

Grouper entity data fields for ABAC

Wiki Home	Grouper Release Announcements	Grouper Guides	Grouper Deployment Guide	Community Contributions	Internal Developer Resources
---------------------------	---	--------------------------------	--	---	--

This is in Grouper v5+

- [Terminology](#)
- [Description](#)
- [Data field flow](#)
- [Previous state](#)
- [Data field and row diagram](#)
- [Configure data field privacy realms](#)
- [Configure data fields](#)
- [Configure data rows](#)
- [Configure data providers](#)
- [Configure data provider queries](#)
- [Configure data provider real time query](#)
- [Grouper data field dictionary](#)
- [Grouper dependency SQL caching](#)

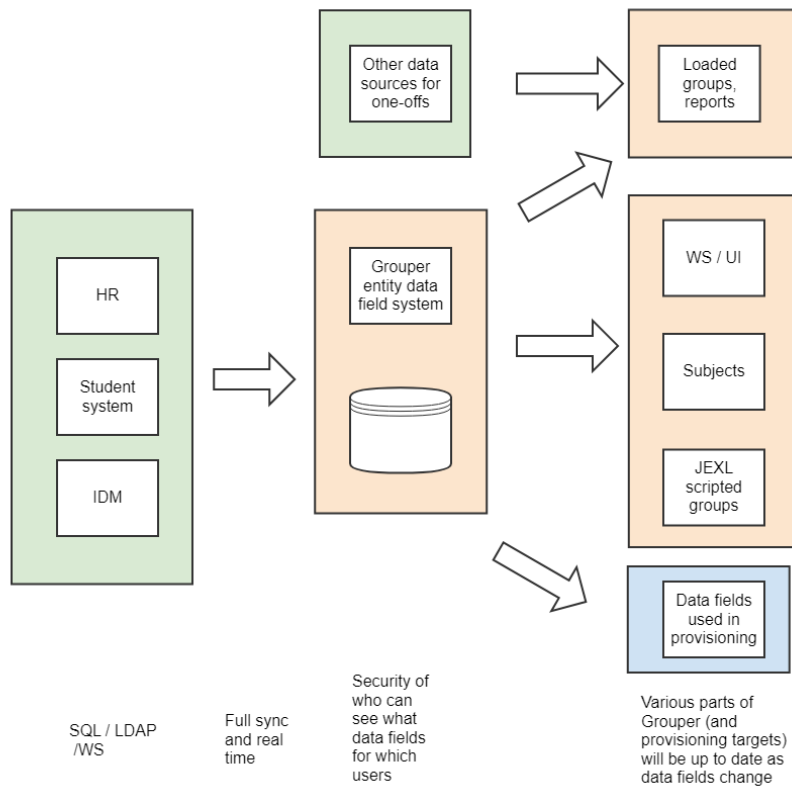
Terminology

- "data field" is a user attribute. This is named "data field" since "attribute" is used in many other places, e.g. the attribute framework

Description

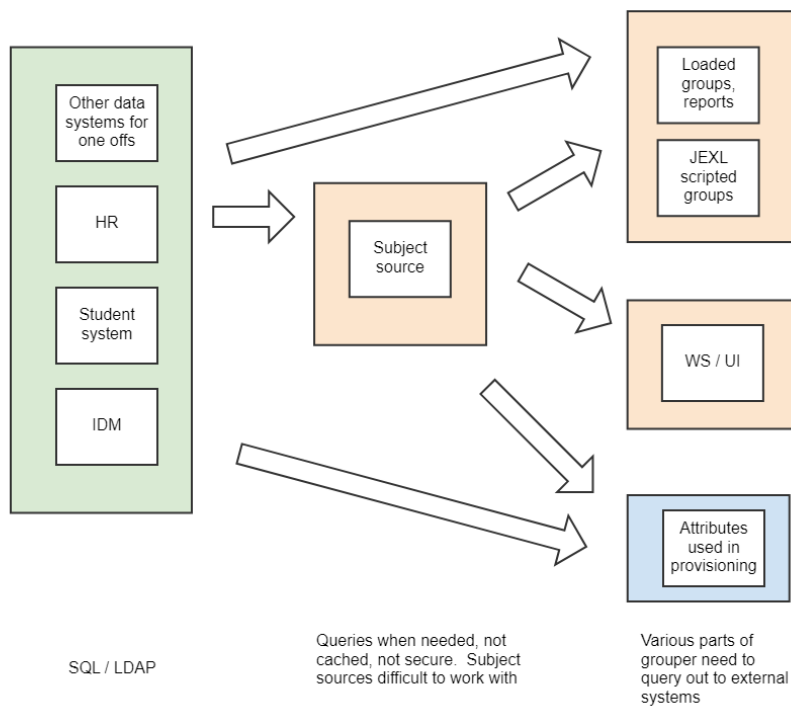
- Data field values are assigned to users, groups, or globally available
- The data can be single or multi-valued
- The data can be structured in a row (the data in a cell can still be multi-valued)
- The data is stored in Grouper, updated in real time and full syncs
- Point in time history will be maintained
- Security on data fields will ensure that private data remains private
- The data will be stored efficiently so it does not take a lot of space and queries are efficient
- Data fields are documented with examples so users can easily request access, see what data fields represent and how to use them
- Data fields can be configured in the UI
- The data can be used:
 - To construct ABAC policies on groups (scripted group based on data field values)
 - Subject sources can be replaced by a data field source (this is the future direction, and all subject sources will eventually need to be migrated)
 - Provisioning data about users to other systems
 - Reports about access and users
 - Etc

Data field flow

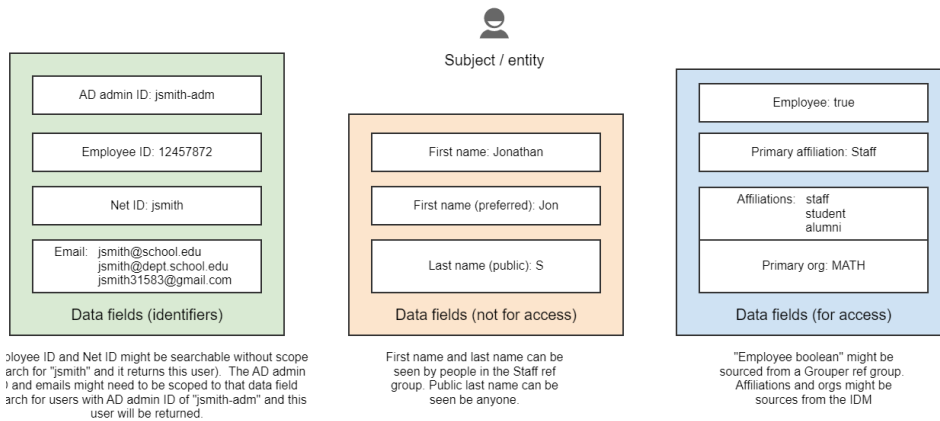


- In this diagram, the green data field resolvers are either cached (e.g. source systems), or not (one off which doesnt need the overhead of PIT etc)
 - A provisioner target could have entity data field values for users

Previous state



Data field and row diagram



Affiliation name	Center	End date
Staff	Arts and Sciences	2023/12/03
Student	Continuing education	2022/05/12
Alumni	Business	2007/05/10
Data row (affiliations)		

Title	Org	Full time
Lecturer	MATH	true
Service provider	PSFT	false
Data row (payroll records)		

Example usages:

- Provision any identifier to a target without having to "resolve the subject"
- Make a JEXL scripted group: People who have a payroll data row where org is "MATH" and have an affiliation data row where affiliation name is Staff or Faculty with an End date in the next month. Put a rule on that group for the Business Analyst to review people who might need to be renewed.
- Load users from Zoom and match accounts to users in Grouper by any of their email addresses
- A staff member creates a report where a column represents if the user in the service is an Employee
- A help desk worker can see the history of affiliations and troubleshoots access by seeing that the user's payroll org recently changed

Configure data field privacy realms

A privacy realm is a configuration for privacy of one or many data fields. Re-use them as much as you can to reduce the number of configurations

Privacy realms

Config id	crashplan	
Privacy realm name	☐ EL?	crashplan *
name of this privacy realm, not really used, just here to configure the realm		
Is realm public?	☐ EL?	☒ Default value (False) ☐ True ☐ False if this field can be seen by anyone, even if not authenticated. Default value is 'false'.
Is realm authenticated?	☐ EL?	☒ Default value (False) ☐ True ☐ False if this field can be seen by anyone who is authenticated. Default value is 'false'.
Can sysadmins access?	☐ EL?	☒ Default value (True) ☐ True ☐ False if this field can be seen by grouper sysadmins who are not in the viewers group. Default value is 'tru
Viewers group name	☐ EL?	*
group name of the group who can view and use fields marked in this privacy realm		
Updaters group name	☐ EL?	*
group name of the group who can update and use fields marked in this privacy realm		
Readers group name	☐ EL?	*
group name of the group who can read and use fields marked in this privacy realm		

Configure data fields

Each data field or row column is configured as a data field.

Entity data fields

Data field actions ▾

Config id cp_active

Field aliases

☐ EL?

cp_active *

aliases that this field is referred to as

Field privacy realm

☐ EL?

crashplan ▾ *

privacy realm for people who can see or use this data field

Description html

☐ EL?

If the user account is active in Crashplan. If so then the user is eligible to be billed for space used. *

Description html

Data owner html

☐ EL?

Crashplan service owner

Data owner html

How to get access html

☐ EL?

Open a ticket with a@b.c

How to get access html

Zero to many examples

☐ EL?

entity.hasRow('cp_user', "cp_active")

Zero to many examples html

Field multivalued?

☐ EL?☒ Default value (False) ☐ True ☐ False
if this field can have multiple values. Default value is 'false'.

Field datatype

☐ EL?

boolean ▾

data type for this field. Default value is 'string'.

Field data structure

☐ EL?

rowColumn ▾

data structure for this field. Default value is 'attribute'.

Field data use

☐ EL?

access ▾

use of this field. If it is access related then it will be available in an abac script. Default value is 'access'.

Field data calculated

☐ EL?☒ Default value (False) ☐ True ☐ False
if this field is calculated from multiple providers. If it is calculated from one provider, it can be configured in that provider. Default value is 'false'.

Field data calculated script

☐ EL?

script to build this data field value from multiple providers. Default value is 'false'.

Field data source

☐ EL?

provider ▾

field source. Could be sourced from a data provider. Could be sourced from a group membership. Default value is 'provider'.

Field data store in PIT

☐ EL?☒ True ☐ False

should this field be stored in PIT. All identifiers will be in PIT.

Days to store in PIT

☐ EL?

200

how many days to store PIT. Default value is '730'.

Field data assignable to

☐ EL?

individuals



Is this assignable to groups or individuals.

Field data LOV

☐ EL?

☒ Default value (False) ☐ True ☐ False

Are the values for this data fields a list of values with less than 100ish options? Used for validation of ABAC scripts. Default value is 'false'.

Submit

Cancel

Configure data rows

Rows are configured as a "table". The columns are data fields.

Data rows

Data row actions ▾

Config id cp_user

Description html

☐ EL?

A user row has attribute for the account such as active and matched to Penn account. There is one row per user account in crashplan *

Description html

Data owner html

☐ EL?

Crashplan service owner

Data owner html

How to get access html

☐ EL?

Open a ticket with a@b.c

How to get access html

Zero to many examples

☐ EL?

Crashplan users who are active in Crashplan, not blocked, known to match a user at Penn, in the med school, not an active member at Penn (e.g. they left) or their account is locked, and

Zero to many examples html

Data row config

Data row config

Row aliases

☐ EL?

cp_user *

aliases that this row is referred to as

Row privacy realm

☐ EL?

crashplan *

privacy realm for people who can see or use this data row

Row number of data fields

☐ EL?

5 *

number of fields in this row

Row data field 1

Configure row data field

1 - Col data field config id

☐ EL?

cp_user_id *

data field for this column

1 - Row key field

☐ EL?

☐ Default value (False) ☒ True ☐ False

If this single valued column is the key or part of composite key to uniquely identify this row for this entity. Default value is 'false'.

Row data field 2

Configure row data field

2 - Col data field config id

☐ EL?

cp_active *

data field for this column

2 - Row key field

☐ EL?

☒ Default value (False) ☐ True ☐ False

If this single valued column is the key or part of composite key to uniquely identify this row for this entity. Default value is 'false'.

Row data field 3

Configure row data field

3 - Col data field config id

☐ EL?

cp_blocked *

data field for this column

3 - Row key field

☐ EL?
 ☒ Default value (False)
 ☐ True
 ☐ False

If this single valued column is the key or part of composite key to uniquely identify this row for this entity. Default value is 'false'.

Row data field 4

Configure row data field

4 - Col data field config id

☐ EL?

cp_known

data field for this column

*

4 - Row key field

☐ EL?
 ☒ Default value (False)
 ☐ True
 ☐ False

If this single valued column is the key or part of composite key to uniquely identify this row for this entity. Default value is 'false'.

Row data field 5

Configure row data field

5 - Col data field config id

☐ EL?

cp_org

data field for this column

*

5 - Row key field

☐ EL?
 ☒ Default value (False)
 ☐ True
 ☐ False

If this single valued column is the key or part of composite key to uniquely identify this row for this entity. Default value is 'false'.

Submit

Cancel

Institute of Higher Education

Configure data providers

A data provider is a set of queries that load data into Grouper in real time or full sync

Home > Miscellaneous > Entity data fields > Data providers > Edit data provider

Data providers

Data provider actions ▼

Config id

crashplan

Name

☐ EL?

crashplan

data provider name, not really needed or used, but there to setup the provider

*

Submit

Cancel

Institute of Higher Education

Configure data provider queries

These select data from the target to populate Grouper with data field values. A single provider can have multiple queries. Each query has one provider.

Data provider queries

Data provider query actions ▾

Config id crashplan_user

Data provider query config

dataProviderQueryConfig

Provider config id ☐ EL? crashplan *

data provider config id

Provider query type ☐ EL? sql *

data provider query type

Provider query sql config id ☐ EL? grouper *

SQL config id

Provider query sql query ☐ EL? select subject_id, cp_user_id, cp_active, cp_blocked, cp_known, cp *

SQL query

Provider query data structure ☐ EL? row *

Data structure

Provider query row config id ☐ EL? cp_user *

Data row to link to

Provider query subject id attribute ☐ EL? subject_id *

Attribute which links this data to subjects

Provider query subject id type ☐ EL? subjectId

Which type of subject id

Provider query subject source id ☐ EL?

which subject source this is a subject id for

Provider query number of data fields ☐ EL? 5 *

number of fields in this row

Provider query data field 1

Configure provider query data field

1 - Provider data field config id ☐ EL? cp_user_id *

data field for this column

1 - Provider mapping type ☐ EL? attribute *

mapping type for this data field, e.g. could be translation eventually

1 - Provider data field attribute ☐ EL? cp_user_id *

mapping type for this data field

Provider query data field 2

Configure provider query data field

2 - Provider data field config id ☐ EL? cp_active *

data field for this column

2 - Provider mapping type ☐ EL? attribute *

mapping type for this data field, e.g. could be translation eventually

2 - Provider data field attribute

EL?

cp_active

✖

mapping type for this data field

Provider query data field 4

Configure provider query data field

4 - Provider data field config id

EL?

cp_known

▼

*

data field for this column

4 - Provider mapping type

EL?

attribute

▼

*

mapping type for this data field, e.g. could be translation eventually

4 - Provider data field attribute

EL?

cp_known

✖

mapping type for this data field

Provider query data field 5

Configure provider query data field

5 - Provider data field config id

EL?

cp_org

▼

*

data field for this column

5 - Provider mapping type

EL?

attribute

▼

*

mapping type for this data field, e.g. could be translation eventually

5 - Provider data field attribute

EL?

cp_org

✖

mapping type for this data field

Submit

Cancel

➤ Institute of Higher Education

Configure data provider real time query

This helps the change log know which data to update

Data provider change log queries

Data provider change log query action

Config id

crashPlan

*

The Config id is an alphanumeric key for the data provider change log queries that will be referred to from places that use the data provider change log query. It is also used in the configuration keys.

Data provider change log query configuration

GrouperDataProviderChangeLogQueryConfiguration

*

Data provider change log query configuration

Data provider change log query config

dataProviderChangeLogQueryConfig

Provider config id

☐ EL?

crashplan

*

data provider config id

Provider change log query type

☐ EL?

sql

*

data provider change log query type

Provider change log query sql config id

☐ EL?

grouper

*

SQL config id

Provider change log query sql query

☐ EL?

*

SQL change log query

Provider change log query primary key attribute

☐ EL?

*

Change log attribute that is the primary key

Provider change log query timestamp attribute

☐ EL?

*

Change log attribute that contains the timestamp for when this row was added, e.g. a timestamp or number field (number of millis since 1970). Should be indexed where possible.

Provider change log query subject id attribute

☐ EL?

*

Change log attribute which links this data to subjects

Provider change log query subject id type

☐ EL?

▼

Which type of subject id

Provider change log query subject source id

☐ EL?

which subject source this is a subject id for

Submit

Cancel