

## FUNCTIONAL REQUIREMENTS

	FUNCTIONAL REQUIREMENTS	DESCRIPTION
1.	<p>The system should provide a variety of query building modes, to support novice users as well as highly technical power users. Typical user of the system include:</p> <p>Public</p> <p>Partners</p> <p>Internal Staff</p> <p>Oversight (OMB, Congress)</p>	<p>When constructing a query, the user should have the option to choose an appropriate query building method:</p> <p>With Menu Assisted mode, the system should provide toolbars or menu items of common operators so that the user can construct the query using a point and click method.</p> <p>With Advanced Command-Based mode, the user should be able to construct a query using the native syntax.</p> <p>With Saved Query mode, the user should be able to create new query by customizing a previously created query.</p>
2.	<p>The system should provide the ability to query a wide range of cross-program content including:</p> <ul style="list-style-type: none"> <li>▪ Projects</li> <li>▪ Proposals</li> <li>▪ Grants</li> <li>▪ Reviewers</li> <li>▪ Reviews</li> <li>▪ Plans of Work</li> <li>▪ Annual Reports</li> </ul>	
3.	<p>The system should provide the ability to query both current and historical information.</p>	<p>For example when searching projects, the user should be able to search active projects, terminated projects, or both.</p>

	FUNCTIONAL REQUIREMENTS	DESCRIPTION
4.	<p>The system should support a wide range of content format including</p> <ul style="list-style-type: none"> <li>▪ Databases</li> <li>▪ PDFs</li> <li>▪ Documents</li> <li>▪ Spreadsheets</li> <li>▪ Others</li> </ul>	<p>Though much of the agency information is captured in databases, there is still a significant amount of information in PDFs, spreadsheets, documents, stored on file servers. These include, for example, Plans of Work and Accomplishment reports from 2000-2006 captured in PDF files and Project Reviews in word documents. The system must provide the ability to search these files as well.</p>
5.	<p>There should be at least four independent steps to the search process: (1) Constructing the query, or asking the initial questions; (2) Getting the query results; (3) Refining the results; (4) Reporting the refined results.</p>	<p>The system should provide a range of features to streamline each of these steps.</p>
6.	<p>The system should provide the user with the ability to perform spelling checks on any word or phrase used to construct a query.</p>	<p>The query spelling checker suggests alternatively spelled words and allows the user to select the correct spelling.</p> <p>The spelling checker should use all the vocabulary contained in CSREES information systems, and works with proper names and technical words used by the agency. All actions must occur in real-time and are transparent to the user.</p>
7.	<p>The system should provide the user with the ability to perform queries using a ‘sounds like’ (“soundex”) function.</p>	<p>This enables the user to find other words that sound similar to the ones provided in a query, compensating not only for the user’s spelling errors, but also for errors within the data itself.</p>

	FUNCTIONAL REQUIREMENTS	DESCRIPTION
8.	The system should be able to automatically pick up common words with similar meanings and synonyms.	<p>The use of a custom synonym dictionary or “thesaurus” improves query results and assists the user by expanding query terms with related keywords that may not have occurred to the user but were intended and perhaps implicit.</p> <p>A thesaurus that can be custom-defined to include terms and phrases relevant to USDA in general, and CSREES in particular.</p>
9.	<p>The system should support wide range of query building features including:</p> <ul style="list-style-type: none"> <li>▪ Starts with, Ends with</li> <li>▪ Wildcarding (*,?, %, ...)</li> <li>▪ Stemming</li> <li>▪ Proximity searching (ex., “plants” within 2 words of “animals”)</li> <li>▪ Boolean operators</li> </ul>	
10.	The user should be able to save a query and share it with other users.	The system should allow the user to annotate the query describing its purpose and use, so that others can benefit from it.

	FUNCTIONAL REQUIREMENTS	DESCRIPTION
11.	The system should provide some form of relevancy ranking, so that the most relevant results are presented at the top of the results list.	<p>For example:</p> <p>Term frequency: How frequently a query term appears in a record should determine record's relevance.</p> <p>Location of terms: The location (data elements) — in which a term occurs— should indicate its significance to the record. Terms occurring in the title of a project that match a query term should be weighted more heavily than terms occurring in the detail summary of the project.</p> <p>Relative Proximity of query terms: When the terms in a query occur near to each other within a record; it should be more likely that the record is relevant to the query than if the terms occur at greater distance.</p>
12.	The system should provide the ability to display useful attributes such as relevancy rating and number of 'hits' (query words found) within each item in the results list.	
13.	The system should provide the ability to refine or narrow the results list by conducting another query on the first results list returned; i.e. a query within a query.	
14.	The system should provide the ability to “peek” at the contents of each item in the returned results list without going through the time-consuming process of clicking on each result to determine the relevancy of the content.	The words in the search query are used to generate a dynamic summaries based on the phrases

	FUNCTIONAL REQUIREMENTS	DESCRIPTION
15.	The system should provide the ability to sort/re-sort the results list by dynamically generated query attributes such as number of hits and relevancy ranking, as well as by a variety of data elements that are specific to the content or domain of information being queried upon.	
16.	The system should provide the ability to navigate through the results set by going to the first hit within unstructured data fields in the record or segment of the document, and thereon to the next or previous hits as required.	This means that the user can avoid having to read through large text fields of records before finding the relevant section, making it much more efficient.
17.	The system should provide the ability to identify hits by highlighting the keywords used in search criteria.	Hit highlighting is when the keywords in a record or document are highlighted in a different color. When combined with hit-to-hit navigation, hit highlighting enables the user to immediately see the relevant section of the record or document.
18.	The system should provide the ability to refine (narrow down or expand) the results list returned by drilling up and drilling down using a pre-defined hierarchical category or category tree.	For example, the user should be able to filter or narrow the results list returned to a specific region, or to specific state or to a specific institution type:
19.	The system tool should quickly adapt to new or updated data sources.	
20.	The system should provide the ability to tag relevant results, annotate them with notes for later search and reference, and extraction for further analysis or discovery.	
21.	The system should provide the ability to enterprise query results from variety of contents and time frames.	To account for information scattered across silos of data sources, the ability to federate results generated would allow user to have a complete picture of all relevant cross-program activities.

	FUNCTIONAL REQUIREMENTS	DESCRIPTION
22.	The system should provide the ability to save query results in a variety of common formats.	To further review and analyze query results, the user should have the ability to save or export query results to a variety of file types including excel spreadsheet, documents, and PDFs
23.	The system should provide the ability to export the results list, or a portion of the results list, for further review, analysis and manipulation.	
24.	The system should provide the ability to measure and analyze query activity and behavior of users to enhance the query process.	This in turn enables them to optimize the query experience.
25.	The system should provide the ability to restrict access to query results or functionality based on authenticated access level or permission	

# CRIS Basic Query

Select domain of information you would like to query:

- Query all active and recently terminated projects
  Query all historical projects (1998 - 2006)

Field Query:

**Project:**

Project Type:  ...

Project Status:  ...

**Classifications:**

Knowledge Area:  ...

Subject of Investigation:  ...

Field of Science:  ...

Program:  ...

FDC:  ...

Keywords:  ...

**Geographic Area:**

City:  ...

State:  ...

Region:  ...

Country:  ...

**Project Associated Years:**

Reporting Fiscal year:  ...

Award Fiscal Year:  ...

**Unique Identifiers of CRIS Projects:**

Accession Number:

Project Number:

Multi-state Project Number:

Proposal Number:

Award Number:

**Organization:**

Agency:  ...

Division Station:  ...

Department:  ...

Institution:  ...

Investigator:  ...

Select Knowledge Areas you would like to query for

Selection	Code	Knowledge Area
<input checked="" type="checkbox"/>	101	Appraisal of Soil Resources
<input checked="" type="checkbox"/>	102	Soil, Plant, Water, Nutrient Relationships
<input type="checkbox"/>	103	Management of Saline and Sodic Soils and Salinity
<input type="checkbox"/>	104	Protect Soil from Harmful Effects of Natural Element
<input type="checkbox"/>	111	Conservation and Efficient Use of Water
<input type="checkbox"/>	112	Watershed Protection and Management
<input type="checkbox"/>	121	Management of Range Resources
<input type="checkbox"/>	122	Management and Control of Forest and Range Fires
<input checked="" type="checkbox"/>	123	Management and Sustainability of Forest Resources
<input checked="" type="checkbox"/>	124	Urban Forestry
<input checked="" type="checkbox"/>	125	Agroforestry
<input checked="" type="checkbox"/>	131	Alternative Uses of Land
<input checked="" type="checkbox"/>	132	Weather and Climate
<input checked="" type="checkbox"/>	133	Pollution Prevention and Mitigation
<input type="checkbox"/>	134	Outdoor Recreation
<input type="checkbox"/>	135	Aquatic and Terrestrial Wildlife
<input type="checkbox"/>	136	Conservation of Biological Diversity
<input type="checkbox"/>	141	Air Resource Conservation and Management
<input type="checkbox"/>	201	Plant Breeding, Genome, Genetics, and Genetic Mec
<input type="checkbox"/>	202	Plant Genetic Resources and Biodiversity
<input type="checkbox"/>	203	Plant Biological Efficiency and Abiotic Stresses Affe
<input type="checkbox"/>	204	Plant Product Quality and Utility (Preharvest)

Apply Cancel

Text Query:

Projects that contain at least one of:  AND ALSO contain at least one of:  BUT DO NOT contain any of:

Query Expansion Options:

Use Synonyms
  Use Stemming
  Use Sounds Like

Submit Cancel

Your current Query is:

All active and recently terminated projects that meet the Text Query Criteria AND ALSO meet all of the following criteria

1) Knowledge Area is in (101, 102, 123-133)    2) Reporting Fiscal year is 2006    3) Agency is "CSREES"    4) State is in ("VA", "MD", "NJ")

# CRIS Advanced Query

Select domain of information you would like to query: \_\_\_\_\_

- Query all active and recently terminated projects
  Query all historical projects (1998 - 2006)

## Field Query

Data Field	Operator	Query Value(s)	Join Operator
<input checked="" type="checkbox"/> Knowledge Area	Does Not Match Any Of	101; 102; 123-133	And Also
<input checked="" type="checkbox"/> State	Matches At Least One Of	VA; MD; NJ; PA; NY	And Also
<input checked="" type="checkbox"/> Grant Award Amount	Greater Than	1,000,000	Or
<input checked="" type="checkbox"/> Project Length/Duration	Less Than	365	And Also
<input checked="" type="checkbox"/> Termination Date	Greater Than or Equal To	1/1/1998	And Also
<input checked="" type="checkbox"/> Termination Date	Less Than or Equal To	12/31/2004	

## Text Query

Data Field	Operator	Query Term	Operator	Query Term	Join Operator
<input checked="" type="checkbox"/> Objectives	Contains	Oak	Within 2 Words Of	Tree	And Also
<input checked="" type="checkbox"/> Impact Statement	Contains	improvement	Within The Same Sentence	resistance	
<input checked="" type="checkbox"/>					

Query Expansion Options:

- Use Synonyms
  Use Stemming
  Use Sounds Like

Submit

Cancel

Your current query is: \_\_\_\_\_

Query Translation: TBD



# Enterprise Query

Select domain of information you would like to query:

CRIS:

- Projects
- Proposals

C-REEMS:

- Proposals
- Reviewer Expertise

POW:

- |                                           |                                           |
|-------------------------------------------|-------------------------------------------|
| Plans of Work                             | Annual Reports                            |
| <input type="checkbox"/> Overviews        | <input type="checkbox"/> Overviews        |
| <input type="checkbox"/> Planned Programs | <input type="checkbox"/> Planned Programs |

Text Query:

Records that contain at least one of:

AND ALSO contain at least one of:

BUT DO NOT contain any of:

Query Expansion Options:

- Use Synonyms
- Use Stemming
- Use Sounds Like

Submit

Cancel

Your current query is:

All CRIS projects and proposals that contain at least one of the following ("sustainable", "green") and also contain at least one of the following ("energy", "fuels") but do not contain "corn".

# CRIS Query Results

Your query for "Organic" and "Farm" found:

[Active and recently terminated CRIS projects \(44\)](#)

[CRIS project history \(100\)](#)

View: [Detail](#) | [Summary](#)

1- 25 of 144 CRIS Projects

<< [Previous](#) | [Next](#) >>

	Accession No	Project Title	Institution	Hits	Related Records	
					POW	Awards
<input type="checkbox"/>	<a href="#">0085372</a>	Degradation of agricultural and natural toxins present in livestock	UNIV OF CALIFORNIA	11	<a href="#">1</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0081132</a>	Dynamics of Livestock Insects and Associated Vector-borne F	UNIV OF MINNESOTA	9	<a href="#">3</a>	<a href="#">1</a>
<input checked="" type="checkbox"/>	<a href="#">0089841</a>	USE OF NMR SPECTROSCOPY IN CONFORMATIONAL ANALY	TEXAS A&M UNIV	9	<a href="#">2</a>	<a href="#">1</a>
<input checked="" type="checkbox"/>	<a href="#">0068123</a>	GENETIC STUDIES IN VEGETABLE CROPS	OREGON STATE UNIVERSITY	8	<a href="#">3</a>	<a href="#">2</a>
<input checked="" type="checkbox"/>	<a href="#">0074045</a>	PLANT GENETIC RESOURCES CONSERVATION AND UTILIZAT	VIRGINIA POLYTECHNIC INSTITUT	6	<a href="#">1</a>	<a href="#">1</a>
<input checked="" type="checkbox"/>	<a href="#">0084840</a>	WEED CONTROL IN CROPLAND AND NON-CROPPED AREAS C	OREGON STATE UNIVERSITY	8	<a href="#">1</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0081625</a>	DEVELOPMENT OF CROPPING PRACTICES FOR PROFITABILIT'	OREGON STATE UNIVERSITY	10	<a href="#">1</a>	<a href="#">2</a>
<input type="checkbox"/>	<a href="#">0014308</a>	PLANT GENETIC RESOURCE CONSERVATION AND UTILIZATIC	OREGON STATE UNIVERSITY	4	<a href="#">3</a>	<a href="#">2</a>
<input type="checkbox"/>	<a href="#">0081620</a>	SUSTAINABLE MANAGEMENT FOR POTATO PRODUCTION AN	OREGON STATE UNIVERSITY	9	<a href="#">2</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0079420</a>	ANTHOCYANIN PIGMENTS AND POLYPHENOLICS IN FRUITS A	OREGON STATE UNIVERSITY	3	<a href="#">3</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0080957</a>	ROOTSTOCK AND INTERSTEM EFFECTS ON POME AND STON	OREGON STATE UNIVERSITY	6	<a href="#">1</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0057761</a>	BIOLOGICAL CONTROL OF WEEDS	OREGON STATE UNIVERSITY	8	<a href="#">3</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0056503</a>	Breeding and Genetics of Hazelnut	OREGON STATE UNIVERSITY	5	<a href="#">2</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0014885</a>	ORCHARD FLOOR MANAGEMENT PRACTICES FOR IMPROVIN	OREGON STATE UNIVERSITY	7	<a href="#">3</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0011635</a>	REPRODUCTIVE PERFORMANCE IN DOMESTIC RUMINANTS	OREGON STATE UNIVERSITY	7	<a href="#">1</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0082978</a>	THE NATIONAL ATMOSPHERIC DEPOSITION PROGRAM	UNIV OF MASSACHUSETTS	6	<a href="#">1</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0004446</a>	Weed Management in Horticultural Crops	OREGON STATE UNIVERSITY	9	<a href="#">1</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0030355</a>	Multistate Research Coordination, Northeastern Region	UNIV OF MASSACHUSETTS	4	<a href="#">3</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0087962</a>	ROOTSTOCK AND INTERSTEM EFFECTS ON POME AND STON	UNIV OF MASSACHUSETTS	11	<a href="#">2</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0071358</a>	Multistate Research Coordination, Southern Region	UNIVERSITY OF GEORGIA	9	<a href="#">3</a>	<a href="#">2</a>
<input type="checkbox"/>	<a href="#">0075198</a>	THE NATIONAL ATMOSPHERIC DEPOSITION PROGRAM (NADF	UNIVERSITY OF ILLINOIS	9	<a href="#">1</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0073727</a>	ROOTSTOCK AND INTERSTEM EFFECTS ON POME- AND STON	UNIVERSITY OF ILLINOIS	8	<a href="#">3</a>	<a href="#">1</a>
<input type="checkbox"/>	<a href="#">0079288</a>	MANAGEMENT OF GRAIN QUALITY AND SECURITY FOR WOF	UNIVERSITY OF ILLINOIS	6	<a href="#">2</a>	<a href="#">2</a>
<input type="checkbox"/>	<a href="#">0073640</a>	CONSERVATION, MANAGEMENT, ENHANCEMENT AND UTILIZ.	UNIVERSITY OF ILLINOIS	8	<a href="#">3</a>	<a href="#">2</a>

Select All

[Back](#)

[Customize Display](#)

[Save Query](#)

[Save List](#)

[Export Results](#)

[Create Report](#)

Refine Results (keyword):

In

[Query](#)

Knowledge Area

- 102 Soil, Plant, Water Nutrient (22)
- 212 Pathogens and Nematodes (55)
- 311 Animal Diseases (77)
- All Knowledge Area

Subject of Investigation

- 3299 Poultry, gneral/other (33)
- 3310 Beef cattle, live animal (70)
- Sheep, live animal (43)

Geographic Area

- Eastern Region (74)
  - VA (44)
  - MD (30)
- Weastern Region (30)
  - CA (15)
  - OR (15)
- More >>

[Apply](#)

[Clear](#)

## PROCESS REQUIREMENTS

The Enterprise Search process should be comprised of four distinct but inter-related sub-processes. Each process is designed to streamline the search process, to ensure the accuracy of the search results, and to maximize access to wide range of content critical to strategic and operational decision making.

Sub-Process 1: Construct the Query (ask the initial question)

Sub-Process 2: Execute the Query (get the initial results)

Sub-Process 3: Sift and Sort through Results (analyze & pinpoint the information)

Sub-Process 4: Generate the Results (create reports, export/save the result)

### **Sub-Process 1: Construct the Query (ask the initial questions):**

The process of constructing a query should be intuitive and cater to a diverse group of users from the very basic all the way through to highly technical power users. Whenever appropriate, the system should provide two different methods of constructing queries for each content area, each method enhanced by a range of features (such as, synonyms, sounds like, and stemming) to streamline the process.

- Basic Query
- Advanced Query

Basic Query: The process of constructing a basic query should require no prior knowledge of the content. The system should present users with a list of most commonly used search fields comprised of both standard data fields and free-form text fields. Each standard data field should be accompanied by a list of associated values for users to choose from. While users should have the option to pick one, multiple, or range of values from the list, they should also be able to enter the values directly if needed.

Advanced Query: The advanced query process should allow more experienced users to construct queries using Boolean operators and proximity searching, so that users who know exactly what they want can be extremely specific with their search. By making all data fields available for use in constructing the query, the advanced query process should allow users to better define, narrow searches and deliver better result sets.

### **Sub-Process 2: Execute the Query**

The process of executing the initial query should begin by examining the query syntax to ensure that users are immediately notified and prompted to correct any error in the query syntax. Once records that match the query criteria have been identified, the process should restrict access to the query results based on authenticated access level or permissions of the user. The query execution process should also provide the means for ranking of the results based on a pre-determined set of criteria to ensure that the most relevant results are presented first.

### **Sub-Process 3: Sift and Sort Through Results**

As the overall success of the query process is measured by the accuracy and efficiency of the query results, simply returning a list of records that match the query criteria will not be sufficient. The process must allow the user to pinpoint precise content. Therefore, the Sift and Sort process should include a range of features designed to streamline and continuously narrow the query through refinements based on keywords as well as structured and hierarchical filters. Furthermore, the process should allow for sorting /re-sorting the results list by dynamically generating search attributes such as the number of hits and relevancy as well as by a variety of calculated data elements.

### **Sub-Process 4: Generate Output**

The process of generating output for the query results should allow the user to save the query as well as the results (list of selected record identifiers) within the system; to export the information to a variety

# Enterprise Search Process Model

Thursday, February 14, 2008

