



## KE EMu Documentation

# KE Software Portal

Document Version 1





# Contents

|                  |                                                                                    |           |
|------------------|------------------------------------------------------------------------------------|-----------|
| <b>SECTION 1</b> | <b>Introduction</b>                                                                | <b>3</b>  |
|                  | The Search page                                                                    | 4         |
|                  | Example Search                                                                     | 8         |
|                  | The Results List page                                                              | 11        |
| <b>SECTION 2</b> | <b>Administration of the Portal</b>                                                | <b>13</b> |
|                  | Caching of Data                                                                    | 13        |
|                  | Customization                                                                      | 14        |
|                  | How to disable/enable an existing Data Source                                      | 14        |
|                  | How to change the display name of a Data Source                                    | 14        |
|                  | How to change the Data Source color                                                | 14        |
|                  | How to change the fields returned by a search, set example values, map field names | 14        |
|                  | How to add a new Data Source                                                       | 15        |
|                  | How to change the look and feel                                                    | 16        |
|                  | XML responses                                                                      | 17        |
|                  | The Search page XML                                                                | 17        |
|                  | Results List XML                                                                   | 19        |



## SECTION 1

# Introduction

KE Software's Portal provides searching and collation of data from multiple sources simultaneously. Data Sources can be both internal and external to an institution, and include EMu, DiGIR and OAI, with the potential for many more.

The Portal has two modes of operation. At its core it is an XML query engine, comprising two customizable XML *responses*:

1. Query information: what data sources the Portal recognizes, what fields can be queried, etc.
2. Query response: what is returned in response to a query; this includes general information about the sources, a list of fields provided by the sources and a list of matching records.

This core XML component can be plugged in to another web or computer system as its data collection engine.

On the other hand, through implementation of XSL styling, the XML engine is a fully-fledged browser-based Search and Display utility for accessing multiple Data Sources. This is the default configuration of the Portal. The interface is easily modified and there is much potential to incorporate interesting and powerful features, such as record selection, linking and web mapping.

This document describes the default configuration of the Portal.

---

## The Search page

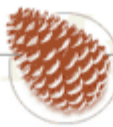
The following description applies to the default Portal Search page.



Keep in mind that it is a simple matter to customize the Portal browser to provide alternative functionality, a different interface, or even to dispense with the interface entirely if the Portal is used as part of another system.

On the Search page it is possible to:

- Select any of the specified Data Sources.
- Set the order of importance of each selected Data Source.
- Use a mixture of AND and OR query terms.
- Limit the number of records found.
- Specify how results will be sorted.


The Field  
Museum



## Search

### The Field Museum Data Portal

|                     |                                                                                                                                                   |
|---------------------|---------------------------------------------------------------------------------------------------------------------------------------------------|
| Status              | Local Database Access: on                                                                                                                         |
| Data Source(s)      | <b>POTENTIAL (select in priority order)</b> <b>ACTUAL (in priority order)</b>                                                                     |
|                     | <input type="checkbox"/> FMNH Botany EMu                                                                                                          |
|                     | <input type="checkbox"/> FMNH Fish EMu                                                                                                            |
|                     | <input type="checkbox"/> FMNH Invertebrates EMu                                                                                                   |
|                     | <input type="checkbox"/> FMNH Insects EMu                                                                                                         |
| Records             | Max records to display <input type="text" value="80"/>                                                                                            |
| Record Distribution | <input checked="" type="radio"/> Get records from as many sources as possible<br><input type="radio"/> Get records from sources in priority order |
| Timeout Per Source  | <input type="text" value="45"/> Seconds                                                                                                           |

|                                           | Family               | Genus                | Species              | Latitude             | Longitude            |
|-------------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Query Terms<br><i>blank terms ignored</i> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
|                                           | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
|                                           | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
|                                           | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
|                                           | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |

|      |                                                                                                                                                      |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sort | By <span style="border: 1px solid black; padding: 2px;">Family</span> in <span style="border: 1px solid black; padding: 2px;">Ascending</span> order |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------|

The following table describes each component of the Search page:

| Control                    | Description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Status</b>              | Indicates the status of the EMuWeb services on the EMu database server. If EMuWeb is not running, the system will not accept queries.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <b>Data Sources</b>        | Each searchable Data Source is listed here. Users select which Source(s) to search. The sequence in which Sources are selected determines the order in which results are displayed; this may be important depending on the number of records returned by the search (see <b>Records</b> and <b>Record Distribution</b> below).                                                                                                                                                                                                                                                                                                                                                                                                 |
| <b>Records</b>             | Specifies the maximum number of records to be returned by the search.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <b>Record Distribution</b> | <p>These options work in conjunction with <b>Records</b> and <b>Data Sources</b>:</p> <ul style="list-style-type: none"> <li>• <b>Get Records From As Many Sources As Possible</b><br/>The search will run against each Data Source selected by the user and will return records equally (as far as possible) from each Source up to the <b>Max Records To Display</b>.</li> <li>• <b>Get Records From Sources In Priority Order</b><br/>The search will run against the first Data Source that the user selected and will return as many records as possible; if this is less than the <b>Max Records To Display</b>, the second Data Source will be searched, and so on, up to the <b>Max Records To Display</b>.</li> </ul> |
| <b>Timeout Per Source</b>  | Specify how long (in seconds) to wait for a response from each Data Source. Searches of each Source are run simultaneously. Results will be displayed when all Sources have responded or when the timeout value has been reached.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| <b>Query Terms</b>         | <p>Specify the terms for which to search:</p> <ul style="list-style-type: none"> <li>• Entering values in a single row specifies a Boolean AND search (all values in that row must be found).</li> <li>• Each row is a Boolean OR search.</li> </ul> <p>For example:</p> <ol style="list-style-type: none"> <li>1. In the first row, values are entered in Family and Genus.</li> <li>2. In the second row a value is entered in Species.</li> </ol> <p>When the search is run, records that contain both values in the first row, or records that contain the value in the second row will be returned.</p>                                                                                                                   |



|                |                                                                                                 |                                                                            |
|----------------|-------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------|
| <b>Sort</b>    | Specify a field on which to sort results, and whether to sort in ascending or descending order. |                                                                            |
|                | The field and order specified here can be changed on the Results List page.                     |                                                                            |
| <b>Buttons</b> | <b>Search</b>                                                                                   | Commence the search.                                                       |
|                | <b>Fill With Example Values</b>                                                                 | Enter some typical search values appropriate to the Data Sources selected. |
|                | <b>Clear</b>                                                                                    | Clear any search terms specified.                                          |
|                | <b>Help</b>                                                                                     | Display user help.                                                         |

---

## Example Search

In this example we search for records with:

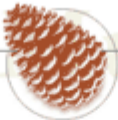
- A Family value of *Solanaceae* AND a Genus of *Solanum*
  - OR with a Genus of *Cestrum*
1. Select the Data Sources that you wish to search.



The order in which the Data Sources are selected determines the priority in which they are searched. In other words, if there are four Data Sources listed and you want the last one in the list to have priority in the search, select it first.

In this example we select the **FMNH Botany EMu** checkbox (and just to demonstrate the priority ordering of Data Sources, we then select **FMNH Insects EMu**, **FMNH Invertebrates EMu** and **FMNH Fish EMu**).

2. Specify the maximum number of records to display.  
The default is 80. In this case we enter 5000.
3. At Record Distribution, select (or accept if it is the default setting) **Get records from sources in priority order**.
4. Specify a Timeout Per Source.  
The default is 45 seconds and in this case we won't change that.
5. In the first row of the Query Terms table, enter *Solanaceae* in the Family cell AND *Solanum* in the Genus cell.  
In the second row, enter *Cestrum* in the Genus cell.
6. Select **Genus** in the Sort By drop list.  
The results will be sorted in alphabetical order by Genus.


The Field Museum



## Search

**The Field Museum Data Portal**

|                     |                                                                                                                                                   |                                                                                                                                                                                                                                |  |  |
|---------------------|---------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|--|
| Status              | Local Database Access: on                                                                                                                         |                                                                                                                                                                                                                                |  |  |
|                     | <b>POTENTIAL (select in priority order)</b>                                                                                                       | <b>ACTUAL (in priority order)</b>                                                                                                                                                                                              |  |  |
| Data Source(s)      |                                                                                                                                                   | <input checked="" type="checkbox"/> FMNH Botany EMu<br><input checked="" type="checkbox"/> FMNH Insects EMu<br><input checked="" type="checkbox"/> FMNH Invertebrates EMu<br><input checked="" type="checkbox"/> FMNH Fish EMu |  |  |
| Records             | Max records to display <input type="text" value="5000"/>                                                                                          |                                                                                                                                                                                                                                |  |  |
| Record Distribution | <input type="radio"/> Get records from as many sources as possible<br><input checked="" type="radio"/> Get records from sources in priority order |                                                                                                                                                                                                                                |  |  |
| Timeout Per Source  | <input type="text" value="45"/> Seconds                                                                                                           |                                                                                                                                                                                                                                |  |  |

|                                           | Family     | Genus   | Species | Latitude | Longitude |
|-------------------------------------------|------------|---------|---------|----------|-----------|
| Query Terms<br><i>blank terms ignored</i> | solanaceae | solanum |         |          |           |
|                                           |            | cestrum |         |          |           |
|                                           |            |         |         |          |           |
|                                           |            |         |         |          |           |
|                                           |            |         |         |          |           |

|      |                                                                                       |  |  |  |
|------|---------------------------------------------------------------------------------------|--|--|--|
| Sort | By <input type="text" value="Genus"/> in <input type="text" value="Ascending"/> order |  |  |  |
|------|---------------------------------------------------------------------------------------|--|--|--|

7. Select the **Search** button.

The Results List page displays.

Each of the selected Data Sources will be searched simultaneously up to the maximum timeout (45 seconds). The display of records is determined by:

- The priority order of the Data Sources (**FMNH Botany EMu** has priority over **FMNH Insects EMu** for instance)
- The Record Distribution option and maximum records to display setting.

In this case the Search results will be listed first from **FMNH Botany EMu** up to 5000 records. If less than 5000 matching records are found in the **FMNH Botany EMu** Data Source, matching records from the next Data Source will be listed, and so on until all matches are returned up to a maximum of 5000.

In this search, records that contain both *Solanaceae* AND *Solanum* (the record must contain both words, not just one of them) OR which contain *Cestrum* will be located.



The Results List page displays:

1. A summary of the Data Sources searched and a count of the records returned:

[Search Again](#)

| Source                 | Records | Status    |
|------------------------|---------|-----------|
| FMNH Botany EMu        | 3570    | completed |
| FMNH Insects EMu       | 0       | completed |
| FMNH Invertebrates EMu | 0       | completed |
| FMNH Fish EMu          | 0       | completed |

2. A table listing the returned records. The columns that display are specified in the XML and can easily be configured to suit an institution's requirements.

Click a column heading to change the sort order (ascending / descending) and the column on which the results are sorted.

Select the **Search Again** link above the Summary table to return to the Search page.

## SECTION 2

# Administration of the Portal

### Caching of Data

In order to reduce the load on the Data Sources, search results are cached for up to 15 minutes from the time that the data was last used. If the Data Source changes within the cache time-frame, results will not immediately reflect the Data Source changes.

It is possible to forcibly clear the cache from a browser by using the url:

```
http://remus.yvr.kesoftware.com/emuwebfmnh/webservices/  
lib/DataCacher.php?action=cleanCache
```

This part of the url:

```
http://remus.yvr.kesoftware.com/emuwebfmnh/
```

is institution specific.



Note that this will clear the cache for all users.

## Customization

In the following examples the Portal client is FMNH and the fetcher is NMNH DiGIR.

First we need to locate and edit the fetcher:

1. Locate the NMNH DiGIR fetcher - this is located in the Portal's fetcher directory, e.g.:  

```
cd ~/web/webservices/portal/fmnh/fetchers
```
2. An `ls` of this directory might return:  

```
DefaultTexxml.php  
FmnhBotTexxml.php  
FmnhFishesTexxml.php  
FmnhInsectTexxml.php  
FmnhIzTexxml.php  
NmnhDigir.php
```
3. Edit the `NmnhDigir.php` fetcher:  

```
vi NmnhDigir.php
```

### How to disable/enable an existing Data Source

Data Sources can be disabled/enabled temporarily on the back-end.

To enable or disable a Data Source that has already been specified, locate its fetcher in the Portal's fetcher directory and change the `enabled` flag to `true` or `false`. In this case we enable the NMNH DiGIR fetcher:

1. Locate the line that reads:  

```
var $enabled = false;
```

  
and change it to:  

```
var $enabled = true;
```
2. Save the file.  
The NMNH source will be added to the list of Data Sources on the Search page.

### How to change the display name of a Data Source

1. Locate and edit the following line:  

```
var $sourceName = "Name to change";
```

### How to change the Data Source color

1. Locate and edit a line similar to:  

```
var $preferredRGB = '#ffaaaa';
```

### How to change the fields returned by a search, set example values, map field names

1. Locate lines similar to:  

```
$this->setKnownConcept("Family","DarFamily");
```



```
$this->setKnownConcept( "Genus", "DarGenus", "Acacia" );
```

Each line that calls `setKnownConcept`:

- Adds a field.
- Specifies how to extract it from the Data Source's native list of fields.
- Specifies an example value (if any) for test queries.

The format of these lines is:

```
setKnownConcept( "Common Field Name", "Source Field", "Example  
(if any)" );
```

where:

"Common Field Name" is what the Portal will call the field.

"Source Field" is what the Data Source calls the field.

"Example" is an optional value for specifying a suggested value to run in test queries.



The Portal looks at all the "Common Field Names" of all enabled fetchers, and those that exist in all fetchers will be made available as search fields.

## How to add a new Data Source

This is achieved by writing a fetcher for the Data Source and placing it in the Portal's fetcher directory. KE Software supplies components for quickly setting up fetchers for EMu, DiGIR and OAI sources. It also supplies lower level components for building fetchers for other web services; assembling these can be done either in-house or by KE Software.

For example - to add a Texxml data source to an EMu database:

1. Find the host name of the database (e.g. the server host might be "hal2001.kesoftware.com").
2. Find the host's texxml port to which we will connect (e.g. 30003).
3. Establish a set of fields that can be used as searchable concepts, (e.g. Family, Genus, Species).
4. Map these concepts to the EMu database, e.g.:  

```
Family <=> DarFamily
Genus <=> DarGenus
Species <=> DarSpecies
```

It is probably easier to copy an existing Texxml fetcher and use this as the basis for a new one:

1. `cd ~/web/webservices/portal/fmnh/fetchers`
2. `cp DefaultTexxml.php MyShinyNewTexxml.php`
3. Edit `MyShinyNewTexxml.php` using the details located above:  

```
var $enabled = true;
var $hostname = "hal2001.kesoftware.com";
$this->port = 30003;
$this->setKnownConcept( "Family", "DarFamily" );
```

```
$this->setKnownConcept( "Genus", "DarGenus", "Acacia" );  
$this->setKnownConcept( "Species", "DarSpecies" );
```

## How to change the look and feel

KE Software provides a basic web browser interface for the Portal. More advanced functionality can be added though a combination of XSL, CSS and Javascript (or other web scripting tools).

The Portal returns XML in response to requests, and institutions can process this XML to generate the user interfaces they require.

Typically the default XSL stylesheets, CSS stylesheets (and any javascript) reside in:

```
~/web/webservices/portal/style
```

To customize any of these, place your files in:

```
~/web/webservices/portal/CLIENT/style
```

where:

CLIENT is your Portal name (e.g. fmnh, mnh, nmnh, mv, etc.).

To change how the query screen is generated from the Portal XML response, typically you would create a customized:

```
~/web/webservices/portal/CLIENT/style/portal_queryscreen.xsl
```

To change how the Results List page is created you would create a customized:

```
~/web/webservices/portal/CLIENT/style/portal.xsl
```

It is possible to configure the Portal to use a stylesheet of a different name for Results Lists by editing the PortalFactory file in:

```
~/web/webservices/portal/CLIENT/PortalFactory.php
```

and adding a line to the local Portal object, e.g.:

```
$defaultStylesheet = 'otherPortal.xsl';
```

It is also possible to use a different stylesheet by calling it with its name as a parameter. For example, for a Search page use:

```
http://HOST/emuwebCLIENT/webservices/portal.php?queryScreen=url_to  
_stylesheet
```

To display a Results page with an alternative stylesheet, pass the stylesheet as an argument to the Portal when making the search request.

## XML responses

### The Search page XML

This provides information about how the Portal can be queried and lists Data Sources and fields that can be queried. For example:

```
<xml>
  <queryScreen destinationUrl='/emuwebfmnh/webservices/portal.php'
    name='The Field Museum Data Portal' statusMapAvailable=''
    backendType='fmnh' localDbAccess='on'>
    <sources>
      <source name='FmnhBotTexxml' displayName='FMNH Botany EMu'
        suggestedPriority='0' foregroundRGB='#000000'
        backgroundRGB='#999933' icon='' />
      <source name='FmnhFishesTexxml' displayName='FMNH Fish EMu'
        suggestedPriority='1' foregroundRGB='#000000'
        backgroundRGB='#E7F5F6' icon='' />
      <source name='FmnhIzTexxml'
        displayName='FMNH Invertebrates EMu' suggestedPriority='2'
        foregroundRGB='#000000' backgroundRGB='#e8cecl' icon='' />
      <source name='FmnhInsectTexxml'
        displayName='FMNH Insects EMu' suggestedPriority='3'
        foregroundRGB='#000000' backgroundRGB='#elele5' icon='' />
    </sources>
    <queryableFields>
      <field name='Family:string' displayName='Family'
        type='string'>
        <example src='FmnhBotTexxml' />
        <example src='FmnhFishesTexxml' />
        <example src='FmnhIzTexxml'>Muricidae</example>
        <example src='FmnhInsectTexxml' />
      </field>
      <field name='Genus:string' displayName='Genus' type='string'>
        <example src='FmnhBotTexxml'>Acacia</example>
        <example src='FmnhFishesTexxml' />
        <example src='FmnhIzTexxml' />
        <example src='FmnhInsectTexxml' />
      </field>
      <field name='Species:string' displayName='Species' type='string'>
        <example src='FmnhBotTexxml' />
        <example src='FmnhFishesTexxml'>gracilis</example>
        <example src='FmnhIzTexxml' />
        <example src='FmnhInsectTexxml'>cariosa</example>
      </field>
      <field name='Latitude:float' displayName='Latitude'
        type='float'>
        <example src='FmnhBotTexxml' />
        <example src='FmnhFishesTexxml' />
        <example src='FmnhIzTexxml' />
        <example src='FmnhInsectTexxml' />
      </field>
      <field name='Longitude:float' displayName='Longitude'
        type='float'>
        <example src='FmnhBotTexxml' />
        <example src='FmnhFishesTexxml' />
        <example src='FmnhIzTexxml' />
        <example src='FmnhInsectTexxml' />
      </field>
    </queryableFields>
  </queryScreen>
</xml>
```

```
        </field>
    </queryableFields>
    <suggestedParameters>
        <maxPerSource>20</maxPerSource>
        <timeoutSeconds>45</timeoutSeconds>

    <displayStylesheet>/emuwebfmnh/webservices/portal/fmnh/style/portal.xml
    </displayStylesheet>

    <queryStylesheet>/emuwebfmnh/webservices/portal/fmnh/style/portal_queryscreen.xml
    </queryStylesheet>
    </suggestedParameters>
    </queryScreen>
</xml>
```

## Results List XML

This is returned in response to a query. It has some general information on the sources and their responses, a list of fields provided by the sources and a list of matching records. Some typical XML would be:

```
<xml>
  <mergedData>
    <statusBlock status='success'>
      <instance>Portal.8145.1196378088</instance>
      <systemName>The Field Museum Data Portal</systemName>
      <emuwebBase>emuwebfmnh</emuwebBase>
      <emuBackendType>fmnh</emuBackendType>
      <description>Set of up to 10 matching records sampling across
sources</description>
      <sources>
        <source name='FmnhFishesTexxml' status='completed'
sourceOrder='3' displayName='FMNH Fish EMu'
recordsUsed='10' translatorType='texxml'
backgroundRGB='#E7F5F6' foregroundRGB='#000000' icon='' />
      </sources>
      <groups>
        <group type='base'>description</group>
        <group type='extended'>ScientificName</group>
        <group type='extended'>recordSource</group>
        <group type='extended'>Class</group>
        <group type='extended'>Family</group>
        <group type='extended'>Genus</group>
        <group type='extended'>Species</group>
        <group type='extended'>TypeStatus</group>
        <group type='extended'>Latitude</group>
        <group type='extended'>Longitude</group>
        <group type='extended'>irn</group>
        <group type='extended'>Summary</group>
      </groups>
    </statusBlock>
    <records start='1' limit='10' count='10' sortBy='Family'
sortOrder='ascending'>
      <record index='1' sourceName='FMNH Fish EMu'
backgroundRGB='#E7F5F6' foregroundRGB='#000000' icon=''>
        <description>50644:Hyphessobrycon
gracilis:54::::Characidae:226:KPS1926-36:South
America:Argentina</description>
        <latitude>
        </latitude>
        <longitude>
        </longitude>
        <group name='ScientificName'>Hyphessobrycon
gracilis</group>
        <group name='recordSource'>FMNH Fish EMu</group>
        <group name='Family'>Characidae</group>
        <group name='Genus'>Hyphessobrycon</group>
        <group name='Species'>gracilis</group>
        <group name='irn'>629479</group>
        <group name='Summary'>50644:Hyphessobrycon
gracilis:54::::Characidae:226:KPS1926-36:South
America:Argentina</group>
      </record>
      <record index='2' sourceName='FMNH Fish EMu'
```

```
backgroundRGB='#E7F5F6' foregroundRGB='#000000' icon=''>
  <description>106795:Pimelodella
  gracilis:5::::Pimelodidae:261:B96:South
  America:Bolivia</description>
  <latitude>
</latitude>
  <longitude>
</longitude>
  <group name='ScientificName'>Pimelodella gracilis</group>
  <group name='recordSource'>FMNH Fish EMu</group>
  <group name='Family'>Pimelodidae</group>
  <group name='Genus'>Pimelodella</group>
  <group name='Species'>gracilis</group>
  <group name='irn'>685922</group>
  <group name='Summary'>106795:Pimelodella
  gracilis:5::::Pimelodidae:261:B96:South
  America:Bolivia</group>
</record>
</records>
</mergedData>
</xml>
```