

# Finder module

Cornell Data Management Service Group  
and  
Cornell Information Technologies Custom Development

## Overview

The purpose of the Finder module is to help users choose among a number of similar related alternatives.

In our initial use case, we had 18-20 computer data storage alternatives available for researchers. Our Help Desk created a list of questions to guide the researchers to specify their requirements. The Finder module allowed the researchers to enter requirements and interactively narrow the field of 18-20 to perhaps 2-3 alternatives.

### Describe your data

Answer these questions to help identify data storage services that are suitable for your needs. Checking these boxes will change the list of available services. If you are uncertain how to answer, leave the question blank to maximize your resulting options.

Clear Answers

#### 1. What is the classification of your data? ⓘ

- ☐ Public
- ☒ Sensitive / Moderate Risk
- ☐ Confidential or Restricted / High Risk
- ☐ HIPAA-Regulated

#### 2. Do you need backups, snapshots or replication of your data? ⓘ

- ☒ I need one or more backup/snapshot copies of the data, and need to be able to restore data from previous points in time (high durability).
- ☐ I need to have replicate copies of the data to minimize downtime (high availability).

#### 3. How much data do you have and how fast will it grow? ⓘ

- ☐ Unlikely to exceed 1TB in 2 years
- ☒ Greater than 1TB or likely to exceed in 2 years

#### 4. Do you have special performance needs? ⓘ

- ☒ I am likely to have more than 1,000 files in a single directory within two years.
- ☐ My data interactions demand high transaction or transfer rates.

Select data storage services you would like to compare.

Select All

Clear Selections

<b>Amazon Web Services Elastic Block Store</b> Storage for use with Amazon EC2	<b>Amazon Web Services Elastic File System</b> Storage for use with multiple Amazon EC2 instances	<b>Amazon Web Services Glacier</b> Cloud based archival storage	<b>Amazon Web Services S3</b> Flexible, scalable object storage	<b>BioHPC Cloud</b> Storage for BioHPC lab computing services
<b>CAC Archival Storage</b> Single copy, non-mountable storage	<b>CAC Red Cloud Storage</b> Storage for Red Cloud compute instances	<b>CISER Research Servers</b> Storage for CISER computing environment	<b>CUGIR</b> Publicly shared geospatial data storage	<b>CUL eCommons</b> Publicly shared data repository
<b>Cornell Box</b> Online file sharing and collaboration	<b>Cornell Restricted Access Data Center</b> Storage for CRADC (confidential) computing environment	<b>EZ-Backup Static Storage</b> Archival storage and backup storage	<b>Kaltura Video Platform Service</b> Flexible, scalable video and multi-media storage (customizable)	<b>Kaltura Video on Demand</b> Video and multi-media storage
<b>LabArchives</b> Online electronic lab notebook	<b>Shared File Services</b> File sharing between users and computers	<b>Shared File Services - Confidential</b> File sharing between users and computers for confidential data (non-WCM)	<b>WCM Block Storage</b> High performance storage attached to centrally hosted servers (WCM only)	<b>WCM File Sharing</b> Secure network storage (NFS/CIFS) for research computing (WCM only)
<b>WCM Red Cloud Secure Storage</b> Secure storage for Red Cloud compute	<b>WCM Secure Remote Archive</b> Secure, single copy, non-mountable			

Figure 1 shows the Finder module with criteria checkboxes selected and the eligible results to choose from.

After the users clicks checkboxes to enter their criteria, ineligible results are greyed out. The user can click on remaining eligible results; these are highlighted and detailed characteristics of this smallest group are presented in a table to allow the researchers to make their final choice.

[Compare services that match your selected criteria](#)

Select All

Clear Selections

Amazon Web Services Glacier

WCM Block Storage





	Amazon Web Services Glacier	WCM Block Storage
<b>Brief Description</b>	Amazon Glacier is a secure, durable, and extremely low-cost cloud storage service for data archiving and long-term backup. Amazon Glacier provides three options for access to archives, from a few minutes to several hours.	Enterprise-level SAN for physical and virtual server-based applications such as databases whose architecture or performance requirements cannot be satisfied by network file mounts.
<b>Example Use</b>	Storing an archival copy of a static dataset and when the actively used, primary copy is stored elsewhere.	Running databases and other research computing workloads that require specialized filesystems or high performance.
<b>Cost</b> ⓘ	\$ < \$500/TB/Year	\$\$\$ > \$1000/TB/Year Cost dependent on backup configuration.
<b>Capacity</b> ⓘ	5 TB file size limit. No overall limit (costs incurred). No practical limit to number of files.	No practical file size limit. No practical overall limit (costs incurred). No practical limit to number of files.
<b>Access and Collaboration</b>	Extensive control using AWS IAM (Identity and Access Management) and Cornell Active Directory users and groups.	Restricted to WCM Community (CWID). Very fine grained control.
<b>Data Allowed</b> ⓘ	Allowed: Public data. Sensitive / moderate risk data. Allowed with special configuration: FERPA-protected data. Confidential or restricted / high risk data. Contact <a href="#">Amazon Web Services Glacier</a> for more information. Not Allowed: HIPAA-regulated data.	Allowed: Public data. Sensitive / moderate risk data. Confidential or restricted / high risk data. FERPA-protected data. HIPAA-regulated data. Ithaca researchers needing to store HIPAA-regulated data should contact Cornell's <a href="#">Office of Research Integrity and Assurance</a> before choosing a storage option.
<b>Durability (protection against data loss)</b> ⓘ	 High Very high barrier to accidental deletion of objects, and Glacier performs regular, systematic data integrity checks.	 Medium Can be configured to use external backup services (costs incurred).
<b>Availability (protection against down-time)</b> ⓘ	 High The service redundantly stores data in multiple facilities and on multiple devices within each facility.	 High Replication option protects against availability loss due to infrastructure failure. Software and hardware architecture provide availability through routine maintenance windows.

Figure 2 shows the resulting comparison table containing data in several categories.

Implementing this application in Drupal allows the service managers to easily edit the available services and supporting data without additional programming.

We believe other groups will have a similar need for a self-service product to guide end-users to narrow down a choice within a set of moderately complex alternatives.

## Planning

When using this module to help end-users choose between a number of complex alternatives, based on a number of criteria and the evaluation of information in several categories, expect that gathering the supporting information and designing useful questions will take a significant

amount of planning and time. Changing questions or answers or adding or deleting categories of information after services have been entered requires editing each of the services. We recommend starting with a simple example like that presented below to familiarize yourself with the data layout and editing workflow before beginning a “real” project.

## Install Finder module via Composer

Install Drupal 8 using Composer:

```
composer create-project drupal-composer/drupal-project:8.x-dev testdrupal
```

Require the Finder module:

```
cd testdrupal
composer require cubear/finder
```

Note: If you aren’t using composer to manage your Drupal 8 site, you can clone the module from: <https://github.com/CU-CommunityApps/CD-finder>

From the Extend administrative page (/admin/modules), install the Finder module and its dependencies.

## Setup of the Finder Module

- In the Control Type taxonomy (/admin/structure/taxonomy/manage/control\_type/overview) add two terms: "checkbox" and "radio.". See Figure 3.

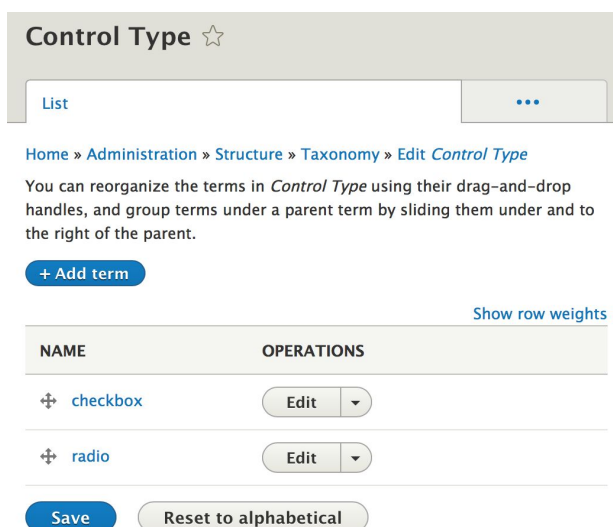


Figure 3.

- In the Facets taxonomy create a two-level hierarchy of terms representing criteria/questions and choices. For each of the criteria (top level facets), specify a control type: either “checkbox” (used when any, all or none of the choices may be selected) or “radio” (used when only one of the choices may be selected). See Figure 4.

The screenshot shows the 'Edit term' interface for the term 'Required Condiments'. At the top, there is a header 'Edit term' with a star icon. Below it are two tabs: 'View' and 'Edit', with 'Edit' being the active tab. A breadcrumb trail shows 'Home » Required Condiments'. The 'Name' field is labeled 'Name \*' and contains the text 'Required Condiments'. The 'Control Type' field is labeled 'Control Type \*' and has a dropdown menu showing 'checkbox (1)'. Below these fields is a section titled 'RELATIONS' with a right-pointing arrow. At the bottom, there are two buttons: 'Save' (in a blue pill shape) and 'Delete' (in a red pill shape).

Figure 4.

- Drag the Facet taxonomy terms into a two level hierarchy, representing the relationship between criteria/questions and choices. See Figure 5.

The screenshot shows the Facet taxonomy interface. It features a table with a header 'NAME' and several rows of terms. Each term is preceded by a blue icon consisting of two crossed arrows. The terms are: 'What must be on your burger?', 'Chicken', 'Hamburger', 'Pork', 'Required Condiments?', 'Ketchup', 'BBQ sauce', and 'Mayonaise'. At the bottom of the interface, there are two buttons: 'Save' (in a blue pill shape) and 'Reset to alphabetical' (in a grey pill shape).

Figure 5. Facet taxonomy terms in a two level hierarchy, where “What must be on your burger?” represents criteria and “Chicken,” “Hamburger” or “Pork” represent choices.

Services (the things being chosen) can contain a lot of data, and so require a bit of planning. Once the end-user has narrowed down their choices, they will be presented with a comparison table containing data in several fields. (See Brief Description, Example Use, Cost, Capacity, Access and Collaboration, etc. in Figure 2.) These fields are represented by a Drupal Paragraph called “Service Paragraphs”. Service Paragraphs should be edited before you start entering Service data at

`/admin/structure/paragraphs_type/service_paragraphs/fields`

When delivered, the Finder module comes with two fields in Service Paragraphs: “First Category” and “Second Category.” You can rename or replace these, and add as many fields as you wish. We have found that Text (formatted, long) fields work best.

**Manage fields** ☆

[Edit](#) [Manage fields](#) [Manage form display](#) [Manage display](#)

[Home](#) » [Administration](#) » [Structure](#) » [Paragraphs types](#) » [Service paragraphs](#)

[+ Add field](#)

LABEL	MACHINE NAME	FIELD TYPE	OPERATIONS
First Category	field_first_category	<a href="#">Text (formatted, long)</a>	<a href="#">Edit</a> ▼
Second Category	field_second_category	<a href="#">Text (formatted, long)</a>	<a href="#">Edit</a> ▼

*Figure 6. Fields in Service Paragraphs*

Now you may add Services. From the images above, you can see that a service has a title, summary (show on the cards in the first image. They are “eligible” to be chosen, depending on the choices the end-user has specified, and so need a set of “facet-matches” to indicate whether they are compatible with specific choices. They also have text data in several categories to be able to populate the comparison table shown in Figure 2.

Create your Services using the Drupal Service content-type. Add the title and summary for each as seen in Figure 7.

**Title \***

Chick-Fil-A

**Summary**

Eat moar chicken.

**Facet Matches**

- ☐ What must be on your burger?
- ☐ -Hamburger
- ☒ -Chicken
- ☐ -Pork
- ☐ Required Condiments
- ☐ -BBQ sauce
- ☐ -Ketchup
- ☒ -Mayo

*Figure 7 shows an example title and summary for each Service.*

Then fill in the text for the Service Paragraphs you have chosen.

Required: Add an additional Service with the title “Help” and put text in each field describing what to expect in that field. This will be used by the Finder module to display help information for the each row in a the comparison table. The comparison table may not appear if the Help Service is not defined.

## Tips

You may wish to remove all sidebar blocks from this page, so that the Finder can use the full width of the page.

The Finder module is displayed at the URL path `/finder`.

There is a configuration page to edit titles and other data on the Finder page at `/admin/config/content/finder`

Check out our production implementation at <http://finder.research.cornell.edu>

You will need to configure the SMTP module to allow the Finder module to send mail.