

Example Feature Investigate - Advanced Mapping

Wednesday, 28 September 2022 8:27 AM

Purpose

I encountered a case where the ingestion process was generating clues with entity codes that I didn't want. Advanced Mapping has been suggested as a low code/no code solution.

Journal

go to settings and enable

 [Administration / Settings](#)

Advanced Mapping

 Enabled

add some data as a source for test mapping

<https://www.briandunning.com/sample-data/>
use au-500.csv



automap

Choose Entity Type

Select from the list or enter name to create new entity type.

Person

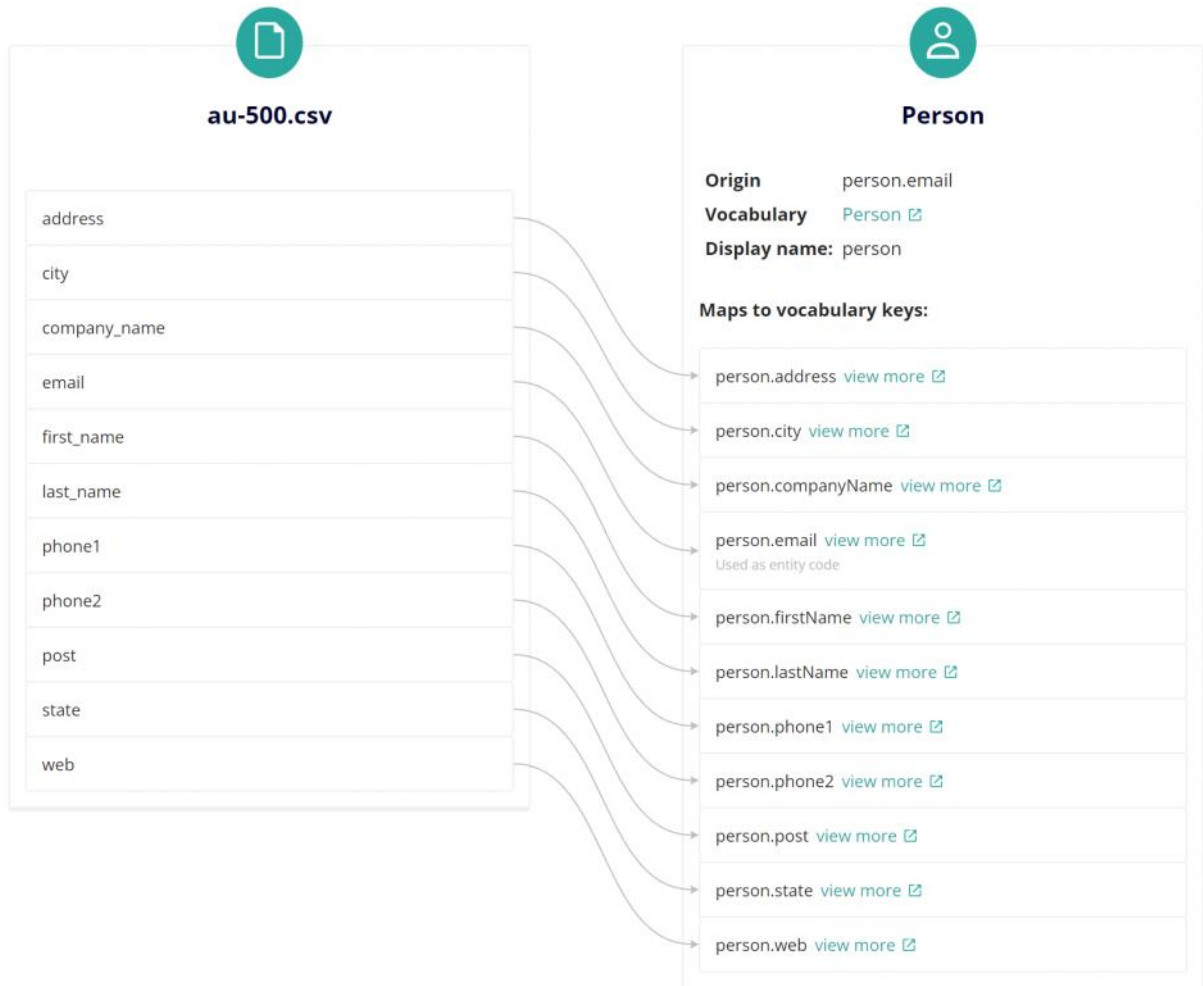
[View more details on Person](#) 

Choose vocabulary

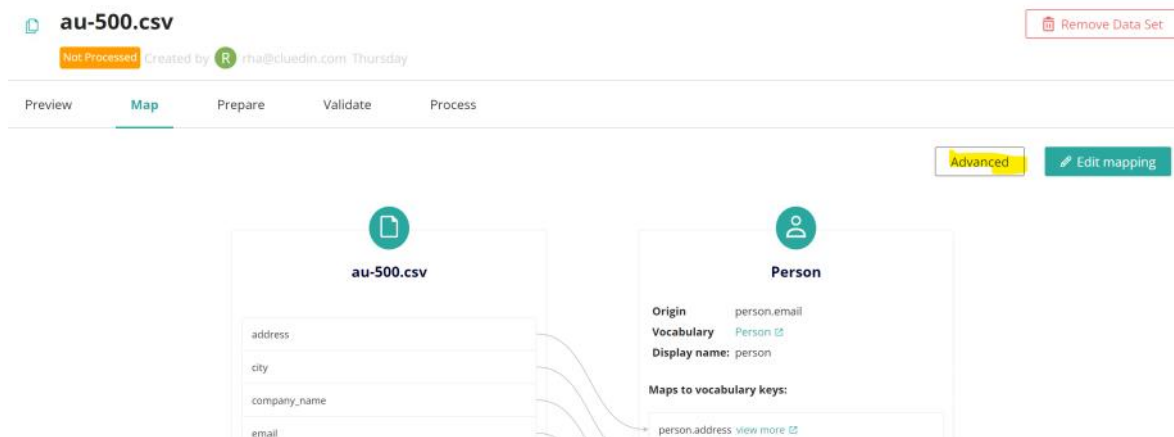
Select from the list or enter name to create new vocabulary.

Person

[View more details on Person](#) 



observe the advanced button on the mapping page



click it and it shows this interface

au-500.csv Remove Data Set

Not Processed Created by Rha@cluedin.com Thursday

Preview **Map** Prepare Validate Process

← Back Run Save Code mapping before sending clues

Press the play button, if valid, the Clues created will appear here.

here's the before sending or before creating choice - let's go with before sending

Code mapping before sending clues

Code mapping before sending clues

Code mapping before creating clues

click run

Code mapping before sending clues

```

{
  "root": [ 10 items
    { 1 item
      "clue": { 4 items
        "attribute-organization": "b81134e3-27d8-4bb5-8437-b5899588184f"
        "attribute-origin":
          "/Person#CluedInImporter(dataset-530F7F70-4D2B-4220-B300-B88AF69145A6):[REDACTED]@[REDACTED].com.au"
        "attribute-appVersion": "1.0.0"
        "clueDetails": { 1 item
          "data": { 7 items
            "attribute-originProviderDefinitionId":
              "7A1ABE56-6645-4813-A34E-182C6BF9DED9"
            "attribute-origin":
              "/Person#CluedInImporter(dataset-530F7F70-4D2B-4220-B300-B88AF69145A6):[REDACTED]@[REDACTED].com.au"
            "attribute-appVersion": "1.0.0"
            "attribute-inputSource": "cluedin-annotation"
            "createdDate": "2022-10-04T22:29:13.504Z"
            "modifiedDate": "2022-10-04T22:29:13.504Z"
            "entityData": { 11 items
              "": ""
              "": ""
              "": ""
              "": ""
              "": ""
              "": ""
              "": ""
              "": ""
              "": ""
              "": ""
              "": ""
            }
          }
        }
      }
    }
  ]
}

```

some doco

<https://github.com/CluedIn-io/C.E.I>

simple example

```

value.clue.clueDetails.data.entityData.description =
  value.clue.clueDetails.data.entityData.name + " is the best!";
[value]

```

← Back

▶ Run

Save

Code mapping before sending clues

```

1 value.clue.clueDetails.data.entityData.description =
2 value.clue.clueDetails.data.entityData.name + " is the best!";
3 [value]

```

```

"createdDate": "2022-10-05T03:30:31.104Z"
"modifiedDate": "2022-10-05T03:30:31.104Z"
"entityData": {
  "name": "Rebecca"
  "description": "Rebecca is the best!"
  "entityType": "/Person"
  "attribute-origin":
    "/Person#CluedInImporter(dataset-530F7F70-4D2B-4220-

```

let's remove all the entity codes and set the origin to something that is source centric - for us to modify later in the code

2

Entity Origin

Origin: **au500**

Edit origin

Select key

Choose the property that will be used to produce entity origin code

✓

▼

Add Property

Status

Empty Values

None ✓

Duplicate Values

None ✓

3

Entity Codes

✗

🔍

what does that yield?

```

{
  "attribute-organization": "b81134e3-27d8-4bb5-8437-b5899588184f",
  "attribute-origin": "/Person#au500:rebecca.didio@didio.com.au",
  "attribute-appVersion": "1.0.0",
  "clueDetails": {
    "data": {
      "attribute-originProviderDefinitionId": "7A1ABE56-6645-4813-A34E-182C6BF9DED9",
      "attribute-origin": "/Person#au500:rebecca.didio@didio.com.au",
      "attribute-appVersion": "1.0.0",
      "attribute-inputSource": "cluedin-annotation",
      "createdDate": "2022-10-05T03:37:52.790Z",
      "modifiedDate": "2022-10-05T03:37:52.790Z",
      "entityData": {
        "name": "Rebecca",
        "description": "Rebecca is the best!",
        "entityType": "/Person",
        "attribute-origin": "/Person#au500:rebecca.didio@didio.com.au",
        "attribute-appVersion": "1.0.0",
        "attribute-source": "au500",
        "codes": [
          "/Person#CluedIn(hash-sha1):99598bb2e945a77a6a98cba5525d0207a05c3af5"
        ],
        "modifiedDate": "2022-10-05T03:37:52Z",
        "discoveryDate": "2022-10-05T03:37:52Z",
        "edges": {},
        "properties": {
          "attribute-type": "/Metadata/KeyValue",
          "property-person.address": "171 E 24th St",
          "property-person.city": "Leith",
          "property-person.companyName": "Brandt, Jonathan F Esq",
          "property-person.email": "rebecca.didio@didio.com.au",
          "property-person.firstName": "Rebecca",
          "property-person.lastName": "Didio",
          "property-person.phone1": "03-8174-9123",
          "property-person.phone2": "0458-665-290",
          "property-person.post": "7315",
          "property-person.state": "TAS",
          "property-person.web": "http://www.brandtjonathanfesq.com.au"
        }
      }
    }
  }
}

```

what will the resultant entity look like?
what entity codes will it have?

ingest data

500 clues will be created.

Stats

Success 500 ✓

query in search - our custom description is present

All

Filter by

Clear 1 filters

EntityType

Person

Type to filter entityType

User

Organization

Advanced Vocabulary Term Filter

500 results found

Sort By Relevance

Tile

Tabular

Customize Tile view

Clean

Prepare

Select for merge

Alida Helger

Person

Last modified 3 minutes ago, created 3 minutes ago

description

Alida is the best!

Kent Ivans

Person

Last modified 3 minutes ago, created 3 minutes ago

description

Kent is the best!

query in Engine Room -> Processing Pipeline

CLUE PROCESSING

HIDE 1

Default

Message Rates

Total

0

Incoming

0.00/s

Outgoing

0.00/s

Big Clue

EDGE PROCESSING

HIDE 1

Edge Processing

Message Rates

Total

0

Incoming

0.00/s

Outgoing

0.00/s

inspect rebecca entry

< Back

Person

Rebecca Didio

View codes

Overview

All Properties

Relations

Pending Changes

History

Explain Log

Rebecca Didio

Last modified 4 minutes ago, created 4 minutes ago

city

Leith

post

7315

address

171 E 24th St

phone2

0458-665-290

state

TAS

phone1

03-8174-9123

web

http://www.brandtjonathanfesq.com.au

Entity data KPIs

Accuracy over time

100 %

Improve this metric

Not enough data to show trend.

Connectivity over time

100 %

Completeness over time

- %


No data to show yet

Improve this metric

Not enough data to show trend




Relevance over time


13 %


< Back |  Person > **Rebecca Didio**

Overview	All Properties	Relations	Pending Changes	History	Explain Log
----------	----------------	-----------	-----------------	---------	--------------------


Golden Record

>  Entity type:  Person
 Origin code: `/Person#au500:rebecca.didio@didio.com.au`  Copy code


 **Data Parts**

 **Data Part**
 Data Part ID: 20


Process names

 Status: **Successful**
 Duration: 00:00:00.038s

Detect mime type

 Status: **Successful**
 Duration: 00:00:00.005s

Extract entity codes

 Status: **Successful**
 Duration: 00:00:00.005s

Summaries


Table: Changes


Change type	Key	Before	After
Added	Entity code		<code>/Person#CluedIn(email):rebecca.didio@didio.com.au</code>

Events and Operations

"ChangeType: Added, Key: Entity code, Before: , After: /Person#CluedIn(email):rebecca.didio@didio.com.au"

this is the codes tab

Entity codes 

Origin code  Copy code

`/Person#au500:rebecca.didio@didio.com.au`

Codes:

`/Person#au500:rebecca.didio@didio.com.au`

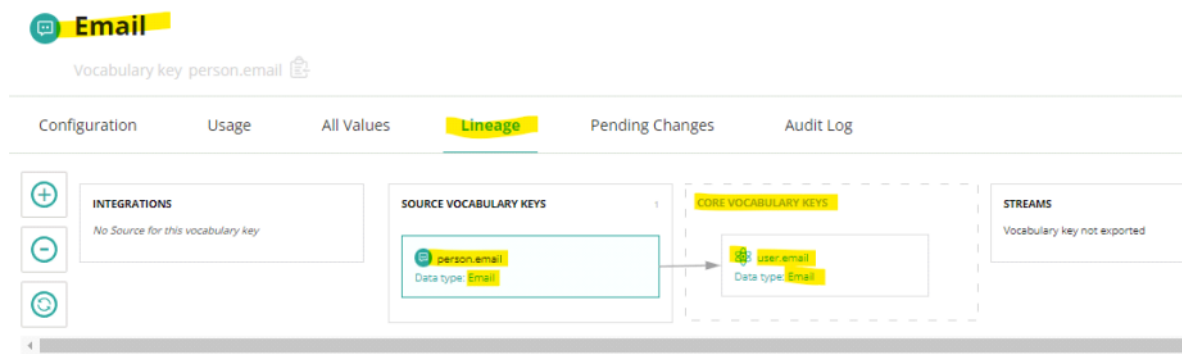
`/Person#CluedIn(email):rebecca.didio@didio.com.au`

`/Person#CluedIn(hash-sha1):99598bb2e945a77a6a98cba5525d0207a05c3af5`

let's explore the CluedIn(email) code

it's generated in the processing phase by the "Extract entity codes" part

it is because **person.email maps to user.email**



let's update person.email to not map to user.email and reprocess our example entry to see what happens

Mmmm we cannot edit it


Maps to



Choose a vocabulary key where the current vocabulary key should be mapped. Once mapped, the current vocabulary key's value will flow towards the selected vocabulary key. Once a value is defined, please visit the lineage tab to view the data flow.

 [Email](#)    Edit


let's delete all and try again - nope can't edit

create new vocab key via the mapping UI and ensure it doesn't map to user.email

Add mapping to "email" 

Vocabulary Key
Key for the field, should be unique
 
 Full vocabulary key will be **person.emailAddress**

Display name*
User-friendly text used in the user interface to help with readability.


Data type*
 


Use as entity code
Specify is this field can be used a single identifier for a clued (eg: a SQL foreign key)
☐ 0

Use as alias
Specify is this field can be used an alias for a clue
☐ 0

2 Entity Origin

Origin: **au500**

 Edit origin

Select key 

Choose the property that will be used to produce ent

person.emailAddress



 Add Property

emailAddress

Vocabulary key person.emailAddress  - by  rha@cluedin.com - a minute ago

Configuration

Usage

All Values

Lineage

Pending Changes

Audit Log

Display name

User-friendly text used in the user interface to help with readability.

emailAddress

Name

The vocabulary 'name' will be used to create the full vocabulary 'key'.

emailAddress

Full vocabulary key will be **person.emailAddress**

Group name

Metadata

Is Visible



Data type

Email

Classification

/

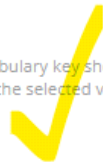
Description

No description has been set

Maps to

Choose a vocabulary key where the current vocabulary key should be mapped. Once mapped, the current vocabulary key's value will flow towards the selected vocabulary key. Once a value is defined, please visit the lineage tab to view the data flow.

Key is not mapped to another vocabulary key




delete all

add admin role to enable deletes

Administrator Settings for **rha@cluedin.com**

[← Back](#) [Settings](#) [Roles](#) [Integration Pt](#)

 Add role to user

Role Name

Admin

OrganizationAdmin

signout/signin

consume -> GraphQL

query all /Person

```
{
  search(query:"entityType:/Person")
  {
```



```

    totalResults
  }
}

response
{
  "data": {
    "search": {
      "totalResults": 500
    }
  }
}

```

delete all /Person

```

{
  search(query: "entityType:/Person", pageSize: 500)
  {
    totalResults
    entries {
      actions {
        deleteEntity
      }
    }
  }
}

```

re-run query until

```

{
  "data": {
    "search": {
      "totalResults": 0,
      "entries": []
    }
  }
}

```

update mapping via advance mapping to change the origin and entity codes accordingly

first clue

```

{
  "attribute-organization": "b81134e3-27d8-4bb5-8437-b5899588184f",
  "attribute-origin": "/Person#au500:rebecca.didio@didio.com.au",
  "attribute-appVersion": "1.0.0",
  "clueDetails": {
    "data": {
      "attribute-origInProviderDefinitionId": "7A1ABE56-6645-4813-A34E-182C6BF9DED9",
      "attribute-origin": "/Person#au500:rebecca.didio@didio.com.au",
      "attribute-appVersion": "1.0.0",
      "attribute-inputSource": "cluedin-annotation",
      "createdDate": "2022-10-05T05:50:07.631Z",
      "modifiedDate": "2022-10-05T05:50:07.631Z",
      "entityData": {
        "name": "Rebecca",
        "description": "Rebecca is the best!",
        "entityType": "/Person",
        "attribute-origin": "/Person#au500:rebecca.didio@didio.com.au",
        "attribute-appVersion": "1.0.0",
        "attribute-source": "au500",
        "codes": [
          "/Person#CluedIn(hash-sha1):1e85c79fd45ffa0aa3935290dd71ababc96d86f6"
        ],
        "modifiedDate": "2022-10-05T05:50:07Z",
        "discoveryDate": "2022-10-05T05:50:07Z",
        "edges": {},
        "properties": {
          "attribute-type": "/Metadata/KeyValue",
          "property-person.address": "171 E 24th St",
          "property-person.city": "Leith",
          "property-person.companyName": "Brandt, Jonathan F Esq",
          "property-person.emailAddress": "rebecca.didio@didio.com.au",
          "property-person.firstName": "Rebecca",
          "property-person.lastName": "Didio",
          "property-person.phone1": "03-8174-9123",
          "property-person.phone2": "0458-665-290",
          "property-person.post": "7315",
          "property-person.state": "TAS",
          "property-person.web": "http://www.brandtjonathanfesq.com.au"
        }
      }
    }
  }
}

```

try updating the origin entity code

```

// debug - test description is updated and therefore no errors in the below code
value.clue.clueDetails.data.entityData.description =

```

```

        value.clue.clueDetails.data.entityData.name + " is the best!";

// get the current origin code and update it
originCode = value.clue["attribute-origin"];
originCode = originCode.replace("#au500:", "#email:");

// replace origin code in all 3 places
value.clue["attribute-origin"] = originCode;
value.clue.clueDetails.data["attribute-origin"] = originCode;
value.clue.clueDetails.data.entityData["attribute-origin"] = originCode;

[value]

```

results

```

{
  "attribute-organization": "b81134e3-27d8-4bb5-8437-b5899588184f",
  "attribute-origin": "/Person#email:rebbecca.didio@didio.com.au",
  "attribute-appVersion": "1.0.0",
  "clueDetails": {
    "data": {
      "attribute-originProviderDefinitionId": "7A1ABE56-6645-4813-A34E-182C6BF9DED9",
      "attribute-origin": "/Person#email:rebbecca.didio@didio.com.au",
      "attribute-appVersion": "1.0.0",
      "attribute-inputSource": "cluedin-annotation",
      "createdDate": "2022-10-05T06:16:28.667Z",
      "modifiedDate": "2022-10-05T06:16:28.667Z",
      "entityData": {
        "name": "Rebecca",
        "description": "Rebecca is the best!",
        "entityType": "/Person",
        "attribute-origin": "/Person#email:rebbecca.didio@didio.com.au",
        "attribute-appVersion": "1.0.0",
        "attribute-source": "au500",
        "codes": [
          "/Person#CluedIn(hash-sha1):1e85c79fd45ffa0aa3935290dd71ababc96d86f6"
        ],
        "modifiedDate": "2022-10-05T06:16:28Z",
        "discoveryDate": "2022-10-05T06:16:28Z",
        "edges": {},
        "properties": {
          "attribute-type": "/Metadata/KeyValue",
          "property-person.address": "171 E 24th St",
          "property-person.city": "Leith",
          "property-person.companyName": "Brandt, Jonathan F Esq",
          "property-person.emailAddress": "rebbecca.didio@didio.com.au",
          "property-person.firstName": "Rebecca",
          "property-person.lastName": "Didio",
          "property-person.phone1": "03-8174-9123",
          "property-person.phone2": "0458-665-290",
          "property-person.post": "7315",
          "property-person.state": "TAS",
          "property-person.web": "http://www.brandtjonathanfesq.com.au"
        }
      }
    }
  }
}

```

let's also get rid of the codes

add this near the bottom

```

// remove all codes
value.clue.clueDetails.data.entityData.codes.length = 0;

```

results

```

{
  "name": "Rebecca",
  "description": "Rebecca is the best!",
  "entityType": "/Person",
  "attribute-origin": "/Person#email:rebbecca.didio@didio.com.au",
  "attribute-appVersion": "1.0.0",
  "attribute-source": "au500",
  "codes": [],
  "modifiedDate": "2022-10-05T06:18:57Z",
  "discoveryDate": "2022-10-05T06:18:57Z",
  "edges": {},
  "properties": {
    "attribute-type": "/Metadata/KeyValue",
    "property-person.address": "171 E 24th St",
    "property-person.city": "Leith",
    "property-person.companyName": "Brandt, Jonathan F Esq",
    "property-person.emailAddress": "rebbecca.didio@didio.com.au",
    "property-person.firstName": "Rebecca",
    "property-person.lastName": "Didio",
    "property-person.phone1": "03-8174-9123",
    "property-person.phone2": "0458-665-290",
    "property-person.post": "7315",
    "property-person.state": "TAS",
    "property-person.web": "http://www.brandtjonathanfesq.com.au"
  }
}

```

final code

```

// debug - test description is updated and therefore no errors in the below code
value.clue.clueDetails.data.entityData.description =
  value.clue.clueDetails.data.entityData.name + " is the best!";

// get the current origin code and update it
originCode = value.clue["attribute-origin"];

```

```

originCode = originCode.replace("#au500:", "#email:");

// replace origin code in all 3 places
value.clue["attribute-origin"] = originCode;
value.clue.clueDetails.data["attribute-origin"] = originCode;
value.clue.clueDetails.data.entityData["attribute-origin"] = originCode;

// remove all codes
value.clue.clueDetails.data.entityData.codes.length = 0;

[value]

```

let's ingest and check out



just the one code!!! success!

The screenshot shows a user profile for "Dexter Prosienski" with tabs for Overview, All Properties, Relations, Pending Changes, and History. The Overview tab is active, showing a profile picture and a card for "Entity data KPIs" with "Accuracy over time" at "100 %". A modal window titled "Entity codes" is open, showing the "Origin code" as "/Person#email:dexter@prosienski.net.au" and a list of "Codes" with the same value. A large yellow checkmark is drawn over the code in the modal.

winner winner 🍷 dinner!