Unified Data Catalog for Imaging Analytics

Scenario 1: Create Tags from AutoTag Policy, Custom-Tagging and Filtering

Set up Spectrum Discover to Manage Data Sources	

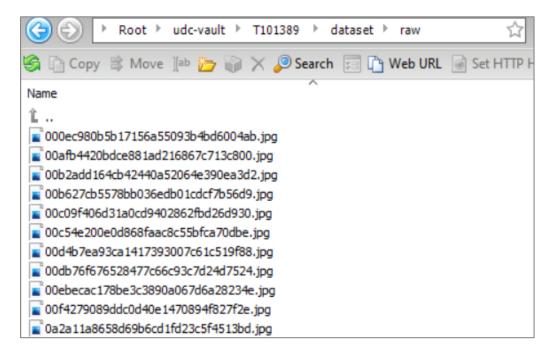


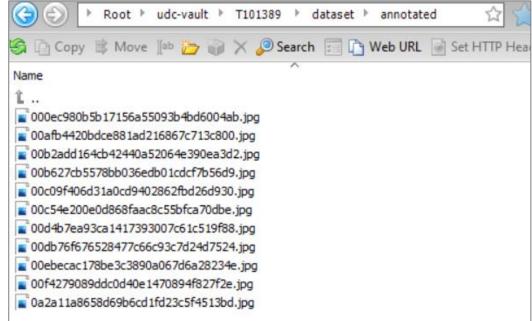
Data source management

Connections Scan history Discover databases

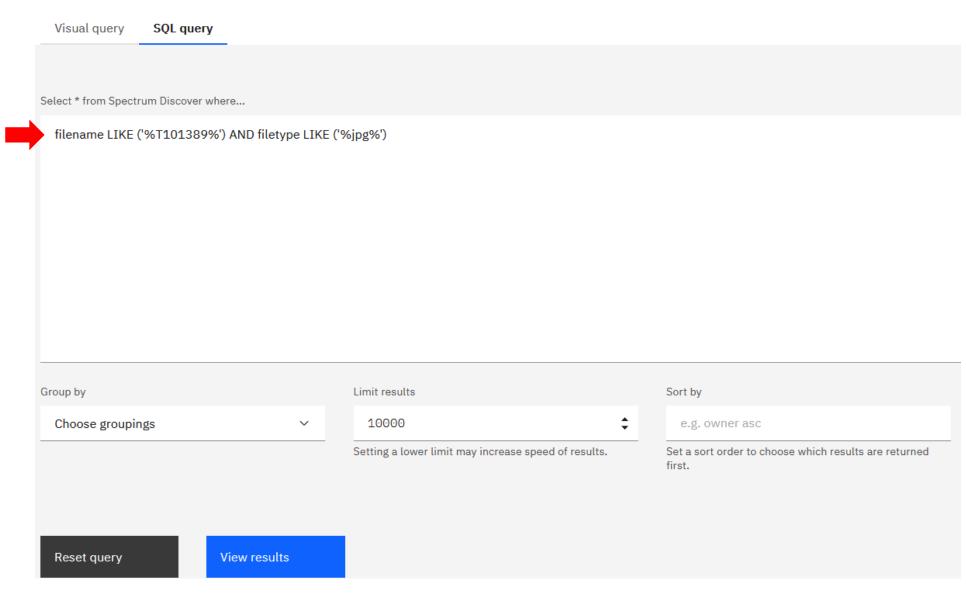
							Q	Add conne	ection	+
Connection name	Connection type	Cluster	Data source	Site	State	Scan status		Ne	ext scan	
UDCt-ClockTower-ScaleS3	\$3	http://129.40.101.35	BreaKHis_v1	ClockTower	Online					
UDCt-haproxy-rhceph01- rgw0	\$3	https://9.11.221.111:443	udcuser-bucket		Online	•				
UDCt-rhceph01-rgw0	\$3	http://rhceph-01.storage- ocp.tuc.stglabs.ibm.com:8080	udc-RGW-bucket		Online	•				
UDCt-rhceph03-rgw0	\$3	rhceph-03.storage-ocp.tuc.stglabs.ibm.com	udc-RGW3- bucket		Online					
UDSt-CoS	IBM COS	df3b7ae7-ab82-756b-01c4-c9f3ac8fae42	udc-vault	Tucson	Online	•				
UDSt-SpectrumScale	Spectrum Scale	scale-bm.storage-ocp.tuc.stglabs.ibm.com	bm_filesystem_2		Online					
Items per page: 20 ∨	1-6 of 6 items						1 ~	of 1 pages	4	>

2. Upload Dataset to Object Storage





Query builder



Results view: ③					Matchea	l 22 records from n	netaocean_view table in 0	.084276 s	second
Condensed Individual								Q	©
Condensed Individual	path	filenam	e	udc dem2	u2 type	u2 id	udc dem3	u2 wid	dth
Inspecting records	udc- vault/		89/dataset/annotated 980b5b17156a55093b4bd6004ab.jpg	NULL	NULL	NULL	NULL	NI	ULL
Grouped By	udc- vault/		89/dataset/annotated 1420bdce881ad216867c713c800.jpg	NULL	NULL	NULL	NULL	NI	ULL
Cluster Datasource	udc- vault/		39/dataset/annotated dd164cb42440a52064e390ea3d2.jpg	NULL	NULL	NULL	NULL	NI	ULL
Filegroup Filename	udc- vault/		89/dataset/annotated 7cb5578bb036edb01cdcf7b56d9.jpg	NULL	NULL	NULL	NULL	NI	ULL
Fileset	udc- vault/		89/dataset/annotated f406d31a0cd9402862fbd26d930.jpg	NULL	NULL	NULL	NULL	NI	ULL
Show more	udc- vault/		89/dataset/annotated e200e0d868faac8c55bfca70dbe.jpg	NULL	NULL	NULL	NULL	NI	ULL
	udc- vault/		89/dataset/annotated 7ea93ca1417393007c61c519f88.jpg	NULL	NULL	NULL	NULL	NI	ULL
	udc- vault/		89/dataset/annotated 6f676528477c66c93c7d24d7524.jpg	NULL	NULL	NULL	NULL	NI	ULL
	<								>
	Items per page:	10 ∨	1–10 of 22 items				1 v of 3 pages	4	٠

3. Set up AutoTag Policy in Spectrum Discover

Create the autotag policy in json

```
{
    "pol_id": "T101389_autotag_u2source_pol",
    "action_id": "AUTOTAG",
    "action_params": {
        "rule": {
            "name":"setFromExistingField",
            "action": "regex",
            "existingField": "filename",
            "regexPattern": "[^/]+",
            "matchNo": 3,
            "newField": "u2-source"
        }
    },
    "schedule": "NOW",
    "pol_filter": "filename LIKE ('%T101389%') AND filetype LIKE ('%jpg%')",
    "pol_state": "active",
    "explicit": "true",
    "collections": []
}
```

Execute the policy on server

```
[moadmin@discover1 T101389]$ pwd
/home/moadmin/T101389
[moadmin@discover1 T101389]$ ls
T101389 aircraft import scenario1.json T101389 autotag u2source.json
T101389 aircraft import scenario2.json
[moadmin@discover1 T101389]$
[moadmin@discover1 T101389]$ gettoken
                                                                Time Current
          % Received % Xferd Average Speed
                                               Time
  % Total
                                                       Time
                                Dload Upload
                                                                Left Speed
                                              Total
                                                       Spent
                                          0 --:--:--
[moadmin@discover1 T101389]$
[moadmin@discover1 T101389]$
[moadmin@discover1 T101389]$ curl -k -H "Authorization: Bearer ${TOKEN}" https://localhost/policyengine/v1/
policies/T101389 autotag u2source pol -X POST -d @./T101389 autotag u2source.json -H "Content-Type: applica
tion/json"
Policy 'T101389 autotag u2source pol' added[moadmin@discover1 T101389]$
```

Tag management

Tags **Policies** Policy history

							Q Add Policy +
Policy	Туре	Schedule (UTC)	State	State	Collections	Last modified by	Last modified
ap_cs_body	CONTENTSEARCH	Done	Active	Complete		sdadmin	9/17/2021, 2:51:16 PM
ap_cs_meta	CONTENTSEARCH	Done	Active	✓ Complete		sdadmin	9/16/2021, 1:18:06 PM
T101389_autotag_u2source_pol	AUTOTAG	Done	Active	Complete		sdadmin	9/21/2021, 2:32:38 PM
u3-iphone-gps	CONTENTSEARCH	Done	Active	Complete		sdadmin	9/17/2021, 12:35:42 AM
Items per page: 20 ∨ 1–4 of	of 4 items					1	. ✓ of 1 pages

Tag management

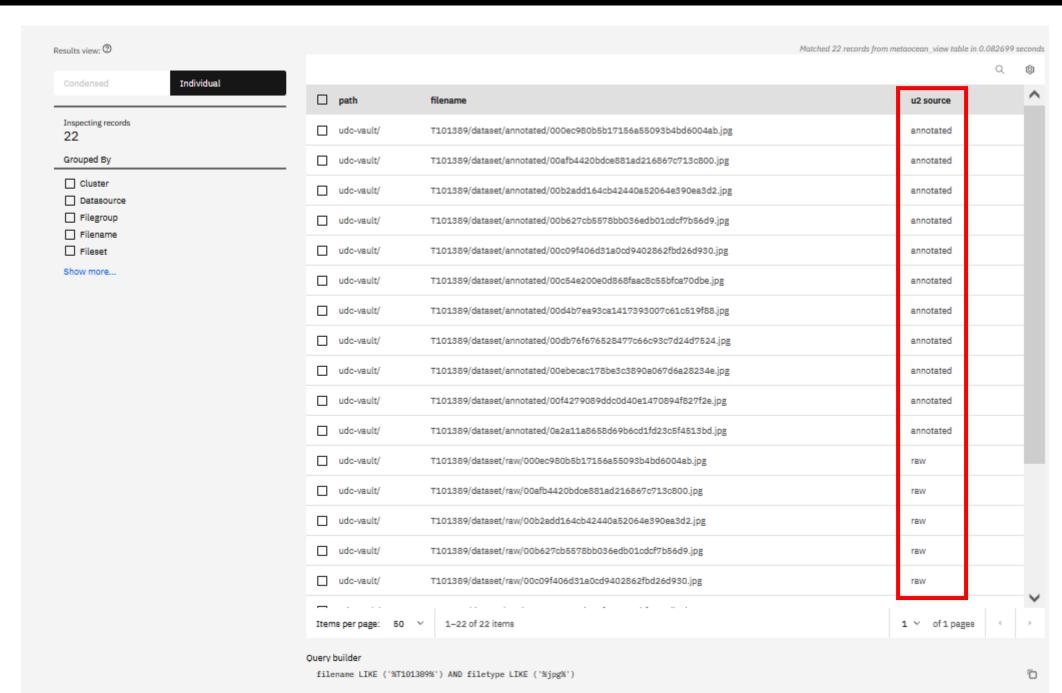
Tags

Policies

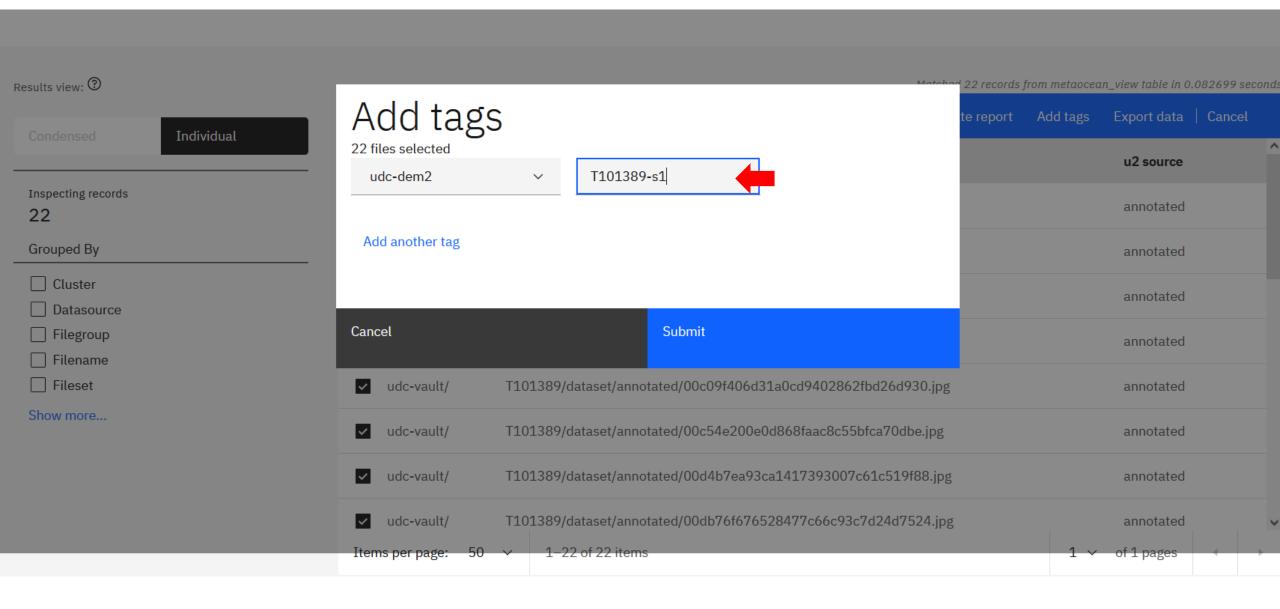
Policy history

								Q	(3)
Policy name	Туре	Started	Ended	Status	Total records	Successful	Skipped	Failed	
T101389_autotag_u2source_pol	AUTOTAG	9/21/2021, 2:32:39 PM	9/21/2021, 2:32:41 PM	Complete	22	22	0	0	
T101389_contentsearch_u2source_pol	AUTOTAG	9/21/2021, 2:27:44 PM	9/21/2021, 2:27:46 PM	Complete	22	22	0	0	
T101389_aircraft_import_scenario2_pol	IMPORT_TAGS	9/21/2021, 1:03:52 AM	9/21/2021, 1:04:11 AM	Complete	22	22	0	0	
T101389_aircraft_import_scenario1_pol	IMPORT_TAGS	9/20/2021, 11:21:32 PM	9/20/2021, 11:21:49 PM	Complete	11	11	0	0	
T101389_aircraft_import_scenario1_pol	IMPORT_TAGS	9/20/2021, 11:17:06 PM	9/20/2021, 11:17:22 PM	Complete	11	11	0	0	
T101389_aircraft_import_pol	IMPORT_TAGS	9/20/2021, 9:18:41 PM	9/20/2021, 9:18:59 PM	Complete	11	11	0	0	
T101389_aircraft_import_pol	IMPORT_TAGS	9/20/2021, 9:16:06 PM	9/20/2021, 9:16:13 PM	Complete	0	0	0	0	
T101389_aircraft_import_pol	IMPORT_TAGS	9/20/2021, 9:08:05 PM	9/20/2021, 9:08:40 PM	Complete	0	0	0	0	
an cs hody	CONTENTSFARCH	9/17/2021,	9/17/2021,	Complete	16	16	n	n	

4. Custom-tag the New Dataset



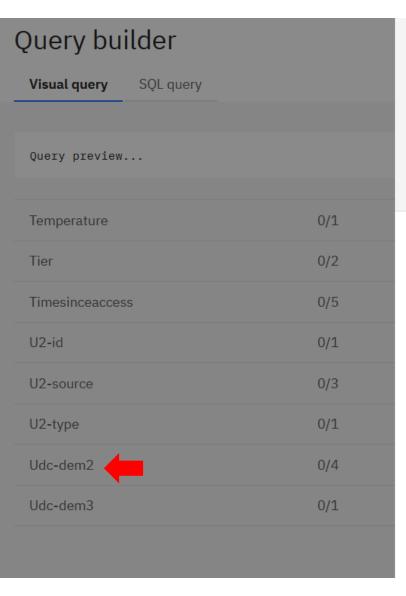
ults view: ①			Matched 2	22 records from naced	ın_view table in 0.	082699 sec
Condensed Individual	22 items select	ed	Generate	e report Add tags	Export data	Cancel
Mariada	path	filenan	е		u2 source	
nspecting records	✓ udc-vault/	T10138	39/dataset/annotated/000ec980b5b17156a55093b4bd6004ab.jpg	5	annotated	
rouped By	udc-vault/	T10138	39/dataset/annotated/00afb4420bdce881ad216867c713c800.jpg		annotated	
Cluster Datasource	✓ udc-vault/	T10138	39/dataset/annotated/00b2add164cb42440a52064e390ea3d2.jpg		annotated	
☐ Filegroup ☐ Filename	✓ udc-vault/	T10138	39/dataset/annotated/00b627cb5578bb036edb01cdcf7b56d9.jpg		annotated	
Fileset	✓ udc-vault/	T10138	39/dataset/annotated/00c09f406d31a0cd9402862fbd26d930.jpg		annotated	
ow more	✓ udc-vault/	T10138	39/dataset/annotated/00c54e200e0d868faac8c55bfca70dbe.jpg		annotated	
	✓ udc-vault/	T10138	39/dataset/annotated/00d4b7ea93ca1417393007c61c519f88.jpg		annotated	
	✓ udc-vault/	T10138	39/dataset/annotated/00db76f676528477c66c93c7d24d7524.jpg		annotated	
	Items per page:	50 ~	L–22 of 22 items	1 ~	of 1 pages	4

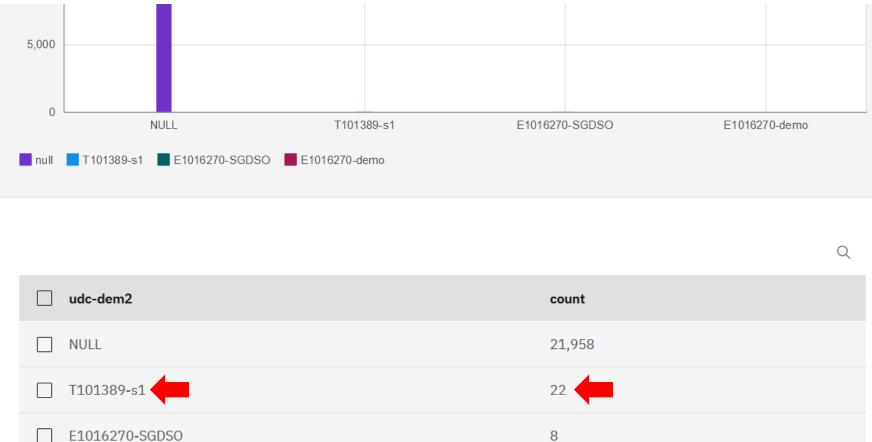


E1016270-demo

1-4 of 4 items

Items per page: 10 ∨

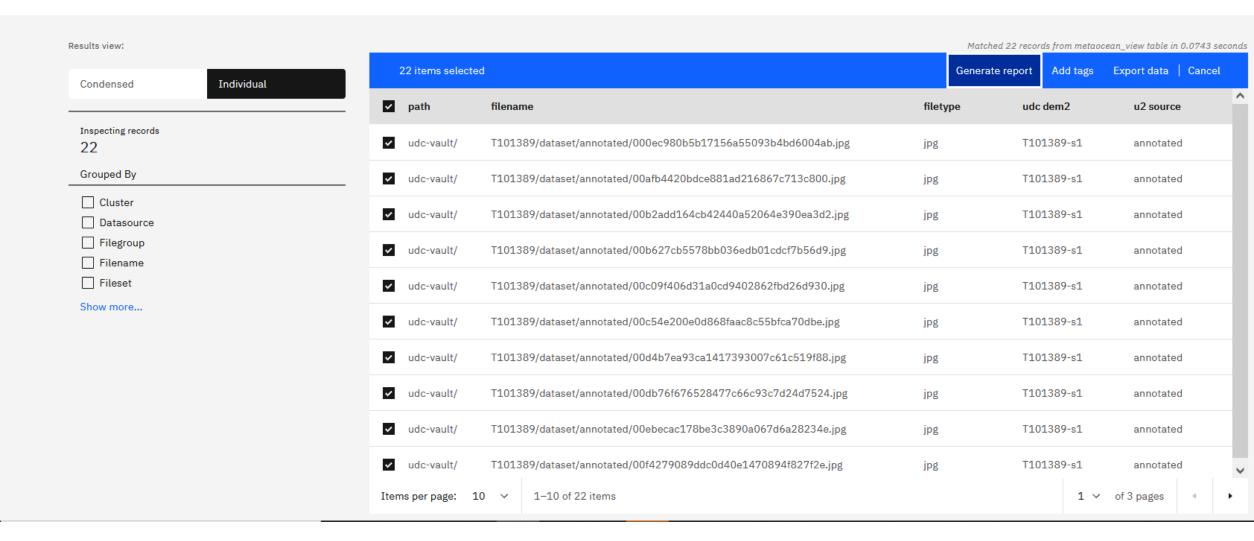


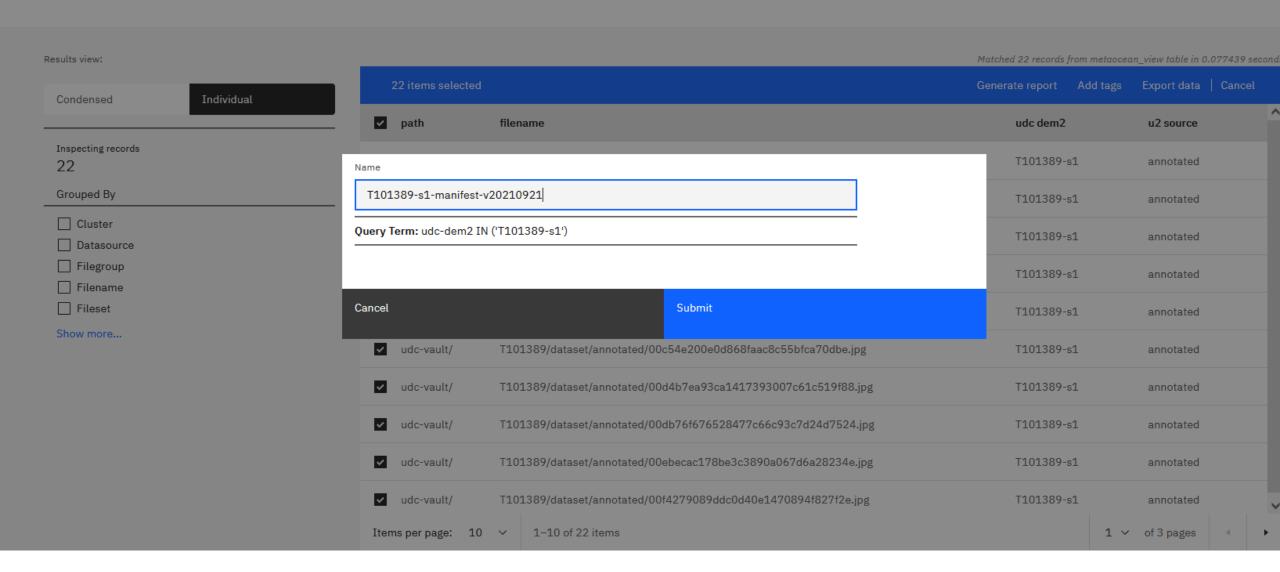


1

1 ∨ of 1 pages







Reports

Report has been Created
The report T101389-s1-manifest-v20210921 was
created

Report	Last run	Duration	Status	Output size	^
T101389-s1-manifest-v20210921	Tue Sep 21 2021 14:49:17 GMT-0500 (Central Daylight Time)	0 seconds	Complete	10.32 KiB	
demo-csv	Thu Sep 09 2021 18:39:49 GMT-0500 (Central Daylight Time)	0 seconds	Complete	2.34 KiB	
E101475_d1-list	Thu Apr 29 2021 08:16:33 GMT-0500 (Central Daylight Time)	0 seconds	Complete	8.17 KiB	
E1014753-data-manifest	Thu Apr 29 2021 00:36:52 GMT-0500 (Central Daylight Time)	0 seconds	Complete	8.08 KiB	
BH01 Dataset Report for FS21EMEA	Tue Feb 09 2021 00:14:30 GMT-0600 (Central Standard Time)	0 seconds	Complete	3.75 MiB	
TD-BH01 Report	Sat Jan 30 2021 18:15:17 GMT-0600 (Central Standard Time)	0 seconds	Complete	3.73 MiB	
BreakHist Summary	Sat Jan 30 2021 17:56:08 GMT-0600 (Central Standard Time)	0 seconds	Complete	3.33 MiB	
Age Report Detail. 90-180 Days	Tue Dec 22 2020 02:24:47 GMT-0600 (Central Standard Time)	0 seconds	Complete	308 Bytes	
Age Report Summary. 90-180 Days	Tue Dec 22 2020 02:24:47 GMT-0600 (Central Standard Time)	0 seconds	Complete	41 Bytes	
Age Report Detail. 720+ Days	Tue Dec 22 2020 02:24:46 GMT-0600 (Central Standard Time)	0 seconds	Complete	308 Bytes	~
Items per page: 20 ∨ 1–20 of 21 items				1 ∨ of 2 pages 4	٠

A1	- : × -	$f_{\!\scriptscriptstyle X}$ path																			
A	ВС	D	E	F	G	Н	1	J	K	L	М	N	0	Р	Q	R	S	Т	U	V	W
1 path	filename filetype	datasource own	er gro	oup	revision	site	platform	cluster	inode	permission	rfileset	uid	gid	recordver	state	migloc	mtime	atime	ctime	tier	size
2 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	()							08:03.0	08:03.0)		100863
3 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	()							08:02.9	08:02.9)		18515
4 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	()							08:03.3	08:03.3	3		109370
5 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:03.5	08:03.5	5		100985
6 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:04.1	08:04.1	L		276382
7 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:04.5	08:04.5	5		37205
8 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:04.8	08:04.8	3		76399
9 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:05.0	08:05.0)		6931
10 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:05.4	08:05.4	1		231879
11 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	()							08:05.7	08:05.7	7		174560
12 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:06.1	08:06.1	L		150729
13 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	()							07:57.6	07:57.6	5		5720
14 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	()							07:58.4	07:58.4	1		595383
15 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							07:58.1	07:58.1	L		66842
16 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							07:58.7	07:58.7	7		124600
17 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:01.1	08:01.1	L		1933439
18 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							07:59.9	07:59.9)		284812
19 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:00.8	08:00.8	3		46243
20 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:01.3	08:01.3	3		25959
21 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:01.7	08:01.7	7		215413
22 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:02.3	08:02.3	3		329510
23 udc-vaul	t/ T101389/c jpg	udc-vault				Tucson	IBM COS	df3b7ae7-	. ()							08:02.4	08:02.4	1		6220
24																					
25 26																					
26																					

5. Create New Tags Based on Filters

Tag management

Tags

Policies

Policy history

							Q Add Policy +
Policy	Туре	Schedule (UTC)	State	State	Collections	Last modified by	Last modified
ap_cs_body	CONTENTSEARCH	Done	Active	Complete		sdadmin	9/17/2021, 2:51:16 PM
ap_cs_meta	CONTENTSEARCH	Done	Active	Complete		sdadmin	9/16/2021, 1:18:06 PM
E1016404_autotag_add_field_pol	AUTOTAG	Done	Active	Complete		sdadmin	9/28/2021, 9:32:20 PM
T101389_autotag_u2source_pol	AUTOTAG	Done	Active	Complete		sdadmin	9/21/2021, 2:32:38 PM
T101389_ImportTags_pol	IMPORT_TAGS	Done	Active	Complete		sdadmin	9/21/2021, 3:17:07 PM
u3-iphone-gps	CONTENTSEARCH	Done	Active	Complete		sdadmin	9/21/2021, 7:44:08 PM
Items per page: 20 ∨ 1–6 of	6 items						1 ∨ of 1 pages

Tag management

Tags

Policies

Policy history

								Q	©
Policy name	Туре	Started	Ended	Status	Total records	Successful	Skipped	Failed	^
E1016404_autotag_add_field_pol	AUTOTAG	9/28/2021, 9:32:21 PM	9/28/2021, 9:32:23 PM	Complete	11	11	0		0
u3-iphone-gps	CONTENTSEARCH	9/21/2021, 7:44:08 PM	9/21/2021, 7:44:16 PM	Complete	1	1	0		0
u3-iphone-gps	CONTENTSEARCH	9/21/2021, 7:36:41 PM	9/21/2021, 7:37:02 PM	Complete	1	1	0		0
T101389_ImportTags_pol	IMPORT_TAGS	9/21/2021, 5:16:35 PM	9/21/2021, 5:16:53 PM	Complete	22	22	0		0
T101389_autotag_u2source_pol	AUTOTAG	9/21/2021, 5:15:42 PM	9/21/2021, 5:15:44 PM	Complete	22	22	0		0
T101389_autotag_u2source_pol	AUTOTAG	9/21/2021, 5:10:11 PM	9/21/2021, 5:10:12 PM	Complete	22	22	0		0
T101389_ImportTags_pol	IMPORT_TAGS	9/21/2021, 3:24:40 PM	9/21/2021, 3:24:56 PM	Complete	22	22	0		0
T101389_ImportTags_pol	IMPORT_TAGS	9/21/2021, 3:22:03 PM	9/21/2021, 3:22:22 PM	Complete	22	44	0		0
T101389_ImportTags_pol	IMPORT_TAGS	9/21/2021, 3:17:08 PM	9/21/2021, 3:17:27 PM	Complete	22	22	0		0
T101389_autotag_u2source_pol	AUTOTAG	9/21/2021, 2:32:39 PM	9/21/2021, 2:32:41 PM	Complete	22	22	0		0
Items per page: 20 ∨ 1–20 of 19	5 items						1 ∨ of 10 page	s 4	•

Query builder

Visual query SQL query Select * from Spectrum Discover where... filename LIKE ('%T101389%') AND path LIKE ('%udc-vault%') AND u2-source LIKE ('%annotated%') Limit results Sort by Group by Choose groupings 10000 e.g. owner asc \sim Setting a lower limit may increase speed of results. Set a sort order to choose which results are returned first.

Reset query

View results

Results view: ②

Inspecting records
11

Grouped By

Cluster
Datasource
Filegroup
Filename
Fileset
Show more...

							Q	(
	path	filename		udc dem2	u2 source	u2 el	u2 pp	
	udc- vault/	T101389/dataset/annotated /00c54e200e0d868faac8c55b	fca70dbe.jpg	T101389-s1	annotated	E1016404	FL	
	udc- vault/	T101389/dataset/annotated /00d4b7ea93ca1417393007c	61c519f88.jpg	T101389-s1	annotated	E1016404	FL	
	udc- vault/	T101389/dataset/annotated /00c09f406d31a0cd9402862	fbd26d930.jpg	T101389-s1	annotated	E1016404	FL	
	udc- vault/	T101389/dataset/annotated /000ec980b5b17156a55093b	4bd6004ab.jpg	T101389-s1	annotated	E1016404	FL	
	udc- vault/	T101389/dataset/annotated /00afb4420bdce881ad21686	7c713c800.jpg	T101389-s1	annotated	E1016404	FL	
	udc- vault/	T101389/dataset/annotated /00db76f676528477c66c93c	7d24d7524.jpg	T101389-s1	annotated	E1016404	FL	
	udc- vault/	T101389/dataset/annotated /00b627cb5578bb036edb01c	dcf7b56d9.jpg	T101389-s1	annotated	E1016404	FL	
	udc- vault/	T101389/dataset/annotated /00f4279089ddc0d40e14708	94f827f2e.jpg	T101389-s1	annotated	E1016404	FL	
	udc- vault/	T101389/dataset/annotated /0a2a11a8658d69b6cd1fd23	e5f4513bd.jpg	T101389-s1	annotated	E1016404	FL	
	udc- vault/	T101389/dataset/annotated /00ebecac178be3c3890a067	d6a28234e.jpg	T101389-s1	annotated	E1016404	FL	
Item	ns per page:	10 ∨ 1-10 of 11 items				1 ∨ of 2 pages	4	•

Query builder

6. View Images Directly with Calculated URL

Results view:

Condensed	Individual
Inspecting records	
Grouped By	
Cluster	
Datasource	
Filegroup	
Filename	
Fileset	
Show more	

					Q	®
	path	filename	cluster	udc dem2	u2 source	
	udc- vault/	T101389/dataset/annotated /000ec980b5b17156a55093b4bd6004ab.jpg	df3b7ae7-ab82-756b-01c4- c9f3ac8fae42	T101389-s1	annotated	
	udc- vault/	T101389/dataset/annotated /00afb4420bdce881ad216867c713c800.jpg	df3b7ae7-ab82-756b-01c4- c9f3ac8fae42	T101389-s1	annotated	
	udc- vault/	T101389/dataset/annotated /00b2add164cb42440a52064e390ea3d2.jpg	df3b7ae7-ab82-756b-01c4- c9f3ac8fae42	T101389-s1	annotated	
	udc- vault/	T101389/dataset/annotated /00b627cb5578bb036edb01cdcf7b56d9.jpg	df3b7ae7-ab82-756b-01c4- c9f3ac8fae42	T101389-s1	annotated	
	udc- vault/	T101389/dataset/annotated /00c09f406d31a0cd9402862fbd26d930.jpg	df3b7ae7-ab82-756b-01c4- c9f3ac8fae42	T101389-s1	annotated	
	udc- vault/	T101389/dataset/annotated /00c54e200e0d868faac8c55bfca70dbe.jpg	df3b7ae7-ab82-756b-01c4- c9f3ac8fae42	T101389-s1	annotated	
	udc- vault/	T101389/dataset/annotated /00d4b7ea93ca1417393007c61c519f88.jpg	df3b7ae7-ab82-756b-01c4- c9f3ac8fae42	T101389-s1	annotated	
	udc- vault/	T101389/dataset/annotated /00db76f676528477c66c93c7d24d7524.jpg	df3b7ae7-ab82-756b-01c4- c9f3ac8fae42	T101389-s1	annotated	
	udc- vault/	T101389/dataset/annotated /00ebecac178be3c3890a067d6a28234e.jpg	df3b7ae7-ab82-756b-01c4- c9f3ac8fae42	T101389-s1	annotated	
	udc- vault/	T101389/dataset/annotated /00f4279089ddc0d40e1470894f827f2e.jpg	df3b7ae7-ab82-756b-01c4- c9f3ac8fae42	T101389-s1	annotated	
Item	ns per page:	10 V 1-10 of 22 items		1 ∨ of 3 pag	es 4	•

Query builder

udc-dem2 like ('%T101389-s1%')







