MAY be supported by service.

¹The IC/DoD Search Specification for SOAP V3.0^[7] drops the use of `cdrs:Query` and use `cdrs:Expression` for search terms.

<p>/cdrs:SearchRequest/@cdrs:timeout</p> <p>As defined in CDR SOAP Search Specification,^[7] i.e. the maximum wait time, in milliseconds, that the consumer expects to receive the search results.</p> <p>The use of this parameter is further described in Section 2.1.3 and Section 2.1.4.</p> <p>Restrictions: The value MUST be a non-negative integer.</p> <p>Default: If the value is not provided by the consumer, the broker SHOULD use its own strategy for timeout (e.g. use of default timeout value).</p>	<p>MAY be supported by service.</p>
<p>/cdrs:SearchRequest/@cdrb:path</p> <p>A comma-separated list of broker identifiers that indicates the sequence of brokers through which a search request has been routed.</p> <p>The purpose of the parameter is to facilitate loop avoidance for broker to broker search. A broker MUST add its identifier to the end of the existing list and it MUST initialize the list if does not exist.</p> <p>Restrictions: The value of the broker identifier MUST be a string of printable ASCII characters and MUST NOT contain commas. The broker identifier MUST be unique across the Brokers. Section 2.1.2 describes the approaches to generate broker id.</p> <p>Default: The value is empty by default and the consumer is considered as a search request initiator.</p>	<p>MUST be supported by service for broker to broker search.</p>

3.1.2.3. Example

The following example illustrates an initial search request message from a search request initiator to the broker, using a notional keyword query (for more information on Query Languages, please reference the IC/DoD Search Specification for SOAP Implementations^[7]):

```
<soap:Envelope>
  <soap:Header>
    ...
    <wsa:Action>
      urn:cdr:broker:2.0:request
    </wsa:Action>
    ...
  </soap:Header>
  <soap:Body>
    <cdrs:SearchRequest
      startIndex="1" count="50" timeout="3000"
```

```

    cdrb:routeTo="abc,def"
    cdrb:maxResults="1000"
    responseFormat="http://www.w3.org/2005/Atom">
      <cdrs:Expression queryLanguage="urn:cdr:queryLanguage:keyword">
        watson ibm
      </cdrs:Expression>
    </cdrs:SearchRequest>
  </soap:Body>
</soap:Envelope>

```

Example initial search request from broker "example.com" to broker "example2.com":

```

<soap:Envelope>
  <soap:Header>
    ...
    <wsa:Action>
      urn:cdr:broker:2.0:request
    </wsa:Action>
    ...
  </soap:Header>
  <soap:Body>
    <cdrs:SearchRequest
      startIndex="1" count="50" timeout="3000"
      cdrb:routeTo="xyz"
      cdrb:maxResults="1000"
      cdrb:path="example.com"
      responseFormat="http://www.w3.org/2005/Atom">
      <cdrs:Expression queryLanguage="urn:cdr:queryLanguage:keyword" >
        watson ibm
      </cdrs:Expression>
    </cdrs:SearchRequest>
  </soap:Body>
</soap:Envelope>

```

3.1.3. Output

The output of the broker is a set of results that describe resources matching the query provided in the search request. The output to this function MUST be conformant with the output defined by the IC/DoD Search Specification for SOAP.^[7]

3.1.3.1. Header

The header of the SOAP message MUST contain the Action element, as defined in WS-Addressing. The purpose of this element is to convey to the receiver which behavior was invoked.

Table 5 - Required Header Elements for Search Response

Element Name and Description	Support
------------------------------	---------

/wsa:Action This element (content is of type xs:anyURI) conveys the value of the [action] property and indicates to a web service which operation should be invoked.	MUST be provided by Service with a value of urn:cdr:broker:2.0:response
--	---

Additional elements, such as other WS-Addressing elements, MAY be added to the SOAP header.

3.1.3.2. Body

The body of the SOAP message MUST consist of a single element, representing a page of the result set. The name and structure of the element is defined by the */cdrs:SearchRequest/@responseFormat*. The details in Table 6 show this in the context of the Atom Result Set Specification.^[1] Using this response format, the single element of the body of the SOAP message is */atom:feed*. If a different response format is specified by */cdrs:SearchRequest/@responseFormat*, the appropriate information as defined by that response format MUST be returned.

Table 6 - Result Set Elements within Atom Feed

Element Name and Description	Support
/cdrs:resultSetID As defined in CDR SOAP Based Search Specification, ^[7] i.e., a unique identifier to identify a result set corresponding to a search request.	MAY be supported by Service.
/cdrb:sourceStatus XML element whose child elements convey status information about the source identified by <i>/cdrb:sourceStatus/@sourceId</i> . When <i>/cdrb:sourceStatus</i> is included, there MUST be one <i>sourceStatus</i> element for each source which the search request was sent by the broker . The content of <i>sourceStatus</i> element MAY be extended with elements from another XML namespace to provide additional information from the broker. The input parameter <i>/cdrb:includeStatus</i> is related with this element (See Table 10).	MAY be supported by Service.

<p>/cdrb:resultSource</p> <p>XML element that MUST be included with each result in order to identify the source from which the result was found. The XML element MUST have <i>sourceId</i> attribute (defined below) and MUST have a value of type string that contains the short name of the source and corresponds to <i>/cdrb:resultSource/@sourceId</i>.</p> <p>The short name of the source is a brief human-readable title that identifies this source. Its value MUST be a string of printable ASCII characters.</p>	<p>MUST be supported by Service.</p>
<p>/cdrb:resultSource/@sourceId</p> <p>XML attribute whose value is a string conveying the source identifier, as defined for the <i>/cdrb:routeTo</i> parameter in Table 4.</p>	<p>MUST be supported by Service.</p>
<p>/cdrb:path</p> <p>A comma-separated list of broker identifiers that indicates the sequence of brokers through which a search result has been returned.</p> <p>This element is intended as an indication of the brokers involved for broker to broker search. A broker MAY add its identifier to the beginning of the existing list and it MAY initialize the list if does not exist.</p> <p>Use of the path parameter for activities such as auditing is outside the scope of this specification.</p>	<p>MAY be supported by Service.</p>

When */cdrb:sourceStatus* is provided in the broker's response to the initial search request, the status information provided **MUST** conform to that shown in Table 7. Additional status information **MAY** also be provided.

Table 7 - cdrb:sourceStatus in the Response

Attribute/Element Name and Description	Support in the context of providing sourceStatus
<p>/cdrb:sourceStatus/@sourceId</p> <p>XML attribute whose value is a string conveying the source identifier as defined for <i>/cdrb:routeTo</i> parameter in Table 4.</p>	<p>MUST be supported by Service</p>
<p>/cdrb:sourceStatus/shortName</p> <p>XML element whose value is a string conveying the short name of this source, where the short name is defined in Table 6 for <i>/cdrb:resultSource</i> and corresponds to <i>/cdrb:sourceStatus/sourceId</i>.</p>	<p>MUST be supported by Service</p>

<p>/cdrb:sourceStatus/status</p> <p>The current status of a single source. It may contain one of the following values:</p> <ul style="list-style-type: none"> • <i>excluded</i>: The source was excluded by the broker. There may be a number of reasons for excluding a source, for example, if a maximum number of sources is exceeded, or if the source does not support query parameters in the request. • <i>waiting</i>: The search request has been sent to the source, and the broker is waiting for a complete response from the source. • <i>error</i>: The source returned an error response. • <i>timeout</i>: The source failed to respond within the configured timeout period. • <i>processing</i>: The broker received a complete response from the source, but is processing the result set (e.g., converting format, merging with other results, re-ranking). • <i>complete</i>: A response was successfully received and the result set from this source has been processed. 	<p>MUST be supported by Service</p>
<p>/cdrb:sourceStatus/resultsRetrieved</p> <p>The number of search results that the broker retrieved from the source</p>	<p>MUST be supported by Service.</p>
<p>/cdrb:sourceStatus/totalResults</p> <p>The number of total results matching the query, as reported by the source.</p>	<p>MAY be supported by Service.</p>
<p>/cdrb:sourceStatus/elapsedTime</p> <p>The elapsed time, in milliseconds, for the source response.</p>	<p>MAY be supported by Service.</p>

3.1.3.3. Example

The results presentation SHOULD follow a documented Result Set specification.

The CDR Specification set includes a single predefined Result Set formatting definition that IC/DoD organizations can leverage in their search service implementations, the IC/DoD ATOM Data Encoding Specification for Content Discovery and Retrieval Result Sets.^[1] Implementers SHOULD consult appropriate policy and implementation guidance to determine requirements or recommendations concerning the use of particular Result Set formatting.

The following represents a sample output using the Atom format for the result set augmented with response elements as identified in CDR-IPT SOAP Search Specification.^[7] Search service implementations SHOULD include these response elements, although the exact mechanism and syntax for including these elements may vary depending on the Result Set specification being used. Consult the Atom specification^[14] and CDR ATOM Specification^[1] for specifics on using the Atom format.

The example below shows the output with two sources: the source "abc" is a source of non-broker type and source "def" is of broker type. The namespaces referenced in the example are listed in Table 2.

```
<soap:envelope>
  <soap:header>
    <wsa:action>urn:cdr:broker:2.0:response</wsa:action>
  </soap:header>
  <soap:body>
    <atom:feed>
      <atom:id>urn:uuid:60a76c80-d399-11d9-b93C-0003939e0af6</atom:id>
      <atom:title>Query Results for &quot;watson ibm&quot; </atom:title>
      <atom:updated>2003-12-13T18:30:02Z</atom:updated>
      <atom:author><atom:name>Enterprise Catalog</atom:name></atom:author>
      <opensearch:totalResults>492420</opensearch:totalResults>
      <opensearch:startIndex>1</opensearch:startIndex>
      <opensearch:itemsPerPage>10</opensearch:itemsPerPage>
      <cdrs:resultSetID>1234</cdrs:resultSetID>
      <cdrs:SearchRequest startIndex="1" count="10"
        cdrb:routeTo="abc,def" cdrb:maxResults="1000"
        responseFormat="http://www.w3.org/2005/Atom">
        <cdrs:Expression queryLanguage="urn:cdr:queryLanguage:keyword">
          watson ibm
        </cdrs:Expression>
      </cdrs:SearchRequest>
      <cdrb:sourceStatus sourceId="abc">
        <cdrb:shortName>non-broker</cdrb:shortName>
        <cdrb:status>waiting</cdrb:status>
        <cdrb:resultsRetrieved>100</cdrb:resultsRetrieved>
        <cdrb:totalResults>222222</cdrb:totalResults>
        <cdrb:elapsedTime>2000</cdrb:elapsedTime>
      </cdrb:sourceStatus>
      <cdrb:sourceStatus sourceId="def">
        <cdrb:shortName>broker</cdrb:shortName>
        <cdrb:status>waiting</cdrb:status>
        <cdrb:resultsRetrieved>50</cdrb:resultsRetrieved>
        <cdrb:totalResults>12345</cdrb:totalResults>
        <cdrb:elapsedTime>1500</cdrb:elapsedTime>
      </cdrb:sourceStatus>
      <atom:entry>
        <atom:id>urn:uuid:60a76c80-d399-11d9-b93C-0003939e0af7</atom:id>
        <atom:title>IBM Watson</atom:title>
        <atom:updated>2011-02-21T00:00:00Z</atom:updated>
        <atom:link rel="alternate"
          href="http://www-03.ibm.com/innovation/us/watson/index.html"/>
        <relevance:score>0.97</relevance:score>
        <cdrb:resultSource cdrb:sourceId="abc">non-broker</cdrb:resultSource>
      </atom:entry>
      <atom:entry>
        <atom:id>urn:uuid:60a76c80-d399-11d9-b93C-0003939e0af8</atom:id>
        <atom:title>IBM Watson</atom:title>
        <atom:updated>2011-02-22T00:00:00Z</atom:updated>
        <atom:link rel="alternate"
          href="http://www-02.ibm.com/innovation/us/watson/index.html"/>
```

```
<relevance:score>0.95</relevance:score>
<cdrb:resultSource cdrb:sourceId="def">broker</cdrb:resultSource>
<cdrb:path>example.com,example2.com</cdrb:path>
</atom:entry>
...
</atom:feed>
</soap:body>
</soap:envelope>
```

3.1.4. Post-conditions

The following conditions MUST be met upon completion of a search.

1. The results available to be returned to the requester are relevant to the input query.
2. The response will consist of a list of results or an appropriate fault.
3. The results are in the correct format.
4. The authenticated requester has been authorized to receive each result in the response.
5. The use of this function has been audited according to applicable policy.

3.1.5. Fault Conditions

In addition to fault conditioned that is defined in CDR Search Specification,^[7] an implementation of the Brokered Search function MAY provide any of the faults listed in Table 8 as a SOAP Fault to the consumer.

Fault response adheres to the SOAP 1.2 specification.^[16] The */soap:Fault/Code/Value* element is used to convey the general type of error condition and MUST be from the enumeration */soap:faultCodeEnum* as described in the SOAP 1.2 specification section 5.4.6. In addition, for the current use of this function, the fault MUST also contain the */soap:Fault/Code/Subcode* element and its child element */soap:Fault/Code/Subcode/Value*.

The */soap:Fault/Code/Subcode/Value* as listed in Table 8 supports automated processing of CDR specific errors. The */soap:Fault/Reason* element has one or more */soap:Fault/Reason/Text* elements as its children. Each */soap:Fault/Reason/Text* element SHOULD be used to provide a human-readable explanation of the fault. The */soap:Fault/Reason/Text* element MUST include the *xml:lang* attribute.

Table 8 outlines the service specific fault conditions that MAY be generated by CDR Brokered Search function implementations. For the fault conditions related with a specific source, the fault is a child element of *cdrb:sourceStatus*.

Table 8 - SOAP Fault Conditions

/soap:Fault/Code/Value /soap:Fault/Code/Subcode/Value /soap:Fault/Reason/Text	Fault Description
soap:Receiver cdr:broker:soap:fault:Timeout Search Timed Out	A fault used if the broker cannot generate the response within the consumer specified timeout period.
soap:Receiver cdr:broker:soap:fault:Federation Federation Results Processing Failure	A fault used if the broker cannot invoke, monitor, or process results from the Federation Results Processing activity.
soap:Sender cdr:broker:soap:fault:DuplicatedSearch Duplicated Search Request	A fault used if the broker already received the same search request
soap:Sender cdr:broker:soap:fault:SourceIdentification Invalid Source	A fault used if the broker cannot identify the source to invoke.
soap:Receiver cdr:broker:soap:fault:SourceInvocation Search Component Invocation Fault	A fault used if the broker cannot invoke the identified source. The fault is a child element of <i>cdrb:sourceStatus</i> .

soap:Receiver cdr:broker:soap:fault:ResultsProcessing Cannot Process the Results	A fault used when the broker cannot process the results set of an individual source. The fault is a child element of <i>cdrb:sourceStatus</i> . This may indicate an error in the returned results set or an inconsistency in interpreting the results set specified format.
---	--

3.1.5.1. Fault Message Example

The following shows a fault message **Search Component Invocation Fault**. The namespaces referenced in the example are listed in Table 2.

```
<soap:envelope>
  <soap:header>
    <wsa:action>urn:cdr:broker:2.0:response</wsa:action>
  </soap:header>
  <soap:body>
    <atom:feed>
      <atom:id>urn:uuid:60a76c80-d399-11d9-b93C-0003939e0af6</atom:id>
      <atom:title>Query Results for "watson ibm"</atom:title>
      <atom:updated>2003-12-13T18:30:02Z</atom:updated>
      <atom:author><atom:name>Enterprise Catalog</atom:name></atom:author>
      <opensearch:totalResults>492420</opensearch:totalResults>
      <opensearch:startIndex>1</opensearch:startIndex>
      <opensearch:itemsPerPage>10</opensearch:itemsPerPage>
      <cdrs:resultSetID>1234</cdrs:resultSetID>
      <cdrs:SearchRequest startIndex="1" count="10"
        cdrb:routeTo="abc,def" cdrb:maxResults="1000"
        responseFormat="http://www.w3.org/2005/Atom">
        <cdrs:Expression queryLanguage="urn:cdr:queryLanguage:keyword">
          watson ibm
        </cdrs:Expression>
      </cdrs:SearchRequest>
      <cdrb:sourceStatus sourceId="abc">
        <cdrb:shortName>My Source </cdrb:shortName>
        <cdrb:status></cdrb:status>
        <cdrb:resultsRetrieved>100</cdrb:resultsRetrieved>
        <cdrb:totalResults>222222</cdrb:totalResults>
      </cdrb:sourceStatus>
      <cdrb:sourceStatus sourceId="def">
        <cdrb:shortName>Other Source</cdrb:shortName>
        <cdrb:status>error</cdrb:status>
        <cdrb:resultsRetrieved>0</cdrb:resultsRetrieved>
        <soap:Fault>
          <soap:Code>
            <soap:Value>soap:Receiver</soap:Value>
          <soap:Subcode>
```

```

        <soap:Value>
            cdr:broker:soap:fault:sourceInvocation
        </soap:Value>
    </soap:Subcode>
</soap:Code>
<soap:Reason>
    <soap:Text xml:lang="en">
        Search Component Invocation Fault.
    </soap:Text>
</soap:Reason>
</soap:Fault>
</cdrb:sourceStatus>
<atom:entry>
    <atom:id>urn:uuid:60a76c80-d399-11d9-b93C-0003939e0af7</atom:id>
    <atom:title>IBM Watson</atom:title>
    <atom:updated>2011-02-21T00:00:00Z</atom:updated>
    <atom:link rel="alternate"
href="http://www 03.ibm.com/innovation/us/watson/index.html"/>
    <relevance:score>0.97</relevance:score>
    <cdrb:resultSource cdrb:sourceId="def">Other Source</
cdrb:resultSource>
</atom:entry>
    ...
</atom:feed>
</soap:body>
</soap:envelope>

```

3.2. Subsequent Brokered Search Function

The broker MAY implement subsequent brokered search function to support stateful search.

The broker MAY support results paging with cached results using subsequent brokered search function. The Brokered Search interface is extending the CDR-IPT Search interface therefore it inherits the support for the results paging function.^[7] Please refer to CDR-SS^[7] for interface definitions of results paging function with the use of */cdrs:PagingRequest*.

3.2.1. Preconditions

In addition to the preconditions of Brokered Search as specified in Section 3.1.1, the subsequent brokered search function **REQUIRES** that:

1. An initial search request has been made and the broker has established the result set associated with this request.
2. The broker has returned a resultSetID in its initial response to the search request.
3. The result set identified by the resultSetID must still be accessible through reference to the resultSetID.

3.2.2. Input

The input to the CDR Search Service MUST be a valid SOAP message that meets criteria identified in this section. The input should be directed to the SOAP Endpoint address identified by the implementer.

3.2.2.1. Header

The header of the SOAP message MUST contain the Action element, as defined in WS-Addressing.^[17] The purpose of this element is to convey to the service which behavior to invoke. Additional elements, such as other WS-Addressing elements, MAY be added to the SOAP header.

Table 9 - Header Elements for Subsequent Search Requests

Element Name and Description	Support
/wsa:Action This element (whose content is of type xs:anyURI) conveys the value of the [action] property and indicates to a web service which operation should be invoked.	MUST be supported by Service. MUST be provided by consumer with a value of urn:cdr:broker:2.0:subsequentRequest

3.2.2.2. Body

The body of the SOAP message MUST contain a single */cdrb:SubsequentRequest* element, as defined in this document. Table 10 shows the child element and attributes of the */cdrb:SubsequentRequest*. The results paging functionality can be realized by using the parameters for the subsequent search requests as defined in Table 10 and varying the values of */cdrs:startIndex* or */cdrs:startPage* (The use of */cdrs:startIndex* and */cdrs:startPage* is described in CDR-SS).^[7]

Table 10 - Subsequent Requests

Element/Attribute Name and Description	Support In the Context of Using Subsequent Search Request

<p>/cdrb:SubsequentRequest/cdrs:resultSetID</p> <p>As defined in Table 6.</p> <p>This child element MUST be provided in a subsequent search request.</p> <p>Restrictions: The value MUST be a string and MUST have been provided in the output of a previous search request.</p> <p>Default: The value is empty by default. A valid value MUST be provided to avoid errors.</p>	<p>MUST be supported by service</p>
<p>/cdrb:SubsequentRequest/@cdrb:sourceFilter</p> <p>A source identifier as defined for <i>/cdrb:routeTo</i> parameter in Table 4, that is used in a subsequent request to constrain a result set to contain only results from the indicated source.</p> <p>Restrictions: The value MUST be a string conveying the source identifier of one source.</p> <p>Default: The value is empty by default, i.e. without source filtering.</p>	<p>MAY be supported by Service.</p>
<p>/cdrb:SubsequentRequest/@cdrb:includeStatus</p> <p>A Boolean value that indicates to the broker whether or not the consumer wishes to receive status information on the sources with the results.</p> <p>If the broker supports this parameter, then a value of 1 MUST result in the status information being returned with the results. A value of 0 MUST result in the status information not being included in the response.</p> <p>Note that for the initial search request, it is up to the broker to decide whether or not to include source status in the response and the consumer cannot specify the includeStatus option.</p> <p>Restrictions: The value MUST be 0, 1.</p> <p>Default: The value is 0 by default.</p>	<p>MAY be supported by Service.</p>
<p>/cdrb:SubsequentRequest/@cdrs:startIndex</p> <p>As defined in CDR SOAP Search Specification^[7]</p>	<p>MAY be supported by service</p>
<p>/cdrb:SubsequentRequest/@cdrs:startPage</p> <p>As defined in CDR SOAP Search Specification^[7]</p>	<p>MAY be supported by service</p>
<p>/cdrb:SubsequentRequest/@cdrs:count</p> <p>As defined in CDR SOAP Search Specification^[7]</p>	<p>MAY be supported by service</p>

3.2.2.3. Example

Example subsequent search request:

```
<soap:Envelope>
  <soap:Header>
    ...
    <wsa:Action>
      urn:cdr:broker:2.0:subsequentRequest
    </wsa:Action>
    ...
  </soap:Header>
  <soap:Body>
    <cdrb:SubsequentRequest
      cdrs:startIndex="100" cdrs:count="10"
      cdrb:sourceFilter="abc" cdrb:includeStatus="1">
        <cdrs:resultSetID>1234</cdrs:resultSetID>
      </cdrb:SubsequentRequest>
    </soap:Body>
  </soap:Envelope>
```

3.2.3. Output

The output of subsequent brokered search function is the same as the output of brokered search function defined in Section 3.1.3. The namespaces referenced in the example are listed in Table 2.

Example output:

```
<soap:envelope>
  <soap:header>
    <wsa:action>urn:cdr:broker:2.0:response</wsa:action>
  </soap:header>
  <soap:body>
    <atom:feed>
      <atom:id>urn:uuid:60a76c80-d399-11d9-b93C-0003939e0af6</atom:id>
      <atom:title>Query Results for "watson ibm" </atom:title>
      <atom:updated>2003-12-13T18:30:02Z</atom:updated>
      <atom:author><atom:name>Enterprise Catalog</atom:name></atom:author>
      <opensearch:totalResults>492420</opensearch:totalResults>
      <opensearch:startIndex>100</opensearch:startIndex>
      <opensearch:itemsPerPage>10</opensearch:itemsPerPage>
      <cdrs:resultSetID>1234</cdrs:resultSetID>
      <cdrb:SubsequentRequest startIndex="100" count="10"
        cdrb:sourceFilter="abc" cdrb:includeStatus="1">
        <cdrs:resultSetID>1234</cdrs:resultSetID>
      </cdrb:SubsequentRequest>
      ...
    </atom:feed>
  </soap:body>
</soap:envelope>
```

3.2.4. Post-conditions

The post-conditions of subsequent brokered search function are the same as the post-conditions of brokered search function defined in Section 3.1.4.

3.2.5. Fault Conditions

The fault conditions of subsequent brokered search function include the fault conditions of brokered search function defined in Section 3.1.5 and the fault conditions of results paging function defined in CDR-SS.[\[7\]](#)

Chapter 4 - CDR Conformance Validation

Conformance is defined herein as “adherence to all relevant organizationally-mandated standards within a related family of specifications such that all implementations of web services based on these standards will achieve measurable levels of increased interoperability with others implementing the same standards.” Conformance is not necessarily compliance. Conformance addresses specific technical interoperability while compliance is associated with standards policy adherence as a whole. While not sufficient by itself, technical conformance is a necessary condition that helps in validating compliance with specifications.

Conformance of an implemented service to this specification will be validated by using certified tests or other means that are approved by the relevant authority by which this standard is mandated. In the absence of a certified test to validate the technical conformance, documentation that enables unambiguous traceability from each specification requirement to satisfactory execution of that requirement in the implementation must be provided. This is often in the form of a mapping matrix called a Conformance Test Matrix (CTM).

Note: The ability to validate conformance is greatly enhanced by the use of the OASIS Test Assertion Guidelines, both for articulating individual specification requirements during specification authoring, and for developing the associated implementation tests.

Appendix A Feature Summary

This document specifies the Brokered Search interface and behaviors and provide the implementation guidance for the following activities as defined in the Specification Framework:^[6]

- The brokered search coordination activity is the primary entry point to the brokered search function and provides coordination of the other activities.
- The source identification activity uses criteria provided by or derived from the consumer (which could be another broker) to identify sources to which a search request should be directed.
- The search component invocation activity sends a search request to each identified source and receives search results or a fault returned by each source. The search component invocation activity does not interact directly with the consumer therefore there is no interface associated with it.
- The federation results processing activity performs the required processing necessary to combine the outputs of the individual search component invocation activities into a single uniform result set per any processing instructions provided in the brokered search properties. The federation results processing activity does not provide standalone functionality to warrant its own interface.

An explicit mapping between this specification to the corresponding sections in CDR-SF^[6] is found in Appendix C.

Appendix B Change History

The following table summarizes the version identifier history for this DES.

Table 11 - Revision History

Doc Revision	Revised By	Revision Date	Revisions
0.1	CDR IPT	17 March 2010	Initial draft for subgroup review.
0.2	CDR IPT	02 June 2010	0.1 Comment Adjudication
0.3	CDR IPT	05 July 2010	0.2 Comment Adjudication
0.3.1	CDR IPT	21 July 2010	0.3 Comment Adjudication
0.4	CDR IPT	19 August 2010	0.3.1 Comment Adjudication
0.4.1	CDR IPT	07 September 2010	0.4 Comment Adjudication
0.4.2	CDR IPT	07 September 2010	Candidate Release
1.0-20100908	CDR IPT	08 September 2010	Minor updates
1.0-20100908	Pam Preaseau	10 September 2010	Tech Edit
1.0-20101026	D. Wigglesworth	26 October 2010	Community Review
2.0-draft v1	Lisa Xie	25 February 2013	Revised towards content re-formatting, consistence across REST/SOAP specification and latest CDR-IPT specifications .
2.0-draft v2	Lisa Xie	11 March 2013	Revised based on the subgroup commentsfor draft v1 and to add broker to broker support
2.0-draft v3	Lisa Xie	25 March 2013	Revised based on the subgroup comment for draft v2
2.0-draft v4	Lisa Xie	10 Apr 2013	Revised based on the subgroup comment for draft v3
2.0-draft v5	Lisa Xie	24 Apr 2013	Revised based on the subgroup comment for draft v4
2.0-draft v6	Lisa Xie	9 May 2013	Revised based on the subgroup comment for draft v5

Appendix C Mapping to Specification Framework

This section explicitly ties the items in this specification to the requirements of the CDR-SF.^[6] The CDR-SF^[6] identifies the requirements for creating specifications, while implementation details are outlined in this document.

C.1. Brokered Search Function

Input

Table 12 shows each input variable defined in this specification, and maps each to the Brokered Search variables as defined in the IC/DoD Content Discovery and Retrieval Specification Framework^[6] (see Table 9 in the CDR-SF^[6]).

Table 12 - Specification Framework Input Variables

Specification Framework Variables	SOAP Brokered Search Specification Variable
Search Activity Inputs	[Search Component Inputs]
Brokered Search Properties	cdrb:routeTo, cdrb:maxResults, cdrb:path

Output

Table 13 Specification Framework Output Variables maps the Brokered Search Function outputs in the CDR-SF^[6] to the elements and attributes defined in this specification.

Table 13 - Specification Framework Output Variables

Specification Framework Variable	SOAP Brokered Search Specification
Merged Results Set and associated output	[Search Component Outputs]
Total Results by Source	cdrb:resultsRetrieved, cdrb:totalResults

C.2. Source Identification Function

Input

Table 14 maps the Brokered Search Source Identification Function inputs in the CDR-SF^[6] to the elements and attributes defined in this specification.

Table 14 - Specification Framework Input Variables

Specification Framework Variable	SOAP Brokered Search Specification
Search Function Inputs	[Search Component Inputs]

Output

Table 15 maps the Brokered Search Source Identification Function outputs in the CDR-SF^[6] to the elements and attributes defined in this specification.

Table 15 - Specification Framework Output Variables

Specification Framework Variable	SOAP Brokered Search Specification
Identified Content Collections and access (as Results Set of Search Function Outputs)	cdrb:sourceStatus

Appendix D Changes from Prior Version

This following section outlines the significant changes that were made from **IC/DoD Content Discovery & Retrieval Brokered Search Service Specification for SOAP V1.1** to the current version. These changes, shown in Table 16 Summary of Changes from V1.1, were made to accommodate changes in the CDR-RA^[2] and CDR-SF^[6], to harmonize the content of this specification with that contained in the other specifications produced by this group, to incorporate feedback on the specification from pilot implementations, to resolve inconsistency between REST and SOAP specification and to introduce new capabilities.

Table 16 - Summary of Changes from V1.1

V1.1	V2.0	Rationale
First service specification developed.	Structure of document revised to reflect experience gained in writing other service specs and to use the latest EI&A template. Format of interface presentation revised to reflect experience gained in writing other service specs.	Leverage consistent structure, consistent format to make specs easier to read
Based on CDR Search specification 1.0	Based on CDR Search specification 3.0. cdrs:query changed to cdrs:expression across the document. cdrs:queryId changed to cdrs:resultSetID cdrs:resultsPerPage changed to cdrs:count Fault code created to meet with CDR Search specification 3.0 requirements. Brokered Search function input, output and fault examples updated accordingly.	Align with the latest CDR IPT specifications.
	broker to broker Search capability added	Add new capability
Section 3 Brokered Search Behavior	Section 2 Development Guidance Redone	Provide clarity and more details to broker behavior

	New request parameters are defined: cdrb:maxResults, cdrb:sourceFilter, cdrb:includeStatus, cdrb:path	Resolve inconsistency between REST and SOAP specification
	wsa:endpointReference dropped	Remove un-used method and resolve inconsistency between REST and SOAP specification
Section 2 sourceId requirement	Section 3.1.3 sourceId requirement changed from OPTIONAL to REQUIRED	Resolve inconsistency between REST and SOAP specification
Section 2 sourceStatus child elements	Section 3.1.2 Additional child elements are defined for sourceStatus.	Resolve inconsistency between REST and SOAP specification
Section 2.4 Faults Conditions	Section 3.1.5 updated with SOAP 1.2 error definition New error types added	Match with Specification Framework V2.0 Requirement Match with Search Specification 3.0 format
	Section 3.2 subsequent brokered search function added	Improve clarity and consistency.
Section 4 Discovery and Publish	Section deleted and contents merged with other sections where appropriate	Leverage consistent structure to make specs easier to read
Section 4.1 SOAP interface	Merged with contents of section 3	Leverage consistent structure to make specs easier to read
Section 4.2 Policy	Merged with contents of section 1.3	Leverage consistent structure to make specs easier to read
Section 4.3 Query Types	Merged with contents of section 2	Leverage consistent structure to make specs easier to read
Section 4.4 Result Sets	Merged with contents of section 3.1.5	Leverage consistent structure to make specs easier to read
Section 4.5 Delivery	Merged with contents of Section 2	Leverage consistent structure to make specs easier to read

Section 4.6 Security Consideration	Security statement in section 1.10	Leverage standard security statement for all service specifications.
Reference	Reference updated with the latest version. Missing references added Reference names modified for consistency.	Align with the latest CDR IPT specifications.
	Appendix added: mapping to specification framework	Align with the latest CDR IPT specifications.

Appendix E Acronyms

This appendix lists all the acronyms referenced in this encoding specification and lists other acronyms that may have been used in other encoding specifications. This appendix is a shared resource across multiple documents so in any given encoding specification there are likely acronyms that are not referenced in that particular encoding specification.

Table 17 - Acronyms

Name	Definition
A&A	Assessment and Authorization
AAS	Authoritative Attribute Sources
ABAC	Attribute Based Access Control
ABNF	Augmented Backus-Naur Form
ACSS	Allied Collaborative Shared Services
ADD	Abstract Data Definition
AICP	Authorized IC Person
AOI	Area of Interest
AOR	Area of Responsibility
API	Applications Programming Interface
APS	Attribute Practice Statement
ARH	Access Rights and Handling
AS	Attribute Service
ATO	Authority To Operate
BBOX	Bounding Box
BNF	Backus-Naur Form
CA	Certification Authority
CAPCO	Controlled Access Program Coordination Office
CAT	Catalog Services Interface Standard
CDR	Content Discovery and Retrieval
CF-NetCDF	Climate and Forecast - Network Common Data Format
CIA	Central Intelligence Agency
CIO	Chief Information Officer
CMS	Cryptographic Message Syntax
CNWDI	Critical Nuclear Weapons Design Information
COMET	Completely Open Mapping Environment
CONOPS	Concept of Operations
CORBA	Common Object Request Broker Architecture

Name	Definition
CQL	Common Catalog Query Language (CQL)
CRL	Certificate Revocation List
CSW	Catalog Service for Web
CTM	Conformance Test Matrix
CUI	Controlled Unclassified Information
CVE	Controlled Vocabulary Enumeration
D & R	Discovery and Retrieval
DAA	Designated Approval Agent
DC MES	Dublin Core Metadata Element Set
DCMI	Dublin Core Metadata Initiative
DDMS	Department of Defense Discovery Metadata Specification
DES	Data Encoding Specification
DI	Digital Identifier
DIA	Defense Intelligence Agency
DISR	DoD Information Technology Standards Registry
DN	Distinguished Name
DNI	Director of National Intelligence
DNS	Domain Name System
DOD	Department of Defense
DOE	Department of Energy
DOI	Digital Object Identifier
DOMEX	Document and Media Exploitation
EA	Enterprise Architecture
EI&A	Enterprise Integration and Archetecture
E.O.	Executive Order
EBNF	Extended Backus-Naur Form
EDH	Enterprise Data Header
EPR	Endpoint Reference
ES&IS	Enterprise Search & Integration Services
ESB	Enterprise Standards Baseline
FD&R	Foreign Disclosure & Release
FOUO	For Official Use Only
FSD	Full Service Directory
FTP	File Transfer Protocol
FY	Fiscal Year

Name	Definition
GENC	Geopolitical Entities, Names, and Codes
GeoRSS	Geographic Really Simple Syndication
GeoTIFF	Geographic Tagged Image File Format
GIF	Graphics Interchange Format
GIS	Geospatial Information System
GML	Geography Markup Language
GNS	Geographic Names Server
GUIDE	Globally Unique Identifiers for Everything
GVS	GEOINT Visualization Services
HDF-EOS	Hierarchical Data Format - Earth Observing System
HTML	HyperText Markup Language
HTTP	Hypertext Transfer Protocol
I2	Information Integration
IC	Intelligence Community
IC.ADD	Intelligence Community Abstract Data Definition
IC CIO	Intelligence Community Chief Information Officer
IC EA	IC Enterprise Architecture
IC ESB	Intelligence Community Enterprise Standards Baseline
IC ITE	IC Information Technology Enterprise
ICD	Intelligence Community Directive
ICEA	Intelligence Community Enterprise Architecture
ICPG	Intelligence Community Program Guidance
ICS	Intelligence Community Standard
ICSR	Intelligence Community Standards Registry
ICTS	Intelligence Community Technical Specification
IdAM	Identity and Access Management
IDM	Interface Data Model
IDMView	Interface Data Model View
IETF	Internet Engineering Task Force
IOC	Initial Operating Capability
IP	Internet Protocol
IPT	Integrated Project Team
IRM	Information Resource Metadata
ISBN	International Standard Book Number
ISM	Information Security Marking

Name	Definition
ISO	International Organization for Standardization
ISOO	Information Security Oversight Office
ITE	Information Technology Enterprise
JPEG	Joint Photographic Experts Group
JPIP	JPEG 2000 Interactive Protocol
JSON	JavaScript Object Notation
JWE	JSON Web Encryption
JWICS	Joint Worldwide Intelligence Communications System
JWT	JSON Web Token
KA	Knowledge Assertion
KML	Keyhole Markup Language
KOS	Knowledge Organization System
KVP	Key Value Pair
LDAP	Lightweight Directory Access Protocol
LIMDIS	Limited Distribution
LNI	Library of National Intelligence
MAC	Multi Audience Collection
MC&GIL	Mapping, Charting, and Geodesy Information Library
MC&GView	Mapping, Charting, and Geodesy View
MIME	Multipurpose Internet Mail Extensions
MTOM	Message Transmission Optimization Mechanism
NARA	National Archives and Records Administration
NATO	North Atlantic Treaty Organization
NCES	Net-Centric Enterprise Services
NGA	National Geospatial Intelligence Agency
NGDS	Net-Centric GEOINT Discovery Services
NGIC	National Ground Intelligence Center
NGT	Next Generation Trident
NIPRNet	Non-Classified Internet Protocol Router Network
NIEM	National Information Exchange Model
NIST	National Institute of Standards and Technology
NITF	National Imagery Transmission Format
NPE	Non-Person Entity
NMEC	National Media Exploitation Center
NRO	National Reconnaissance Office

Name	Definition
NSA	National Security Agency
NSG	National System for Geospatial Intelligence
NSI	National Security Information
NTK	Need-To-Know Metadata
OCIO	Office of the Intelligence Community Chief Information Officer
OCSP	Online Certificate Status Protocol
ODNI	Office of the Director of National Intelligence
OGC	Open Geospatial Consortium
OGCA	Open Geospatial Consortium Australia
OGCE	Open Geospatial Consortium Europe
ONEMail	Optimized Network Email
OPM	Office of Personnel Management
OWS	OGC Web Services
PAP	Policy Administration Point
PAYL	Payload
PDP	Policy Decision Point
PEP	Policy Enforcement Point
PK	Private Key
PKI	Public Key Infrastructure
PNG	Portable Network Graphics
PUBS	Intelligence Publications
PURL	Persistent Uniform Resource Locator
RA	Reference Architecture
RDBMS	Relational Database Management System
REST	REpresentational State Transfer
RFC	Request for Comments
RR-ID	REST Security Encoding Specification for End-to-End Identity Propagation
SAML	Security Assertion Markup Language
SAP	Special Access Program
SCI	Sensitive Compartmented Information
SIPRNet	Secret Internet Protocol Router Network
SLA	Service Level Agreement
SOAP	Simple Object Access Protocol
SQL	Structured Query Language
SSD	Special Security Directorate

Name	Definition
SSL	Secure Sockets Layer
STIL	St Louis Information Library
TCP/IP	Transmission Control Protocol/Internet Protocol
TDC	Trusted Data Collection
TDF	Trusted Data Format
TDO	Trusted Data Object
TGN	Thesaurus of Geographic Names
TIFF	Tagged Image File Format
TIN	Triangulated Irregular Network
TLS	Transport Layer Security
TS	Top Secret
UAAS	Unified Authorization and Attribute Services
UIAS	Unified Identity Attribute Set
UDDI	Universal Description, Discovery and Integration
UML	Unified Modeling Language
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
US	United States
UUID	Universal Unique Identifier
VIRT	Virtual Coverage
W3CDTF	World Wide Web Consortium Date Time Format
WARP	Web Based Access and Retrieval Portal
WCS	Web Coverage Service
WFS	Web Feature Service
WMS	Web Map Service
WSDL	Web Service Definition Language
XACML	eXtensible Access Control Markup Language
XML	Extensible Markup Language
XPath	XML Path Language
XPointer	XML Pointer Language
Xquery	XML Query
XSLT	XML Stylesheet Language for Transformations

Appendix F Bibliography

Bibliography

[1] ATOM.XML

Intelligence Community / Department of Defense Content Discovery & Retrieval Integrated Project Team. *Atom Data Encoding Specification for Content Discovery and Retrieval Result Sets*.

Available online Intelink-U at: <http://purl.org/IC/Standards/ATOM>

Available online at: <http://purl.org/IC/Standards/public>

[2] CDR-RA

Intelligence Community/Department of Defense Content Discovery & Retrieval Integrated Project Team. *Content Discovery and Retrieval Integrated Project Team Reference Architecture (CDR-RA)*.

Available online Intelink-U at: <http://purl.org/IC/Standards/CDR-RA>

Available online at: <http://purl.org/IC/Standards/public>

[3] CDR-RB

Intelligence Community/Department of Defense Content Discovery & Retrieval Integrated Project Team. *REST Data Encoding Specification for Content Discovery and Retrieval: Brokered Search (CDR-RB)*.

Available online Intelink-U at: <http://purl.org/IC/Standards/CDR-RB>

Available online at: <http://purl.org/IC/Standards/public>

[4] CDR-RS

Intelligence Community/Department of Defense Content Discovery & Retrieval Integrated Project Team. *REST Interface Encoding Specification for Content Discovery and Retrieval: Search (CDR-RS)*.

Available online Intelink-U at: <http://purl.org/IC/Standards/CDR-RS>

Available online at: <http://purl.org/IC/Standards/public>

[5] CDR-SD

Intelligence Community/Department of Defense Content Discovery & Retrieval Integrated Project Team. *SOAP Interface Specification for Content Discovery and Retrieval: Describe (CDR-SD)*.

Available online Intelink-U at: <http://purl.org/IC/Standards/CDR-SD>

Available online at: <http://purl.org/IC/Standards/public>

[6] CDR-SF

Intelligence Community/Department of Defense Content Discovery & Retrieval Integrated Project Team. *Data Encoding Specification for Content Discovery and Retrieval: Specification Framework (CDR-SF)*.

Available online Intelink-U at: <http://purl.org/IC/Standards/CDR-SF>

Available online at: <http://purl.org/IC/Standards/public>

[7] CDR-SS

-
- Intelligence Community/Department of Defense Content Discovery & Retrieval Integrated Project Team. *SOAP Interface Encoding Specification for Content Discovery and Retrieval: Search (CDR-SS)*.
Available online Intelink-U at: <http://purl.org/IC/Standards/CDR-SS>
Available online at: <http://purl.org/IC/Standards/public>
- [8] IC ITE INC1 IMPL
Office of the Director of National Intelligence. *Intelligence Community Information Technology Enterprise (IC ITE) Increment 1 Implementation Plan*. July 2012.
Available online Intelink-TS at: <http://go.ic.gov/HvBHBmY>
- [9] IC-ID.XML
Office of the Director of National Intelligence. *Text and XML Data Encoding Specification for Intelligence Community Identifier (IC-ID.XML)*.
Available online Intelink-U at: <http://purl.org/IC/Standards/ICID>
Available online at: <http://purl.org/IC/Standards/public>
- [10] ICD 501
Office of the Director of National Intelligence. *Discovery and Dissemination or Retrieval of Information within the Intelligence Community*. Intelligence Community Directive 501. 21 January 2009.
Available online Intelink-TS at: <http://go.ic.gov/GG61roi>
Available online at: http://www.dni.gov/files/documents/ICD/ICD_501.pdf
- [11] ICPG 710.1
Assistant Director of National Intelligence for . *Application of Dissemination Controls: Originator Control*. Intelligence Community Policy Guidance 710.1. 25 July 2012.
Available online Intelink-TS at: <http://go.ic.gov/yAqVQ0H>
- [12] ICS 500-20
Director of National Intelligence Chief Information Officer. *Intelligence Community Enterprise Standards Compliance*. Intelligence Community Standard 500-20. 16 December 2010.
Available online Intelink-TS at: <http://go.ic.gov/QUDIJKZ>
Available online Intelink-U at: https://intelshare.intelink.gov/sites/odni/cio/ea/library/Data%20Specifications/500-21/500_20_signed_16DEC2010.pdf
- [13] IETF-RFC 2119
Internet Engineering Task Force. *Key words for use in RFCs to Indicate Requirement Levels*. March 1997.
Available online at: <http://tools.ietf.org/html/rfc2119>
- [14] IETF-RFC 4287
M. Nottingham, R. Sayre. *The Atom Syndication Format*. December 2005.
Available online at: <http://www.ietf.org/rfc/rfc4287.txt>
- [15] Joint IC/DoD Memorandum
Intelligence Community Chief Information Officer, and Department of Defence Chief Information Officer. *IC and DoD Commitment to an Interoperable Service-Based Environment*. 13 July 2007.

Available online at: <http://www.docstoc.com/docs/797594/Department-of-Defense-DoD-and-Intelligence-Community-IC-Commitment-to-an-Interoperable-Services-Based-Environment---Enterprise-Services>

[16] SOAP

World Wide Web Consortium (W3C). *SOAP Version 1.2 Part 1: Messaging Framework*. W3C Recommendation 27 April 2007.

Available online at: <http://www.w3.org/TR/soap12-part1/>

[17] WS-A

World Wide Web Consortium (W3C). *Web Services Addressing 1.0 - SOAP Binding*. W3C Recommendation 9 May 2006.

Available online at: <http://www.w3.org/TR/ws-addr-soap/>

Appendix G Points of Contact

The Intelligence Community Chief Information Officer (IC CIO) facilitates one or more collaboration and coordination forums charged with the adoption, modification, development, and governance of IC technical specifications of common concern. This technical specification was produced by the IC CIO and coordinated with these forums, approved by the IC CIO or a designated representative, and made available at DNI-sponsored web sites. Direct all inquiries about this IC technical specification to the IC CIO, an IC technical specification collaboration and coordination forum, or IC element representatives involved in those forums.

Public Website: <http://purl.org/ic/standards/public>

E-mail: <ic-standards-support@intelink.gov> .

Appendix H IC CIO Approval Memo

An Office of the Intelligence Community Chief Information Officer (OCIO) Approval Memo should accompany this enterprise technical data specification bearing the signature of the Intelligence Community Chief Information Officer (IC CIO) or an IC CIO-designated official(s). If an OCIO Approval Memo is not accompanying this specification's version release package, then refer back to the authoritative web location(s) for this specification to see if a more complete package or a specification update is available.

Specification artifacts display a date representing the last time a version's artifacts as a whole were modified. This date most often represents the conclusion of the IC Element collaboration and coordination process. Once the IC Element coordination process is complete, the specification goes through an internal OCIO staffing and coordination process leading to signature of the OCIO Approval Memo. The signature date of the OCIO Approval Memo will be later than the last modified date shown on the specification artifacts by an indeterminable time period.

Upon signature of the OCIO Approval Memo, IC Elements may begin to use this specification version in order to address mission and business objectives. However, it is critical for IC Elements, prior to disseminating information encoded with this new specification version, to ensure that key enterprise services and consumers are prepared to accept this information. IC Elements should work with enterprise service providers and consumers to orchestrate an orderly implementation transition to this specification version in concert with mandatory and retirement usage decisions captured in the IC Enterprise Standards Baseline as defined in Intelligence Community Standard (ICS) 500-20.^[12]