

## IDENTIFYING STUDENTS

To support the collection of quality and usable data resulting in the collation of sector level datasets, HESA has a requirement to identify unique students and their engagements. Part of this requirement extends to identifying where people have multiple interactions with higher education, to ensure continuity of data over time.

## RETURNING IDENTIFIERS

In the Student data model, providers are required to return the following identifiers:

SID	This is allocated by the provider when a student starts studying with them. It uniquely identifies a student at a provider.
Persistent HUSID	An identifier that is unique to students across providers. This is to uniquely identify a student across their entire study in higher education.
Other identifiers	Anything else the provider needs to return against the student, e.g. SSN, TRN or Own student identifier. These are all listed in the PersonIdentifier.IDTYPECODE field.

## WORKED EXAMPLES

### First time we see a student in the Student record

The first time a student starts studying at a provider, they should allocate a SID to the student. This will then become the Persistent HUSID for the student throughout their higher education studies.

YEAR ONE	YEAR TWO	YEAR THREE
SID		
Persistent HUSID = SID		

### Second time we see a student in the Student record

The next time a student studies at a different provider (or any subsequent times), the next provider should allocate a SID to the student. They should then determine what the persistent HUSID for that student was, by using the HUSID lookup service (or equivalent service).

YEAR ONE	YEAR TWO	YEAR THREE
SID		
Persistent HUSID		

If a provider returns a student to HESA and then finds out the student already has a persistent HUSID (either as they didn't know they previously studied in higher education before or are concurrently studying at two providers and so returned the SID as the persistent HUSID; or because the provider wasn't able to find this to start with and so returned nothing), they should then return the field to submit the correct value.

YEAR ONE		YEAR TWO	YEAR THREE
SID			
Persistent HUSID = SID	*Persistent HUSID sent		

\*Please note, this activity can happen at any point in the year; the above example is only halfway as a visual representation.

This means that a provider should not be held up from returning a student to HESA, simply because they haven't been able to find out what the persistent HUSID is.

#### Returning continuing students across the 2018/19 and 2019/20 boundary

Where a student has already started their education before September 2019 HESA will have migrated the data across to allow them to continue to be returned from 2019/20 onwards. The previous HUSID value will be allocated as the persistent HUSID in these cases. The HUSID can also become the SID (the Student.SID field will allow both structures of identifiers to be returned).

YEAR ONE	YEAR TWO	YEAR THREE
SID = HUSID		
Persistent HUSID = HUSID		

#### When a student transfers course at the same provider

If a student starts studying at a provider, and then decides to change to a different course with the same provider, you would keep the SID and persistent HUSID the same. Only their curriculum data would need to change in this case.

YEAR ONE	YEAR TWO	YEAR THREE
SID		
Persistent HUSID = SID		

#### Mergers

Currently we aren't sure what will happen with provider mergers or de-mergers. More details to follow.

#### UNIQUE KEYS

Previously we uniquely identified instances with the following data items:

- HUSID
- UKPRN
- NUMHUS

In the Data Futures model, we will uniquely identify student registrations using the following data items:

- SID (which includes the UKPRN)

- NUMHUS (on the engagement entity)
- NUMREG (on the registration entity)

## STRUCTURES OF IDENTIFIERS

Structure of HUSIDs in the record are (note: this may not apply to persistent HUSIDs):

First 2 digits:	Year of entry into provider (last 2 digits of year)
Next 4 digits:	HESA institution identifier + 1000
Next 6 digits:	6-digit reference number internally allocated by provider
Last digit:	Check digit

Structure of SIDs in the record are:

First 2 digits:	Year of entry into provider (last 2 digits of year)
Next 8 digits:	UKPRN
Next 6 digits:	6-digit reference number internally allocated by provider
Last digit:	Check digit