entrinsicon

Auditing and Tracking Custom Computed Columns

Sam Parsons
Monroe University
Director of Data Strategy, Analytics, and AI

About Monroe University

Tri-semester school:

Fall Term: Sept-Dec
Winter: Jan-Apr
Spring: May-Aug

Four campuses: Bronx, New Rochelle, Online, Saint Lucia

Enrollment: ~8,000 - 9,000 students

Environment: On-premise Unidata Colleague

Where we're at in Informer migration:

Informer 5 is online, all new report requests go to I5

50% reporting comes from live excels –working with each dept to transition to UI for the first time

What goes into the

entrinsicon

Informer 5 datasource

Table: Field

- ❑ *Field ID*
- ❑ *Mapping*
- ❑ *Field display name*
- ❑ *Data type*
- ❑ *DATA JSON object (very valuable!!)*
- ❑ *Restricted*
- ❑ *Hidden*
- ❑ *Field set*

Colleague datasource

Table: RT.FIELDS

- ❑ *Validation code/table*
- ❑ *Association, list, or SV*
- ❑ *Usage in awarding rules*
- ❑ *Usage in billing rules*
- ❑ *Usage in term standing rules*

Excel file datasource

- ❑ *Field description*
- ❑ *Reason for not migrating*
- ❑ *Other IT notes about the field*

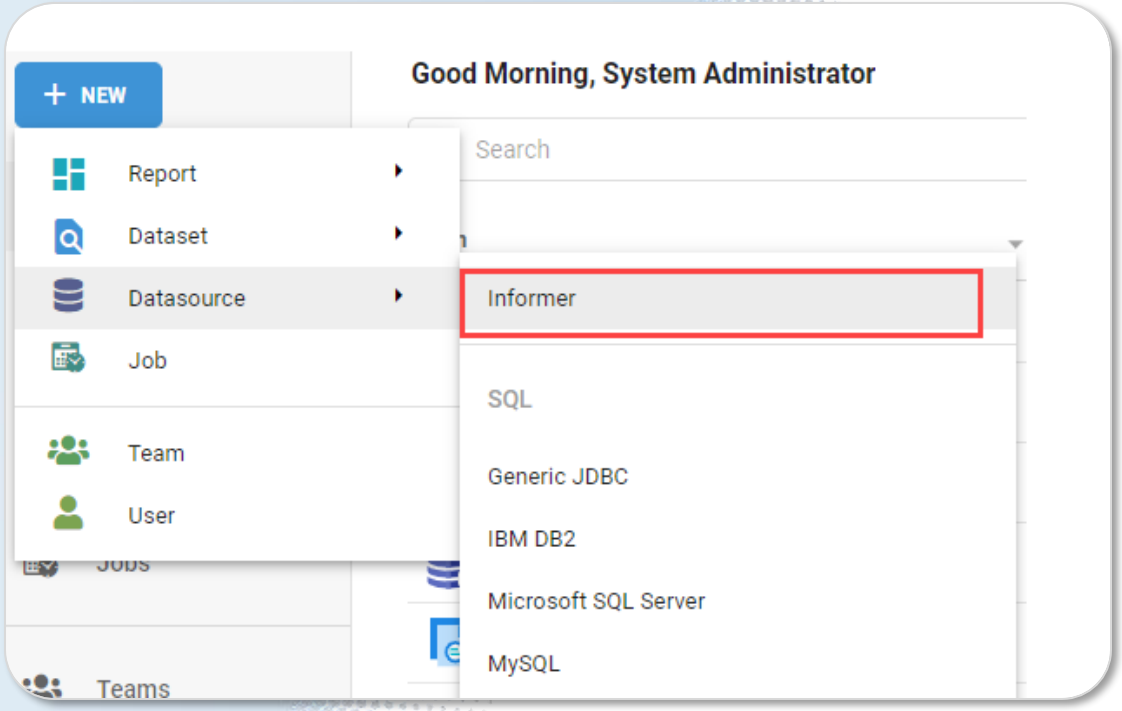
Calculations

- ❑ *Extracting subroutine name or computed column logic*
- ❑ *Extracting fields used in computed column logic*
- ❑ *Written in Envision or Unibasic?*

Initial Setup

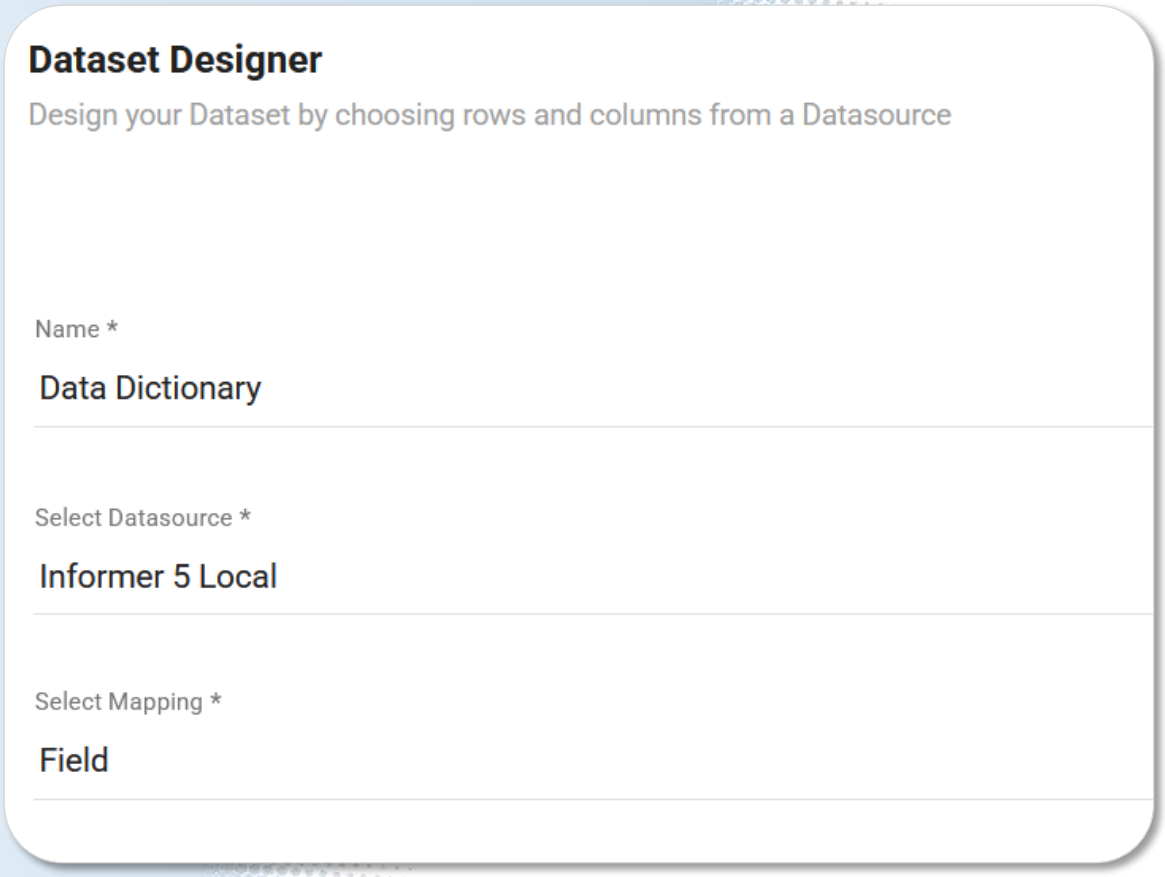
1. If you do not have the Informer datasource setup, follow instructions in article below.

https://informer5.zendesk.com/hc/en-us/articles/11265457384596-Creating-an-Informer-Datasource



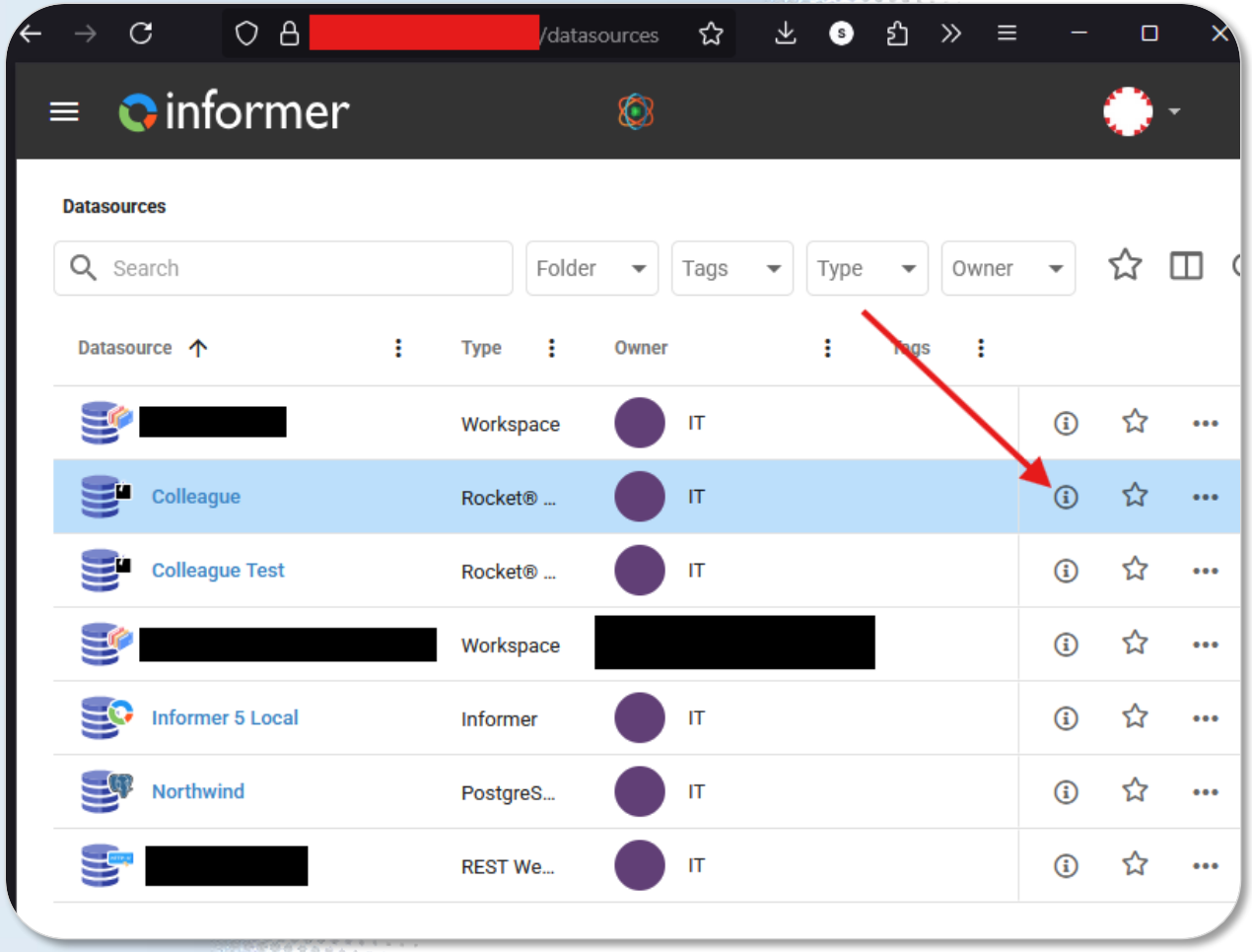
Initial Setup

1. **Create a dataset based in the Informer datasource using the Mapping “Field”.**
2. Find the datasource ID of your Unidata Colleague datasource in Informer 5.
3. Add a filter to your dataset to filter to Colleague. If you have ACYR tables, filter out your individual instance tables. I used the filter “does not contain .19/.20”.



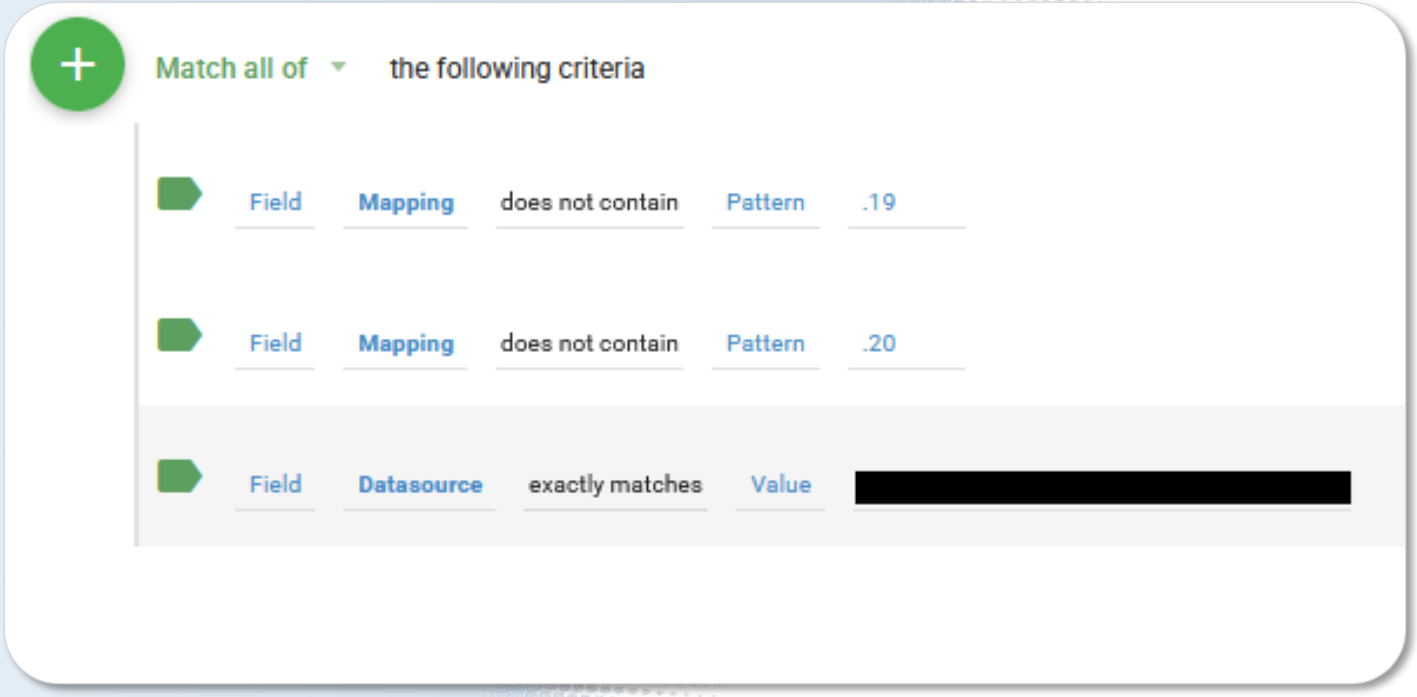
Initial Setup

1. Create a dataset based in the Informer datasource using the Mapping “Field”.
2. **Find the datasource ID of your Unidata Colleague datasource in Informer 5.**
3. Add a filter to your dataset to filter to Colleague. If you have ACYR tables, filter out your individual instance tables. I used the filter “does not contain .19/.20”.



Initial Setup

1. Create a dataset based in the Informer datasource using the Mapping “Field”.
2. Find the datasource ID of your Unidata Colleague datasource in Informer 5.
3. **Add a filter to your dataset to filter to Colleague. If you have ACYR tables, filter out your individual instance tables. I used the filter “does not contain .19/.20”.**



Why include hidden fields?

End users will be able to look up fields that were in Informer 4 but not available in Informer 5. They can check if there are IT notes on why the field was hidden or restricted. The Data Dictionary gives transparency during the migration process.

Informer 5 Metadata

fieldId field name at the database level (sometimes called dictionary name)
mappingId table name
name display name in Informer 5
dataType Informer 5 data type

Restricted – Boolean of Informer setting

Hidden – Boolean of Informer setting

Field Set Name– link to “Field Set” mapping, name

DATA JSON object

FIELDS table

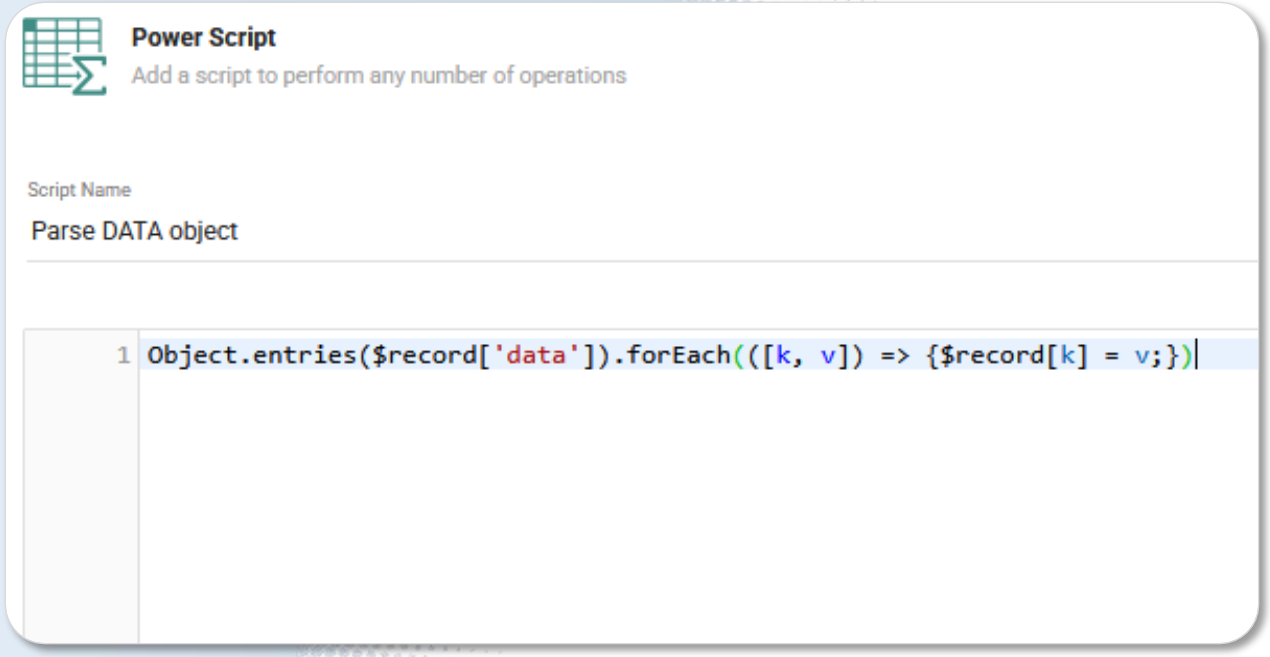
Type	Name	Field		Added ↑	
toggle	Restricted	restricted		ADDED	
quote	Name	name		ADDED	
quote	Mapping	mappingId		ADDED	
toggle	Hidden	hidden		ADDED	
quote	Field Set	fieldSetId		ADDED	
quote	Field	fieldId		ADDED	
quote	Data Type	dataType		ADDED	
{}	Data	data		ADDED	
quote	AltId	altid		ADD	

DATA JSON object

Data


```
{
    "@ID": "STPR.CURRENT.ADDNL.MAJORS",
    "type": "I",
    "width": "5",
    "format": "5L",
    "isSbPlus": false,
    "origType": "I",
    "alignment": "L",
    "attribute": 0,
    "expression": "SUBR(\"CC.GET.DATE.SUBSET\",STPR.ADDNL.MAJORS,STPR.ADDNL.MAJOR.STAR
    "association": "",
    "description": "These are the current additional majors for a student",
    "columnHeader": "CurrentýAddnlýMajors",
    "conversionCode": "",
    "multivalueCode": "S",
    "rawDescription": "CurrentýAddnlýMajors",
    "sbPlusExpression": "",
    "sbPlusDotDictionary": ""
}
```

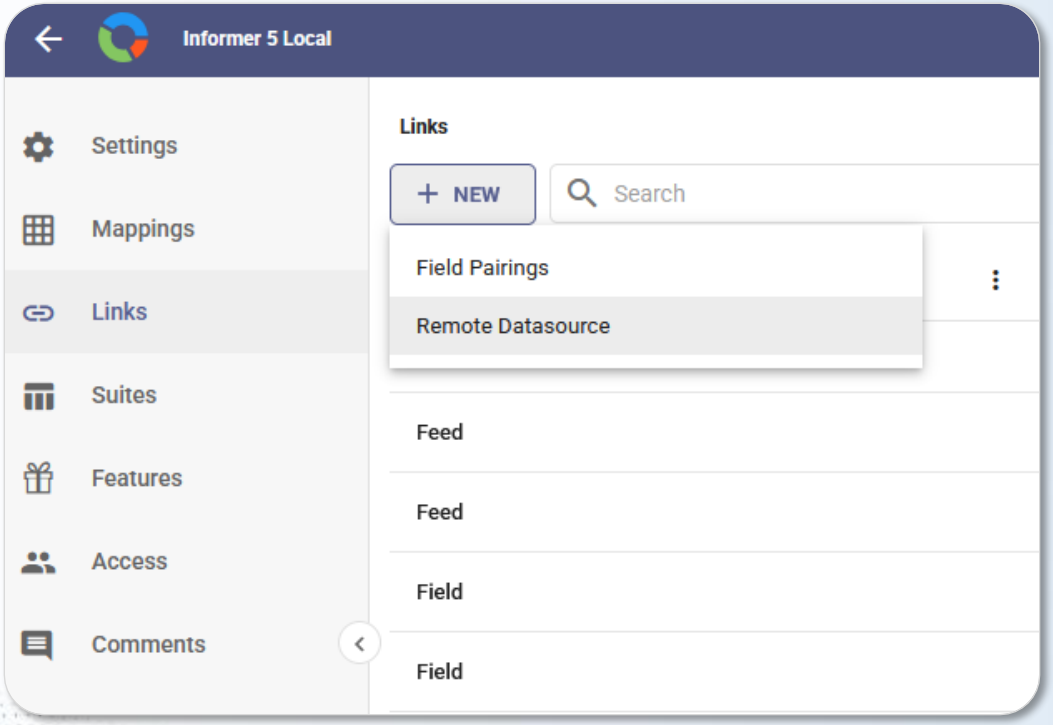

Power Script code to transform any JSON object into columns

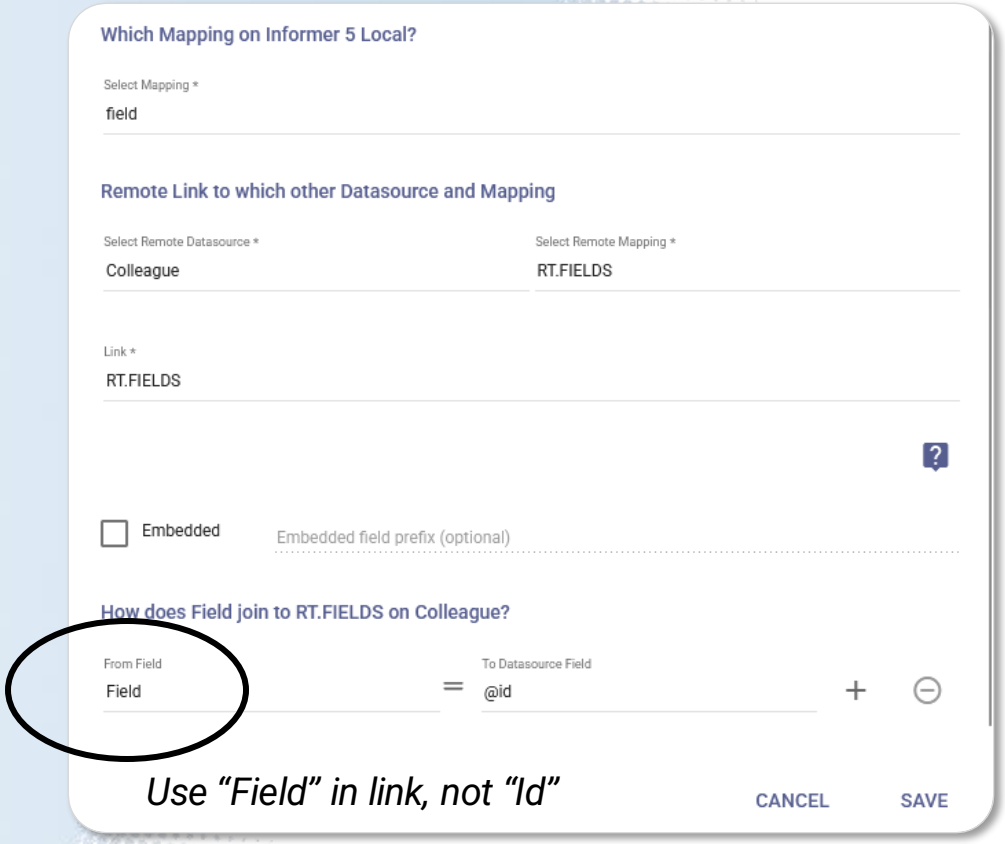


Colleague Metadata “RT.FIELDS”

Add a remote source link to grab Colleague metadata info

Note: This table does not contain all computed columns, only those written in Envision. This will come in handy later. This is also why we do not base the report in RT.FIELDS. Informer is how we can generate a full list.





Use “Field” in link, not “Id”

Colleague Metadata “RT.FIELDS”

Type	Name	Field		Added	
	Rtflds Virtual Field Def	RTFLDS.VIRTUAL.FIELD.DEF		ADDED	
	Rtflds Validation Table	RTFLDS.VALIDATION.TABLE		ADDED	
	Rtflds Validation File	RTFLDS.VALIDATION.FILE		ADDED	
	Rtflds Val Table Application	RTFLDS.VAL.TABLE.APPLICATION		ADDED	
	Rtflds Subroutine Args	RTFLDS.SUBROUTINE.ARGS		ADDED	
	Rtflds Subr Args Pointers	RTFLDS.SUBR.ARGS.POINTERS		ADDED	
	Rtflds Description	RTFLDS.DESCRIPTION		ADDED	
	Rtflds Demand Fields	RTFLDS.DEMAND.FIELDS		ADDED	
	@id	@ID		ADD	

Colleague Metadata “RT.FIELDS”

Virtual Field Def	A handy reference of the computed column logic
Validation Table	If codes maintained on VAL, this is the table name
Val Table Application	If codes maintained on VAL, this is application
Validation File	If validated against PK of a database file/mapping/table
Demand Fields	List of fields declared in Colleague Studio required for logic
Subroutine Args	If a subroutine is called, this lists the fields used as arguments
Description	Description written in Colleague Studio

Why not RTFLDS.SUBROUTINE.NAME?

This is a wrapper that just contains the name of the computed column with a prefix.

Why not base the report in RT.FIELDS?

This table does not contain all computed columns, the Informer scanning does!

Colleague Metadata - Rules

RLDE

Screen where rules are defined as criteria based on Colleague data fields and computed columns

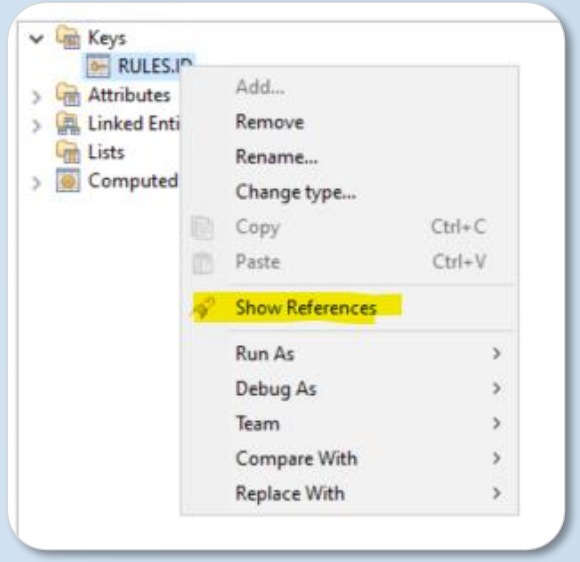
- Rule definition contains the data field names that we can reference in the Data Dictionary dataset
- Possible that rules are defined that aren't actually used

Colleague maintenance screens

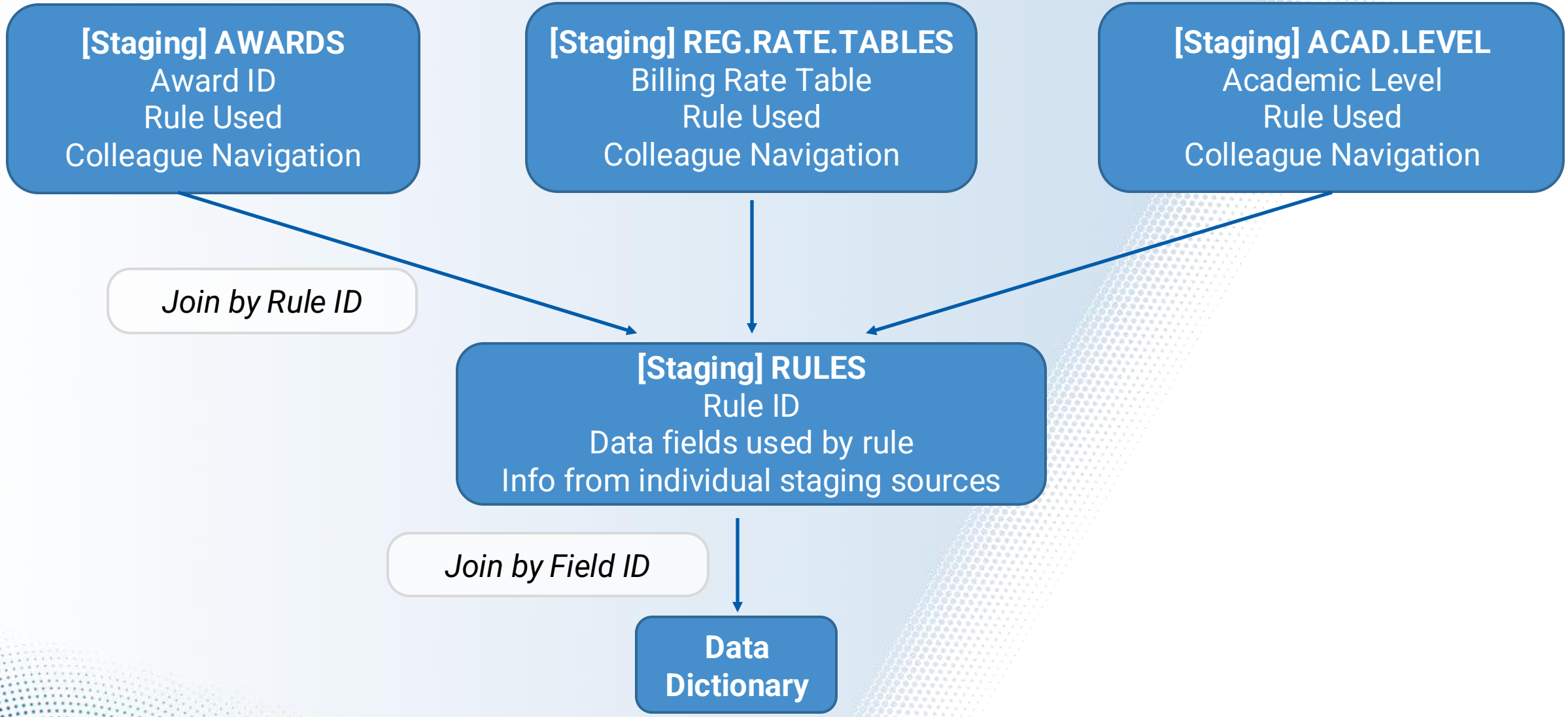
After defining a rule, you can add it to various maintenance screens to tell Colleague how to evaluate records during processes

Easiest way to find all the maintenance places is to use the Colleague Studio "Find references" (right-click RULES.ID field)

- Award eligibility
- Billing rules
- Academic standings
- etc



Colleague Rules Metadata



Calculations

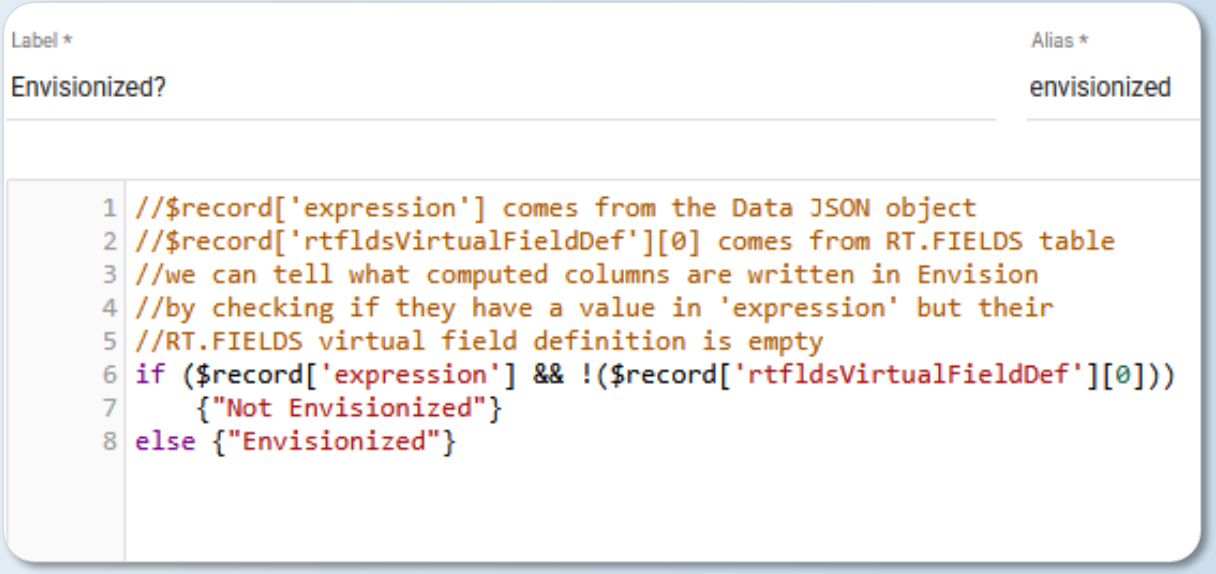
- Formatted Description
- Envisionized?
- Subroutine Name
- Reference Fields
- Validation
- Data field or Computed Column
- Multivalued or Singlevalued

Calculations

- Formatted Description
- Envisionized?
- Subroutine Name
- Reference Fields
- Validation
- Data field or Computed Column
- Multivalued or Singlevalued

Calculations

- Formatted Description
- **Envisionized?**
- Subroutine Name
- Reference Fields
- Validation
- Data field or Computed Column
- Multivalued or Singlevalued



Calculations

- Formatted Description
- Envisionized?
- **Subroutine Name**
- Reference Fields
- Validation
- Data field or Computed Column
- Multivalued or Singlevalued

ubroutine Name

```
var raw_code = null;
var virtual_field = $record['rtfields_rtfldsVirtualFieldDef'];


//The raw code of the computed column, can't use a direct coalesce since
//the virtual def is nested in array so check for existence and if so,
//then pull top value.
if (virtual_field[0])
    {raw_code = virtual_field[0][0];}
else {raw_code = $record['expression'];}

var subr_name = null;

//only try to extract subroutine name if there is code and there is a subroutine call in the code
if (raw_code && raw_code.indexOf("SUBR(") >= 0){
    //find where the subroutine call begins. since indexOf gives the index of the
    //beginning of the substring, add 5 to skip over the literal string SUBR("
        var subr_idx = raw_code.indexOf("SUBR(")+5;
    //find 2nd instance of literal character double quote
        var dbl_qt = raw_code.indexOf("\"",subr_idx+1) ?? 0;
        var sing_qt = raw_code.indexOf("\'",subr_idx+1) ?? 0;
        var subr_end_idx = _.min([dbl_qt, sing_qt].filter(v => v > -1));
    if (raw_code[subr_idx] != "-")
        {
    //extract the subroutine name from the raw code using indexes above
        var subr_name = raw_code.slice(subr_idx+1, subr_end_idx)

    //non-Envision fields have a prefix "CC." so remove that
        if (!(virtual_field[0]))
            {
                subr_name = subr_name.slice(3);
            }
    //S.MIO.TRANS is just a translation to grab a field from another
    //file, remove this reference
        if (subr_name == "S.MIO.TRANS")
            {
                subr_name = null;
            }
        }
}

$record['subroutineName'] = subr_name;

```

Calculations

- Formatted Description
- Envisionized?
- Subroutine Name
- **Reference Fields***
- Validation
- Data field or Computed Column
- Multivalued or Singlevalued

```
if (subr_name)
    {
        var demand_fields = $record['rtfields_rtfldsDemandFields'][0] ?? [];
        var subr_args = $record['rtfields_rtfldsSubroutineArgs'][0] ?? [];
        $record['Referenced Fields'] = _.uniq([...demand_fields, ...subr_args]);
        $record['Referenced Fields'] = $record['Referenced Fields'].filter(v=> v!= "A.KEY")
    }
```

**this is only for computed columns appearing in RT.FIELDS. Getting this info for all computed columns can be retrieved in theory. It would require parsing the 'expression' and validating values as field IDs (e.g. Power Script of some kind).*

Calculations

- Formatted Description
- Envisionized?
- Subroutine Name
- Reference Fields
- **Validation**
- Data field or Computed Column
- Multivalued or Singlevalued

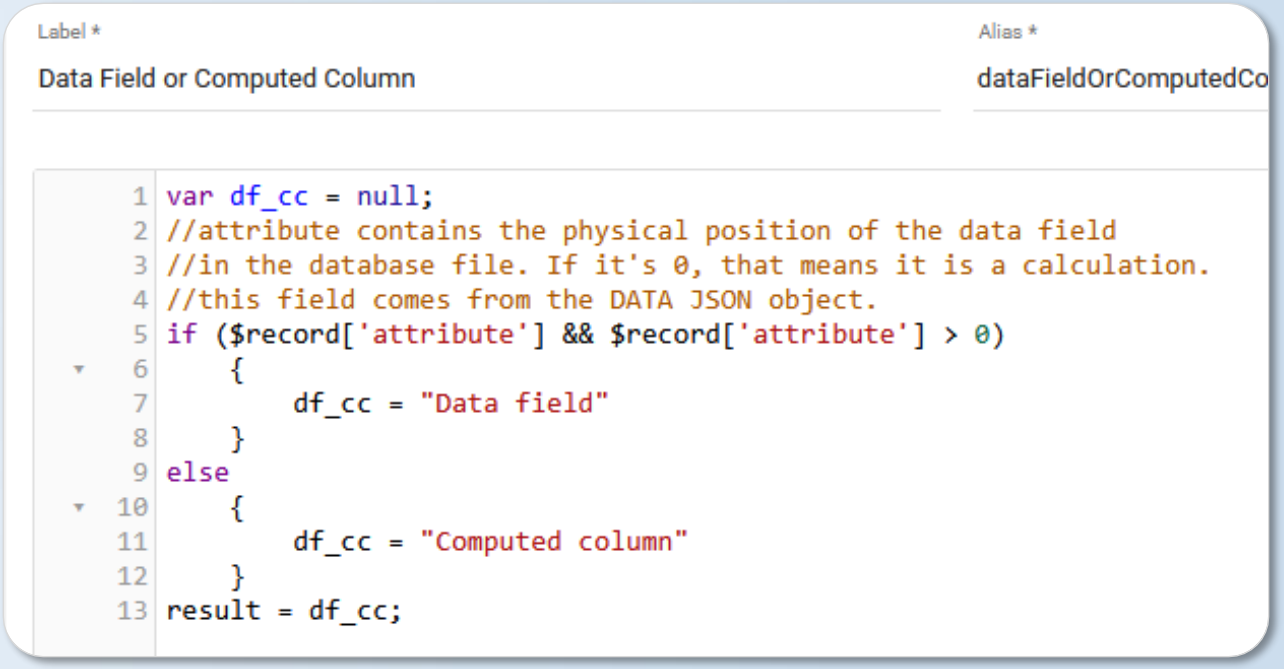
Script Name

Validation

```
1  var val_details = null;
2  //validation can be via a valcode table (maintained on VAL screen)
3  //or a Colleague table(file) key
4  var val_file =  $record['rtfields_rtfldsValidationFile']
5  var val_table = $record['rtfields_rtfldsValidationTable']
6  var val_appl = $record['rtfields_rtfldsValTableApplication']
7
8
9
10 if (val_file[0] || val_table[0])
11     {$record['validated'] = "Yes";
12       if (val_appl[0])
13           {val_details = "Valcode: " + val_appl + "-" + val_table}
14      else {val_details = "Table Key: " + val_file}
15     }
16
17 $record['validationDetails'] = val_details;
```

Calculations

- Formatted Description
- Envisionized?
- Subroutine Name
- Reference Fields
- Validation
- **Data field or Computed Column**
- Multivalued or Singlevalued



Calculations

- Formatted Description
- Envisionized?
- Subroutine Name
- Reference Fields
- Validation
- Data field or Computed Column
- **Multivalued or Singlevalued**

Label *

Multivalued/Singlevalued

```
1 var mv_sv = "Singlevalued";
2 //association and multivalueCode both come
3 //from DATA JSON object
4 if ($record['association'])
5     {mv_sv = "Multivalued (Assoc)";}
6 else if ($record['multivalueCode'] == "M")
7     {mv_sv = "Multivalued";}
8
9 result = mv_sv;
```

Using the dataset

What are our top used subroutines?

Subroutine Name	Count
X.S.PROGRAM.LOANS	41
X.S.HOME.ADDRESS.ID	36
S.GET.ACTIVE.PROGRAMS	35
S.GET.DATE.SUBSET	31

Which fields are using ACAD.PROGRAMS to validate?

Validation Details is Table Key: ACAD.PROGRAMS

Field ID	Cour
ACAD.PROGRAMS.ID	
ACLV.ACAD.PROGRAM	
ACPG.RELATED.PROGRAMS	
APPL.ACAD.PROGRAM	
CEQ.ACAD.PROGRAMS	
CNST.ESIS.STU.PGM.1ST.PGMS	
DABLK.ACAD.PROGRAM	

How many rules use cohort codes?

Used in Rules | Field ID like *COHORT* | Filter

Rows: Used In Award Rules + 2 more | Colur

Used In Award Rules	Used In Billing Rules	Used In Standing/transcript Rules	Count
Yes	Yes	(Empty)	2
Yes	(Empty)	(Empty)	3
(Empty)	Yes	(Empty)	2

How many multivalued computed columns do we have?

Data Field or Computed Column	Singlevalued	Multivalued	Multivalued (Assoc)
Data field	6947	1224	1611
Computed column	4453	1038	152

Auditing Field Sets

entrinsicon

Field Set Name is Birth Date, Disabilities, Do Not Use - Archived... | Restricted is False | Filter | Match All

Add Group | RESET VIEW | Columns | Aggregates | Option

Field ID fieldId	Mapping mappingId	Data Type dataType	Restricted restricted	Hidden hidden	Field Set Name field_set_name	Data Field or Computed Column dataFieldOrComputedColumn	Subroutine Name subroutineName	Multivalued/Singl multivaluedSinglev
SA.SSN	SA.ACYR	keyword_text	false	false	SSN	Computed column		Singlevalued
SSN	SA.ACYR	keyword_text	false	false	SSN	Computed column		Singlevalued
MRR.CRNT.SSN	MRR.ACYR	keyword_text	false	false	SSN	Data field		Singlevalued
MRR.ORIG.SSN	MRR.ACYR	keyword_text	false	false	SSN	Data field		Singlevalued
D01.SSL.SSN	D01.SSL	keyword_text	false	false	SSN	Data field		Singlevalued
OE.SSN.CC	ORG.ENTITY	keyword_text	false	false	SSN	Computed column	S.GET.PERSON.SSN	Singlevalued
TRO.STU.SSN	TRANSCRIPT.ORDERS	keyword_text	false	false	SSN	Data field		Singlevalued
D01.SSL.DATE.OF.BIRTH	D01.SSL	date	false	false	Birth Date	Data field		Singlevalued
STIW.BIRTH.DATE	ST.IPEDS.WORK	date	false	false	Birth Date	Data field		Singlevalued
INSTA.PERSON.FIRST.NAME	INSTITUTIONS.ATTEND	keyword_text	false	false	Legal Name	Computed column		Singlevalued
INSTA.PERSON.SORT.NAME	INSTITUTIONS.ATTEND	keyword_text	false	false	Legal Name	Computed column		Singlevalued
X.SA.SSN	SA.ACYR	keyword_text	false	false	SSN	Computed column		Singlevalued
BIRTH.DATE	PERSON	date	false	false	Birth Date	Data field		Singlevalued
P1098.MASK.SSN	PARM.1098	keyword_text	false	false	SSN	Data field		Singlevalued
P1098.T.SSN.CERT.BOX.CODE	PARM.1098	keyword_text	false	false	SSN	Data field		Singlevalued
APP.SORT.FIRST.NAME	APPLICANTS	keyword_text	false	true	Legal Name	Computed column		Singlevalued
APP.SORT.NAME	APPLICANTS	keyword_text	false	true	Legal Name	Computed column		Singlevalued
CRS.PREREQS	COURSES	keyword_text	false	true	Do Not Use - Archived by Ellucian	Data field		Singlevalued
ETHNIC	PERSON	keyword_text	false	true	Do Not Use - Archived by Ellucian	Data field		Singlevalued
JOINT.MAIL.LABEL.INFORM	PERSON	keyword_text	false	true	Legal Name	Computed column	S.ANY.NAME.ADDRESS	Multivalued
PCC.BIRTH.DATE	PERSON	double	false	true	Birth Date	Computed column		Singlevalued
PCC.BIRTHDAY	PERSON	double	false	true	Birth Date	Computed column		Singlevalued

Future Enhancements

Informer 5 reporting usage

- Statistics on # of ad-hoc queries and datasets that utilize the field
- Which Teams are using the field the most?
- Is the field used in an Informer link?

Permission structure

- Which users/teams have access to each field?
- Is the field in a report that is sent to an external source?

Further parsing of the expression

- Better parsing of all field references

Build out more field references from Colleague

- Flag if a field is referenced in SLCR paragraph
- Parameters/setups/defaults in Colleague that use rules

entrinsicon

Thank you

Sam Parsons

sparsons@monroeu.edu | 646-393-8643