

# CPACC

Caribbean Planning for the Adaptation to global Climate Change

## User Guide for searching the

## CPACC Geospatial



## Data Clearinghouse

Prepared for the **Centre for Geospatial Studies**  
by Rajkumar Singh

July 19, 2000

## 1.0 Background

This document was prepared to support the efforts started by the Caribbean Planning for Adaptation to Global Climate Change (CPACC) project in standardizing the creation of metadata for the twelve countries involved in the CPACC project. The first workshop that started this process was in September 13 - 17, 1999 in Trinidad. It was conducted by Ms. Michelle Anthony, Programmer Analyst, Computer Services, Raytheon, EROS Data Center, Sioux Falls. Since then metadata have been created for most of the countries involved in the project and made searchable on the World Wide Web (WWW) – better known as the web.

## 2.0 Introduction

Simply put metadata are 'data about data'. More specifically metadata are a structured means of describing the content, sources, condition, quality and other appropriate characteristics of acquired data or alternatively, metadata answer **who, what, when, where, why**, and **how** about every facet of the data that are being documented. *They do not, in any way, represent the actual content of the data - they only describe the data.*

The uses of metadata are:

- To tell others what datasets are available. This helps prevent duplication of efforts, especially among government agencies. Some organisations could end up sharing maintenance efforts with other organisations.
- Help organisations document (categorise and inventorise) the information they have. This helps prevent duplication by allowing organisations to easily locate their datasets. It would also provide a smooth transition for new employees – they would know what exist and why they were developed.
- Provide some legal protection in the event of misuse of datasets.
- Aid the receiving organisation/agency in understanding and interpreting the content of the datasets. Also allow them to update their data catalogue(s).
- To advertise and market spatial datasets.

Besides advertising in catalogues, one of the most convenient means of presenting metadata is by utilizing the power of the web. That way metadata can be updated as soon as the datasets change, provide the widest possible coverage and still not be a strain on financial resources.

To make metadata searchable over the web, the National Spatial Data Infrastructure (NSDI) of the United States established a Clearinghouse mechanism. This is a collection of servers in the United States that allow anyone to search for spatial datasets. To understand how this works take a look at the diagram below (Figure 1).

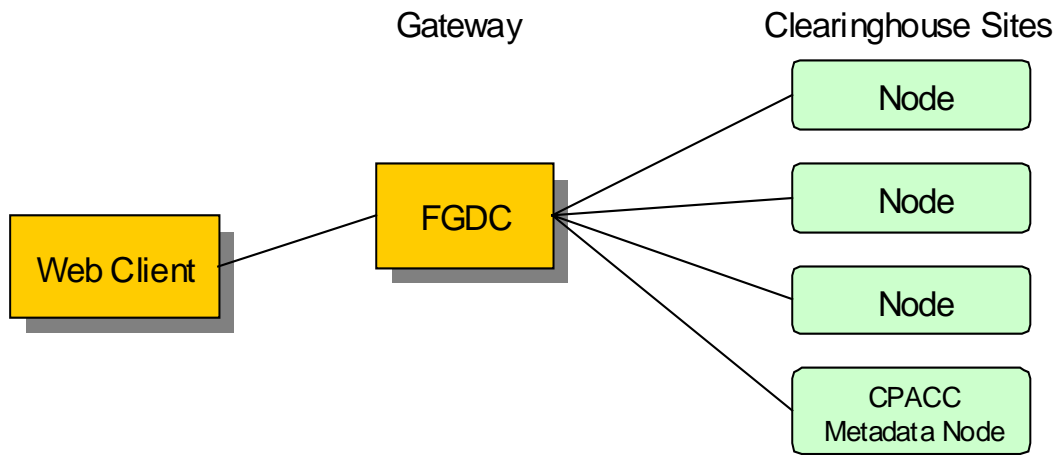


Figure 1: The Clearinghouse Mechanism

An example will be used to illustrate how this process works. To search the clearinghouse a person would go to the FGDC website using their web client (normally a web browser such as Netscape Navigator/Communicator or Microsoft Internet Explorer) and access the page with the search form. The person would then choose the appropriate options and enter keyword(s) to search for (e.g. roads, Belize, etc.). The request would be sent to the 'gateway' (one of the servers – about 6 in all) which would search the nodes<sup>1</sup> specified by the user for the keyword(s) and any other parameter specified in the search form and returns the results to the web client. The user can then view the details of the metadata record(s).

The remaining sections of this guide provide more detailed instructions on searching for spatial datasets through the clearinghouse mechanism.

### **3.0 Instructions on searching for spatial datasets**

#### **3.1 Requirements for performing a search**

A computer system with Internet connection and a web browser (Netscape Navigator/Communicator or Internet Explorer).

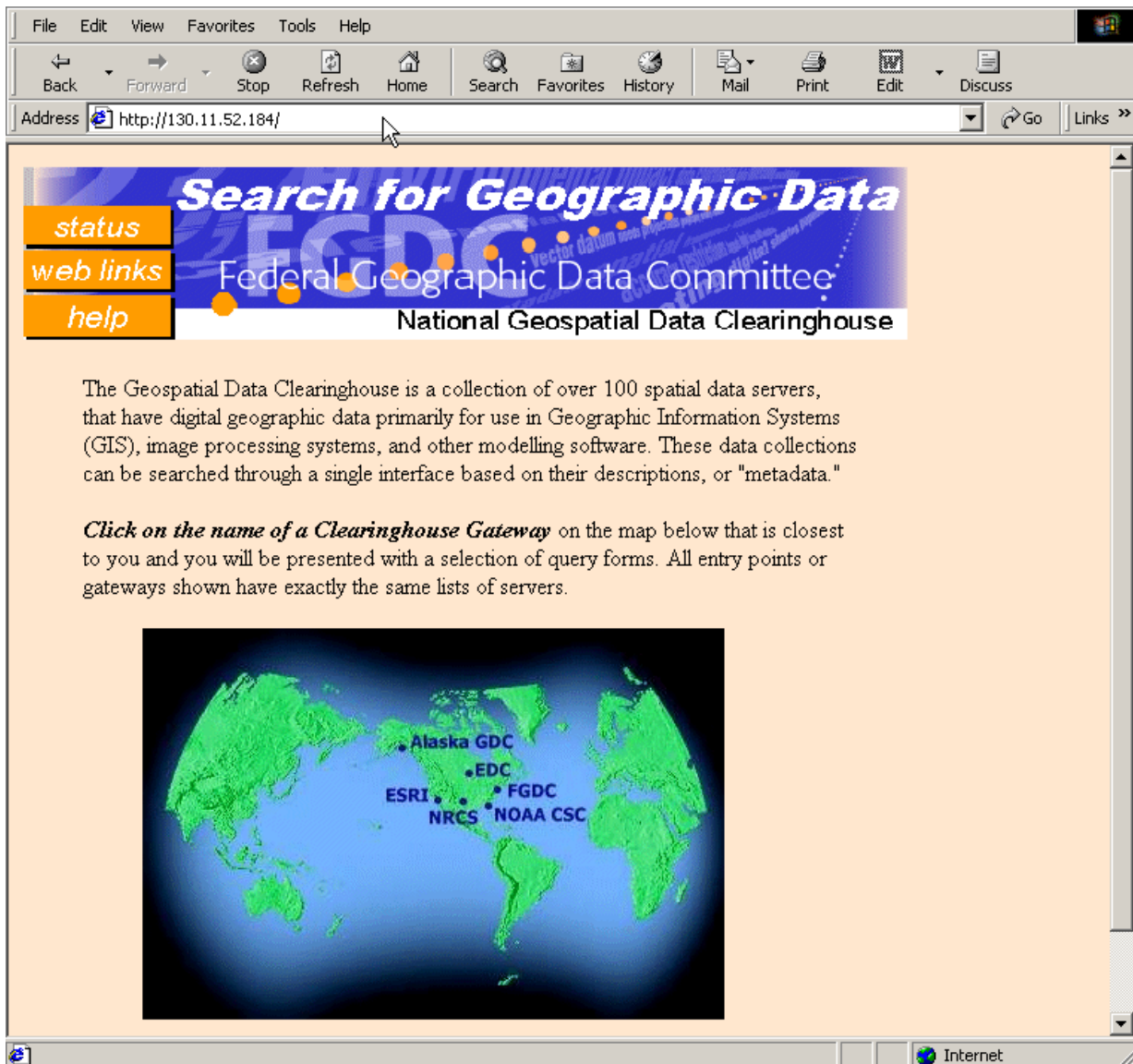
#### **3.2 Performing a search**

Enter the following URL<sup>2</sup> in your web browser:

<http://www.fgdc.gov/clearinghouse/clearinghouse.html> and then click on “[Search for Geospatial Data](#)”. To go directly to the search page use <http://130.11.52.184/> instead. A page like the one displayed below (Figure 2) would appear in your browser.

<sup>1</sup> Each server connects to individual nodes (or smaller servers) that house metadata databases.

<sup>2</sup> Acronym for Universal Resource Locator – it’s a unique address which identifies documents and media over the Internet, e.g. web pages



**Figure 2: The gateway page**

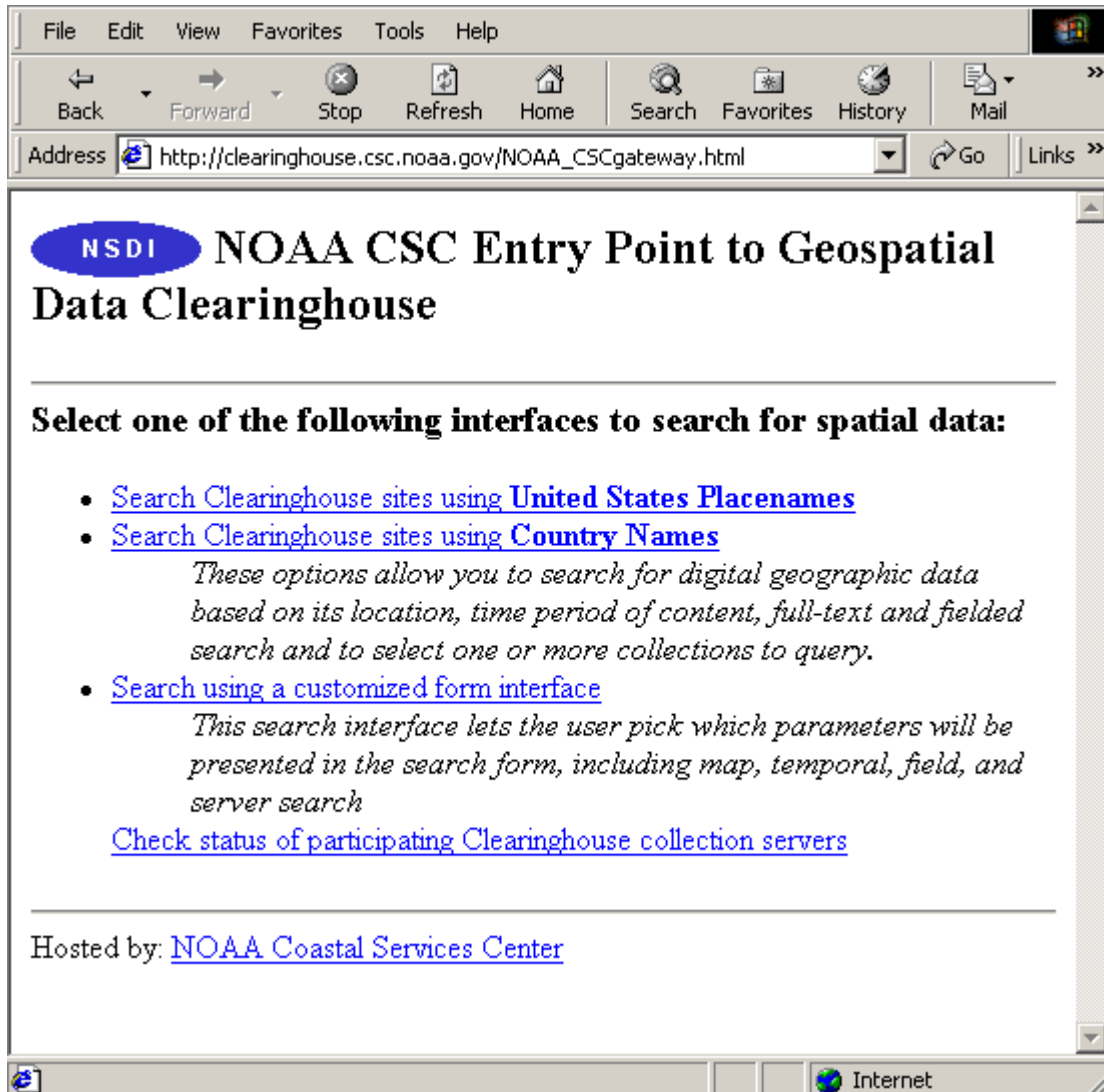
The image just below the text contains six hot spots<sup>3</sup>. Each hot spot (one of the names on the image above) points to a Gateway on the Clearinghouse. To search the CPACC metadata node choose the FGDC Gateway. Click on it. Next you should see the following page (Figure 3) in your web browser.

<sup>3</sup> A small part of a larger image that links to another web page. Clicking on a hot spot would transfer you to that web page.

For the Caribbean region ignore the first option “Search Clearinghouse sites using United States Placenames”. That leaves you with two options. However, at the time of this document’s preparation the second option was not functional. Therefore, we would only consider the first.

### 2.3 3.2.1 Searching the Clearinghouse sites using Country Names

Click on “Search Clearinghouse sites using Country Names”. The top of the web page, that appears, should look like the following (Figure 4).



**Define the Geographic Area of Coverage** Help...

*Specify a query region by selecting or entering values*

Don't search based on location

Use coordinates from a place name:

Cameroon  
 Canada  
**Caribbean**  
 Central African Republic

Enter bounding coordinates:

**North**

**West**                      **East**

**South**

**Specify Time Period of Content** Help...

*Specify a date or date range for desired spatial data by selecting **one** of the methods below.*

Don't search based on time period

Get data whose date is  the Date

Get data from    through

Figure 4: Top of the search page

### **3.2.1.1 Searching is done based on three criteria**

- A. Defining the Geographic Area of Coverage
- B. Specify Time Period of Content
- C. Search in Full-Text (Anywhere) or by Field

### **3.2.1.2 Criterion A: Defining the Geographic Area of Coverage presents three options:**

#### **1. Don't search based on location**

#### **2. Use coordinates from a place name**

*Select the place name (in Figure 4 above, Caribbean is selected)*

#### **3. Enter bounding coordinates**

*Enter the bounding coordinate of the area you are interested in. Values in degrees-minutes-seconds should be converted to Decimal Degrees. For example 6° 45" would be entered as 6.75°.*

### **3.2.1.3 Criterion B: Specify Time Period of Content presents three options:**

#### **1. Don't search based on time period**

#### **2. Search data whose date is "after, before or equal to" a specified date**

#### **3. Get data between two specified dates**

### 3.2.1.4 Criterion C: Search in Full-Text (Anywhere) or by Field (Figure 5)

This criterion allows you to search using keywords in any of four fields. These are standard fields

**Search in Full-Text (Anywhere) or by Field** Help...

*Specify search words by using one or more of the fields below.*

Search for:  in the field

in the field

in the field

in the field

Figure 5: Middle of the search page

that make up the FGDC metadata standard. A list of these fields and a brief description is presented below:

<b>Anywhere:</b> All of the fields listed below
<b>Edition:</b> The version of the dataset
<b>Source_Agency:</b> The Agency where the dataset can be found
<b>Purpose:</b> A summary of the intentions with which the dataset was developed
<b>Theme Keyword:</b> Common-use word or phrase used to describe the subject of the data set
<b>Stratum Keyword:</b> The name of a vertical location used to describe the locations covered by a dataset
<b>Entity_Type_Label:</b> The name of the entity type (e.g. schools, roads)
<b>Lineage:</b> Information about the events, parameters, and source data which constructed the dataset, and information about the responsible parties
<b>Title:</b> The name by which the dataset is known
<b>Originator:</b> The name of an organisation or individual that developed the dataset
<b>Abstract:</b> A brief narrative summary of the dataset
<b>Presentation_Form:</b> The mode in which the geospatial data is represented (atlas, diagram, globe, map, model, profile, remote-sensing image, section and view)
<b>Place Keyword:</b> The geographic name of a location covered by a dataset
<b>Temporal Keyword:</b> The name of a time period covered by a dataset
<b>Attribute_Label:</b> The name of the attribute [e.g. Sch_type (School Type), rd_type (Road Type)]

Let's consider an example on how to search for "roads in Belize". The keywords would be "roads" and "Belize". In this case you want **both** keywords to be present. To begin fill two of the input text boxes with the keywords, ensure that **AND** is selected at the left side and then proceed as describe below. See Figure 5 above. You may want to select 'anywhere' for 'roads' and 'Place\_Keyword' for 'Belize'. Note that in most cases you may end up refining your search since it is likely that you may not get any result. To relax the stringency of the search you should select **OR** instead. That way if either 'roads' **OR** 'Belize' is present you would get a result.

The bottom of the web page should look like the following (Figure 6).

To search the CPACC node you would select it as shown above. To search more than one node, hold down the **Ctrl** key and select the other node(s). The “Maximum Number of records to return from each selected database” and “Maximum Number of records to show on each results page” are self explanatory. Sometimes it may be a good idea to increase the “Maximum Number of seconds to wait for a server response” to 120 seconds especially if your Internet connection is slow. When you are finished click the “Search the Clearinghouse” button and wait for the search results. The page that is returned should look like the following (Figure 7).

### Select Data Servers to Search Help...

*Specify the data sources to query and the number of records that will be retrieved from each source. Use control key combinations to make multiple selections.*

**\* Disclaimer:** You may experience problems when selecting more than 40 servers in a single search session.  
*204 Clearinghouse Nodes Registered (6/19/00 2:42:14 PM)*

- Connecticut – Geospatial Data Clearinghouse
- Cornell University Geospatial Information Repository
- Costa Rica - Instituto Geografico Nacional - Centro Agronomico Tropical de Investigacion y Enseanz
- Costa Rica - Instituto Geografico Nacional de Costa Rica
- Costa Rica Biological Resource Maps (KU)
- CPACC Caribbean Coastal Resources Metadata Node**
- Crossnet Server - for SMMS3 product - sample data
- Delaware Geospatial Clearinghouse Node
- DOD Master Environmental Library
- Dominica - Geospatial Data for Land Management
- Dominican Republic clearinghouse for marine and coastal information
- Eastern Sierra Geospatial Data Clearinghouse
- Ecological Monitoring and Assessment Network (Environment Canada Server)
- El Salvador, CNR Instituto Geografico Nacional
- European Space Agency AVHRR Images

Maximum Number of records to return from each selected database:

Maximum Number of records to show on each results page:

Maximum Number of seconds to wait for a server response:

Figure 6: Bottom of the search page

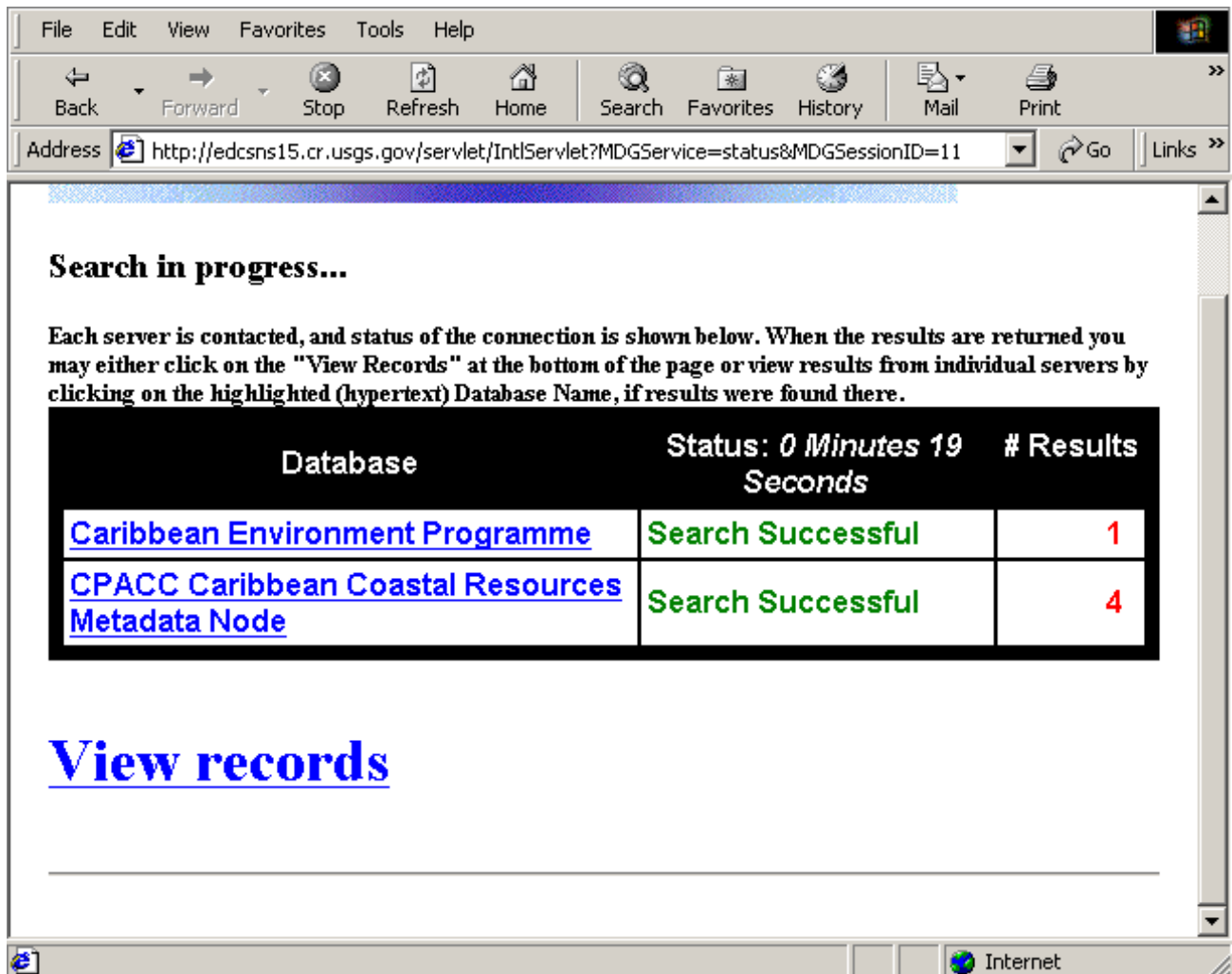


Figure 7: Summary of results

In the example two nodes were selected, "CPACC Caribbean Coastal Resources Metadata Node" and "Caribbean Environment Programme". These are displayed under the "Database" column. The results are displayed under the "# Results" column. What the above is telling us is that there are 1 metadata record on the "Caribbean Environment Programme" node and 4 metadata records on the "CPACC Caribbean Coastal Resources Metadata Node" node that contains the words 'roads' somewhere in the metadata records AND 'Belize' in the Place\_keyword field of the metadata records. The following page (Figure 8) shows what happens if the Internet connection is slow. One of the nodes failed. Failure of a node could also mean that the server where that node is found is temporarily unavailable. You can always check back at a later time.

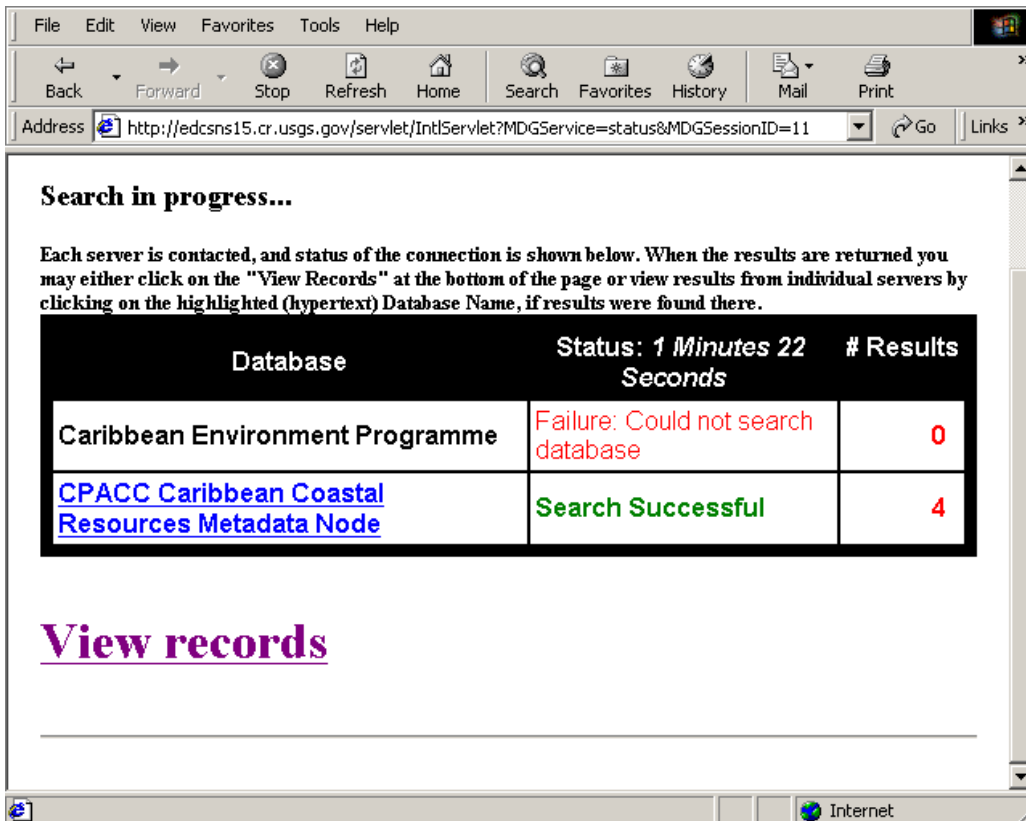


Figure 8: Summary of results with one failed node

Click on "View records" above and the following page (Figure 9) should appear.

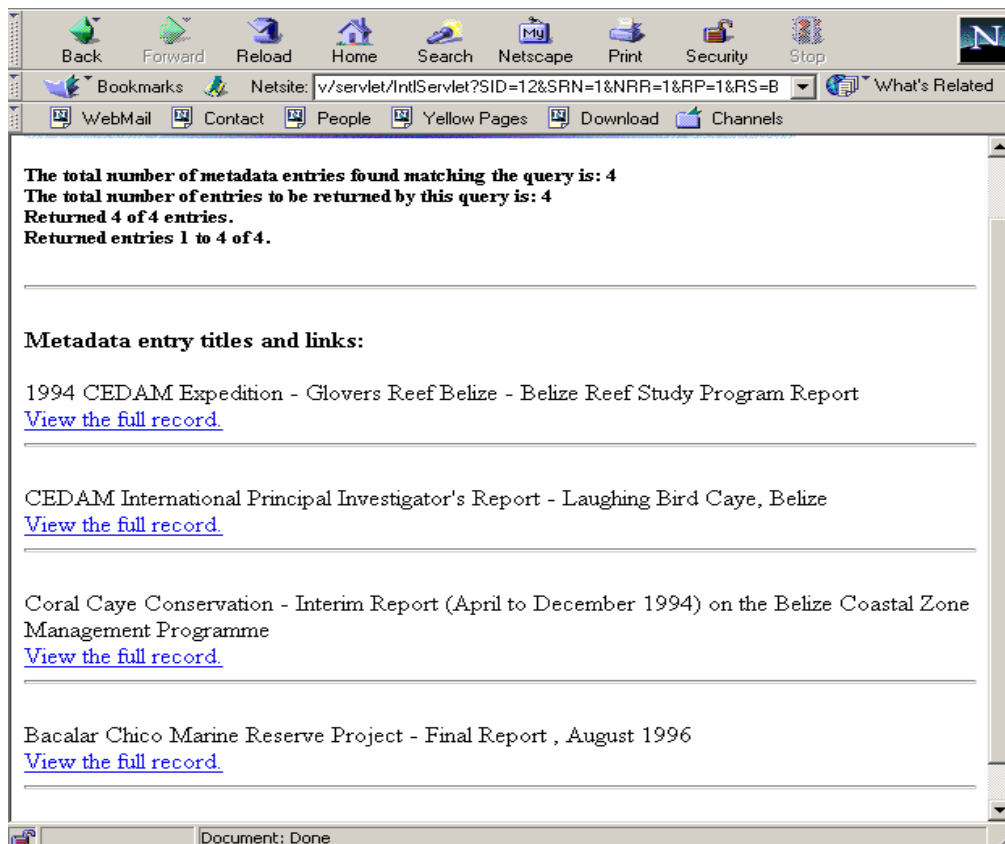


Figure 9: Results page

Clicking on “View records” returns all the metadata records from all the clearinghouse nodes listed under the ‘Database’ column. To view metadata records from individual nodes instead click on that node under the ‘Database’ column. The returned page (illustrated above) only gives a summary of the records. The description just above the ‘View the full record’ hyperlink<sup>4</sup> is taken from the title of the metadata record. To view the full record click on “View the full record”.

The above only illustrates one search example. Always feel free to experiment at first until you get the hang of it. HAVE FUN.

---

<sup>4</sup> An icon, graphic or word when clicked with the mouse automatically opens another file for viewing (e.g. another web page)

#### **4.0 References**

Federal Geographic Data Committee, 1995, *Content Standards for Digital Geospatial Metadata Workbook (June 8)*: Washington, Federal Geographic Data Committee.

Smart Computing, 1999, *Computing Dictionary – The illustrated book of terms and technologies*, 4<sup>th</sup> Edition: Sandhills Publishing, Lincoln, NE