

Demo: Unified Data Catalog for Military Aircraft

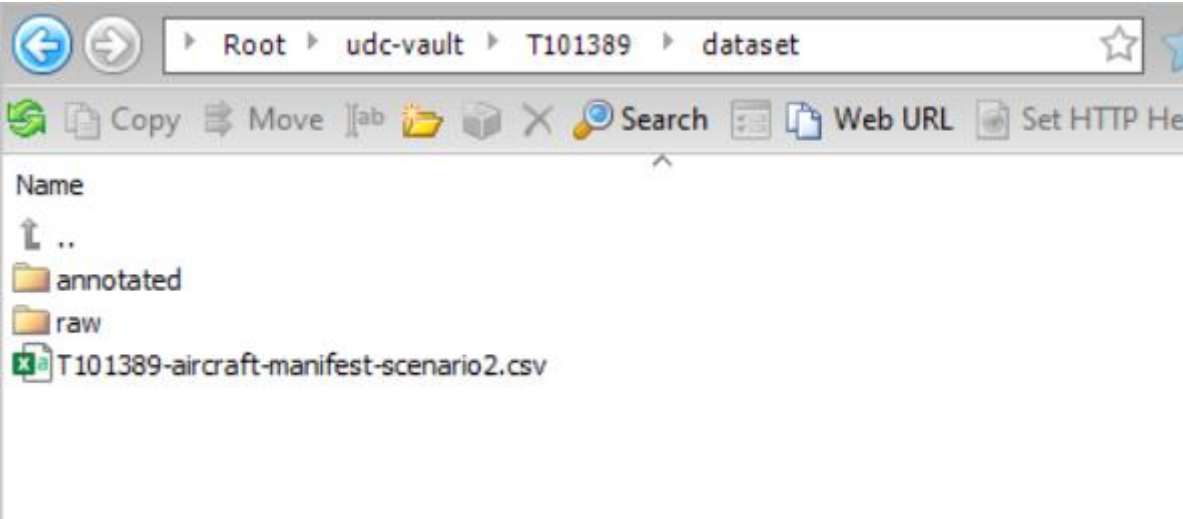
Scenario 2: Import Metadata to Enrich Catalog

Frank Lee

<https://github.com/drfranklee/demo-udc-aircraft>

Create Manifest for New Dataset “T101389-s2”

	A	B	C	D	E	F	G	H	I
1		u2-width	u2-height	u2-type	u2-xmin	u2-ymin	u2-xmax	u2-ymax	u2-id
2	udc-vault/T101389/dataset/annotated/000ec980b5b17156a55093b4bd6004ab.jpg	1000	551	Be200	308	126	651	201	T101389-10000
3	udc-vault/T101389/dataset/annotated/00afb4420bdce881ad216867c713c800.jpg	1200	800	Mig31	299	197	1102	569	T101389-10001
4	udc-vault/T101389/dataset/annotated/00b2add164cb42440a52064e390ea3d2.jpg	1280	850	B1	322	112	893	618	T101389-10002
5	udc-vault/T101389/dataset/annotated/00b627cb5578bb036edb01cdcf7b56d9.jpg	1300	867	F35	607	204	1202	529	T101389-10003
6	udc-vault/T101389/dataset/annotated/00c09f406d31a0cd9402862fbd26d930.jpg	1800	1200	A10	32	60	1800	766	T101389-10004
7	udc-vault/T101389/dataset/annotated/00c54e200e0d868faac8c55bfca70dbe.jpg	1200	1200	V22	647	521	973	768	T101389-10005
8	udc-vault/T101389/dataset/annotated/00d4b7ea93ca1417393007c61c519f88.jpg	1200	675	A10	560	308	647	329	T101389-10006
9	udc-vault/T101389/dataset/annotated/00db76f676528477c66c93c7d24d7524.jpg	1280	857	F117	176	263	1280	553	T101389-10007
10	udc-vault/T101389/dataset/annotated/00ebecac178be3c3890a067d6a28234e.jpg	960	640	C5	58	133	889	397	T101389-10008
11	udc-vault/T101389/dataset/annotated/00f4279089ddc0d40e1470894f827f2e.jpg	2048	1365	C5	875	398	1110	629	T101389-10009
12	udc-vault/T101389/dataset/annotated/0a2a11a8658d69b6cd1fd23c5f4513bd.jpg	960	640	Tu160	1	155	903	506	T101389-10010
13	udc-vault/T101389/dataset/raw/000ec980b5b17156a55093b4bd6004ab.jpg	1000	551	Be200					T101389-10000
14	udc-vault/T101389/dataset/raw/00afb4420bdce881ad216867c713c800.jpg	1200	800	Mig31					T101389-10001
15	udc-vault/T101389/dataset/raw/00b2add164cb42440a52064e390ea3d2.jpg	1280	850	B1					T101389-10002
16	udc-vault/T101389/dataset/raw/00b627cb5578bb036edb01cdcf7b56d9.jpg	1300	867	F35					T101389-10003
17	udc-vault/T101389/dataset/raw/00c09f406d31a0cd9402862fbd26d930.jpg	1800	1200	A10					T101389-10004
18	udc-vault/T101389/dataset/raw/00c54e200e0d868faac8c55bfca70dbe.jpg	1200	1200	V22					T101389-10005
19	udc-vault/T101389/dataset/raw/00d4b7ea93ca1417393007c61c519f88.jpg	1200	675	A10					T101389-10006
20	udc-vault/T101389/dataset/raw/00db76f676528477c66c93c7d24d7524.jpg	1280	857	F117					T101389-10007
21	udc-vault/T101389/dataset/raw/00ebecac178be3c3890a067d6a28234e.jpg	960	640	C5					T101389-10008
22	udc-vault/T101389/dataset/raw/00f4279089ddc0d40e1470894f827f2e.jpg	2048	1365	C5					T101389-10009
23	udc-vault/T101389/dataset/raw/0a2a11a8658d69b6cd1fd23c5f4513bd.jpg	960	640	Tu160					T101389-10010



Create IMPORT_TAGS Policy

```
{
  "pol_id": "T101389_ImportTags_pol",
  "action_id": "IMPORT_TAGS",
  "action_params": {
    "agent": "ImportTags",
    "source_connection": "UDSt-CoS",
    "tag_file_path": "udc-vault/T101389/dataset/T101389_s2_manifest.csv",
    "tag_file_type": "csv"
  },
  "schedule": "NOW",
  "pol_state": "active",
  "pol_filter": "datasource IN ('udc-vault')"
}
```

```
[moadmin@discover1 T101389]$ ls
T101389_autotag_u2source.json  T101389_ImportTags_pol.json
[moadmin@discover1 T101389]$
[moadmin@discover1 T101389]$
[moadmin@discover1 T101389]$ gettoken
  % Total      % Received % Xferd  Average Speed   Time    Time       Time  Current
                                 Dload  Upload   Total   Spent    Left     Speed
   0      0    0     0    0     0      0      0  --:--:--  --:--:--  --:--:--    0
[moadmin@discover1 T101389]$
[moadmin@discover1 T101389]$
[moadmin@discover1 T101389]$ curl -k -H "Authorization: Bearer ${TOKEN}" https://localhost/policyengine/v1
/policies/T101389_ImportTags_pol -X POST -d @./T101389_ImportTags_pol.json -H "Content-Type: application/js
on"
Policy 'T101389_ImportTags_pol' added[moadmin@discover1 T101389]$
```


Tag management

- Tags
- Policies
- Policy history

Policy name	Type	Started	Ended	Status	Total records	Successful	Skipped	↕	Failed
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
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
ap_cs_body	CONTENTSEARCH	9/17/2021, 2:51:17 PM	9/17/2021, 2:51:25 PM	✔ Complete	16	16	0		0

Query builder

Visual query SQL query

Select * from Spectrum Discover where...



filename LIKE ('%T101389%') AND filetype LIKE ('jpg')

Group by

Choose groupings



Limit results

10000



Setting a lower limit may increase speed of results.

Sort by

e.g. owner asc

Set a sort order to choose which results are returned first.

Reset query

View results

Results view: ?

Matched 22 records from metaocean_view table in 0.061466 seconds

Condensed Individual

Inspecting records
22

Grouped By

- ☐ Cluster
- ☐ Datasource
- ☐ Filegroup
- ☐ Filename
- ☐ Fileset

Show more...

udc dem2	u2 id	u2 source	u2 width	u2 height	u2 xmin	u2 ymin	u2 xmax	u2 ymax
T101389-s1	T101389-10000	annotated	1000	551	308	126	651	201
T101389-s1	T101389-10001	annotated	1200	800	299	197	1102	569
T101389-s1	T101389-10002	annotated	1280	850	322	112	893	618
T101389-s1	T101389-10003	annotated	1300	867	607	204	1202	529
T101389-s1	T101389-10004	annotated	1800	1200	32	60	1800	766
T101389-s1	T101389-10005	annotated	1200	1200	647	521	973	768
T101389-s1	T101389-10006	annotated	1200	675	560	308	647	329
T101389-s1	T101389-10007	annotated	1280	857	176	263	1280	553
T101389-s1	T101389-10008	annotated	960	640	58	133	889	397

Query builder

filename LIKE ('%T101389%') AND filetype LIKE ('%jpg%')

Query Results

Results view: ?

Matched 22 records from metaocean_view table in 0.058539 seconds

Condensed Individual

Inspecting records
22

Grouped By

- ☐ Cluster
- ☐ Datasource
- ☐ Filegroup
- ☐ Filename
- ☐ Fileset

Show more...

22 items selected

Generate reportAdd tagsExport dataCancel

<input checked="" type="checkbox"/>	path	filename	udc dem2	u2 id	udc dem3	u2 source	u2 width	u2 l
<input checked="" type="checkbox"/>	udc-vault/	T101389/dataset/annotated/00d4b7ea93ca1417393007c61c519f88.jpg	T101389-s1	T101389-10006	T101389-s2	annotated	1200	
<input checked="" type="checkbox"/>	udc-vault/	T101389/dataset/annotated/00db76f676528477c66c93c7d24d7524.jpg	T101389-s1	T101389-10007	T101389-s2	annotated	1280	
<input checked="" type="checkbox"/>	udc-vault/	T101389/dataset/annotated/00ebecac178be3c3890a067d6a28234e.jpg	T101389-s1	T101389-10008	T101389-s2	annotated	960	
<input checked="" type="checkbox"/>	udc-	T101389/dataset/annotated	T101389-s1	T101389-10009	T101389-s2	annotated	1200	

Items per page: 10 1–10 of 22 items1 of 3 pages

Add tags

22 files selected

udc-dem3

T101389-s2

Add another tag

CancelSubmit

Custom-tag the Selected Data to create a New Dataset “T101389-s2”: Result

Results view: ?

Matched 22 records from metaocean_view table in 0.061466 seconds

Condensed

Individual

Inspecting records

22

Grouped By

- ☐ Cluster
- ☐ Datasource
- ☐ Filegroup
- ☐ Filename
- ☐ Fileset

Show more...

<input type="checkbox"/>	path	filename	u2 id	udc dem3	u2 source	u2 width	u2 height
<input type="checkbox"/>	udc-vault/	T101389/dataset/annotated/000ec980b5b17156a55093b4bd6004ab.jpg	T101389-10000	T101389-s2	annotated	1000	551
<input type="checkbox"/>	udc-vault/	T101389/dataset/annotated/00afb4420bdce881ad216867c713c800.jpg	T101389-10001	T101389-s2	annotated	1200	800
<input type="checkbox"/>	udc-vault/	T101389/dataset/annotated/00b2add164cb42440a52064e390ea3d2.jpg	T101389-10002	T101389-s2	annotated	1280	850
<input type="checkbox"/>	udc-vault/	T101389/dataset/annotated/00b627cb5578bb036edb01cdcf7b56d9.jpg	T101389-10003	T101389-s2	annotated	1300	867
<input type="checkbox"/>	udc-vault/	T101389/dataset/annotated/00c09f406d31a0cd9402862fbd26d930.jpg	T101389-10004	T101389-s2	annotated	1800	1200
<input type="checkbox"/>	udc-vault/	T101389/dataset/annotated/00c54e200e0d868faac8c55bfca70dbe.jpg	T101389-10005	T101389-s2	annotated	1200	1200
<input type="checkbox"/>	udc-vault/	T101389/dataset/annotated/00d4b7ea93ca1417393007c61c519f88.jpg	T101389-10006	T101389-s2	annotated	1200	675
<input type="checkbox"/>	udc-vault/	T101389/dataset/annotated/00db76f676528477c66c93c7d24d7524.jpg	T101389-10007	T101389-s2	annotated	1280	857
<input type="checkbox"/>	udc-vault/	T101389/dataset/annotated/00ebecac178be3c3890a067d6a28234e.jpg	T101389-10008	T101389-s2	annotated	960	640
<input type="checkbox"/>	udc-	T101389/dataset/annotated	T101389-10000	T101389-s2	annotated	1000	551

Items per page:

50



1-22 of 22 items

1



of 1 pages



Query builder

filename LIKE ('%T101389%') AND filetype LIKE ('%jpg%')



Find the Newly Tagged Dataset with Visual Query

Query builder

Visual query

SQL query

Query preview...

Site

0/5

Sizerange

0/4

State

0/2

Td_bh01_tumor

0/10

Temperature

0/1

Tier

0/2

Timesinceaccess

0/5

U2-id

0/12

U2-type

0/10

Udc-dem2

0/5



☐ udc-dem2

count

☐ NULL

21,973

☐ T101389-s1

11

☐ T101389-s2

11

☐ E1016270-SGDSO

8

☐ E1016270-demo

1

Items per page: 10

1-5 of 5 items

1 of 1 pages

Find the Newly Tagged Dataset with Visual Query using new Tag “u2-type”

Query builder

Visual query

SQL query

Query preview...

Size range

U/4

State

0/2

Td_bh01_tumor

0/10

Temperature

0/1

Tier

0/2

Timesinceaccess

0/5

U2-id

0/12

U2-source

0/3

U2-type

0/10

Udc-dem2

0/4

Udc-dem3

0/2

0

NULL

A10

C5

B1

Be200

F117

F35

Mig31

Tu160

V22

null

A10

C5

B1

Be200

F117

F35

Mig31

Tu160

V22

u2-type	count
<input type="checkbox"/> NULL	21,972
<input type="checkbox"/> A10	4
<input type="checkbox"/> C5	4
<input type="checkbox"/> B1	2
<input type="checkbox"/> Be200	2
<input type="checkbox"/> F117	2
<input type="checkbox"/> F35	2
<input type="checkbox"/> Mig31	2
<input type="checkbox"/> Tu160	2
<input type="checkbox"/> V22	2

Items per page: 10


1–10 of 10 items

1 of 1 pages

Query builder

Visual querySQL query

Select * from Spectrum Discover where...

 u2-id LIKE ('%T101389-10005%')

Group by

1 X

Groups selected

▼

Limit results

10000

Setting a lower limit may increase speed of re

Find the Paring Data (raw-annotated) based on newly-crated file ID “u2-id”: result

Results view:

Matched 2 records from metaocean_view table in 0.038424 seconds

Condensed

Individual

Inspecting records

2

Grouped By

- ☐ Cluster
- ☐ Datasource
- ☐ Filegroup
- ☐ Filename
- ☐ Fileset

Show more...

filename	u2 id	udc dem2	u2 type	u2 width	u2 height	u2 xmin	u2 xmax
T101389/scenario2 /raw/00c54e200e0d868faac8c55bfca70dbe.jpg	T101389-10005	NULL	V22	1200	1200	NULL	NU
T101389/scenario2/annotated /00c54e200e0d868faac8c55bfca70dbe.jpg	T101389-10005	T101389-s2	V22	1200	1200	647	9'

Items per page: 10

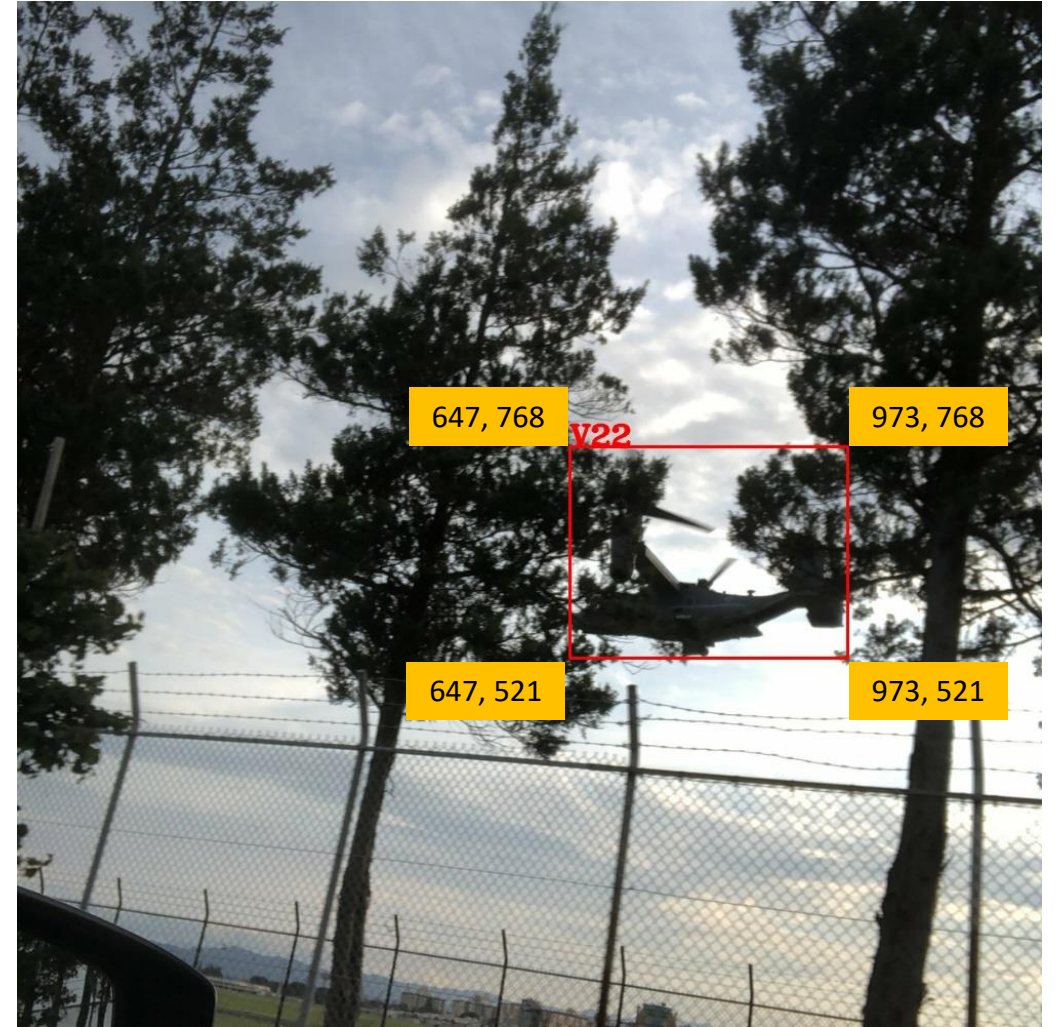
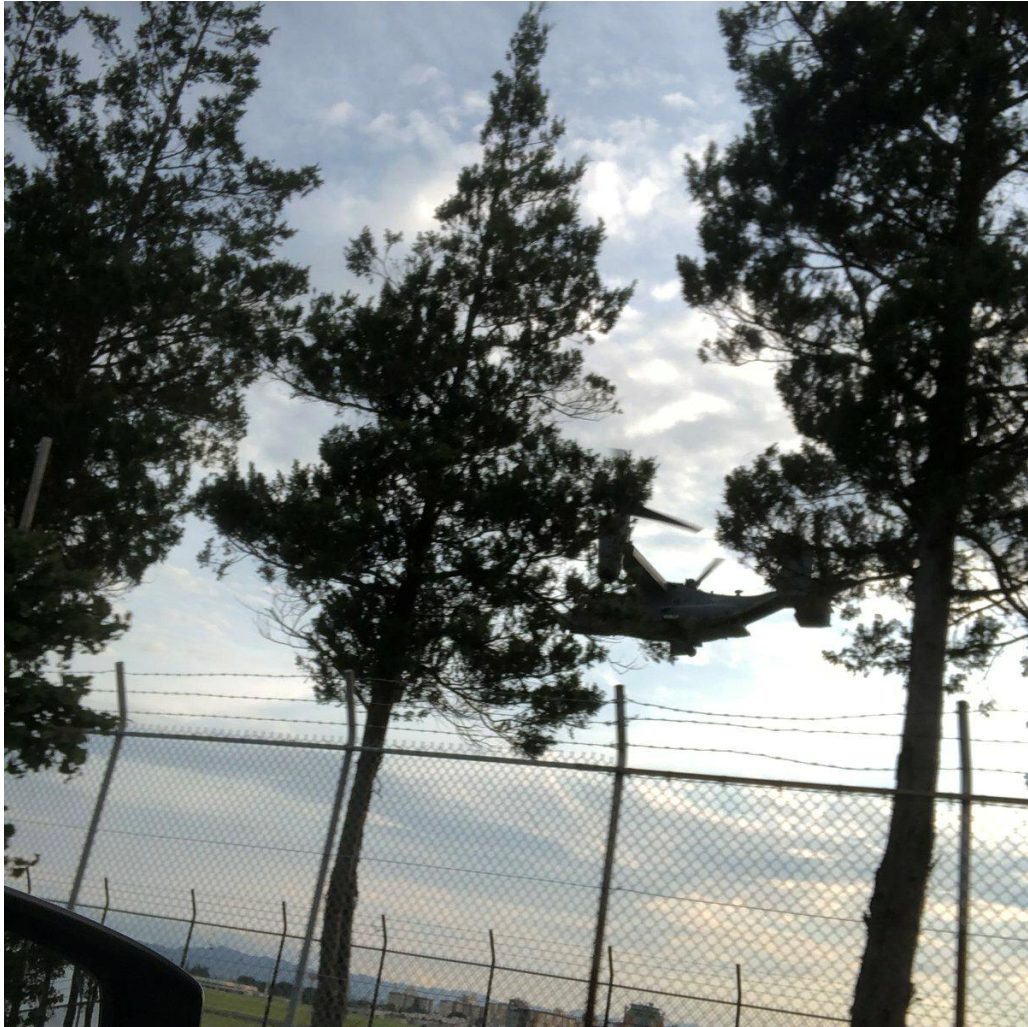
1-2 of 2 items

1 of 1 pages

Query builder

u2-id LIKE ('%T101389-10005%')





u2-id = T101389-10005: filename = 00c54e200e0d868faac8c55bfca70dbe.jpg