

## Specification for the migration of legacy data to the Data Futures data model.

### Migration.Student

Version 0.6.0: Revision history at foot of document

#### Overview

This document provides the specification for the migration of legacy data into the Data Futures Student entity and Disability sub-entity.

This is one of a series of specifications produced to translate or map legacy data submitted against the CXX051/54 collections into the Data Futures data model (056). The data migration activity will allow quality assurance checks to be run on the first Data Futures collection in 2022/23 for data submitted by providers.

#### Entities covered in this specification

Entity name	Entity description
<b>Migration.Student</b>	Provides information on how data has been migrated into the Student entity.
<b>Migration.Student.Disability</b>	Provides information on how data has been migrated into the Disability sub-entity of Student.

**Migration.Student** - This entity represents student information; each student can be identified by a unique combination of student identifier and HE provider identifier.

**Migration.Student.Disability** - This sub-entity of Student represents disability information for each student. In legacy data, only one disability is recorded per student, so this is reflected in the migrated data.

These entities are only created for students within the migrated population (see Z\_POPMIGRATION), this applies to both the Student record (051) and the Student Alternative record (054).

### Depend upon fields

- Student.BIRTHDTE
- Student.CARER (051)
- Student.ETHNIC (051)
- StudentEquality.ETHNIC (054)
- Student.GENDERID (051)
- StudentEquality.GENDERID (054)
- Student.HUSID
- Student.NATION (051)
- StudentEquality.NATION (054)
- Student.OWNSTU
- Student.RELBLF (051)
- StudentEquality.RELBLF (054)
- Student.SCN (051)
- Student.SDEPEND (051)
- Student.SEXID (051)
- StudentEquality.SEXID (054)
- Student.SEXORT (051)
- StudentEquality.SEXORT (054)
- Student.SSN
- Student.TTACCOM (051)
- Student.TTPCODE (051)
- Student.UCASPERID (051)
- Student.UKPRN
- Student.ULN
- Student.DISABLE (051)
- StudentEquality.DISABLE (054)
- Migration.Z\_COLLECTION\_YEAR
- Provider.XINSTCOU01

### Additional information

Where there is a record for the same student (based on HUSID) at the same provider (based on UKPRN) in both the latest collection year and the previous collection year, student information has been taken from the latest collection year.

Where labels, codes or granularity differ for a migrated field, a value mapping table is provided following the Field mapping table below.

Please be aware that whilst field labels may be the same for legacy and Data Futures, the coverage may differ. The Data Dictionary within the Data Specification provides full lists of Data Futures valid entries and coverage of the data.

## Field mapping

Data Futures field	Data Futures label	Legacy field(s)	Applicable to	Migrated?	Direct mapping?
<b>Student</b>					
SID	Student identifier	HUSID	051, 054	Y	Y
BIRTHDTE	Date of birth	BIRTHDTE	051, 054	Y	Y
CARER	Carer	CARER	051	Y	Y
ETHNIC	Ethnicity	ETHNIC	051, 054	Y	N
FNAMES	First name	FNAMES	051, 054	N	Y
GENDERID	Gender identity [Nothing mapped to code 99 Not available]	GENDERID	051, 054	Y	Y
LANGPREF	Language preference (Wales)			N	
NATION	Nationality	NATION	051, 054	Y	N
OWNSTU	Own student identifier	OWNSTU	051, 054	Y	Y
RELIGION	Religion	RELBLF	051, 054	Y	N
SCN	Scottish Candidate Number	SCN	051	Y	Y
SERLEAVE	Service leaver (Scotland)			N	
SERSTU	Service student (Scotland)			N	
SEXID	Sex identifier	SEXID	051, 054	Y	Y
SEXORT	Sexual orientation	SEXORT	051, 054	Y	N
SNAME16	Surname at 16	SNAME16	051, 054	N	Y
SSN	Student support number	SSN	051, 054	Y	Y
STUDEP	Dependant (Scotland)	SDEPEND	051	Y	N
SURNAME	Surname	SURNAME	051, 054	N	Y
TRANS	Transgender			N	
TTACCOM	Term time accommodation type	TTACCOM	051	Y	N
TTPCODE	Term time postcode	TTPCODE	051	Y	Y
UCASPERID	UCAS personal identifier	UCASAPPID	051	Y	Y
ULN	Unique learner number	ULN	051, 054	Y	Y
<b>Disability</b>					
DISABILITY	Disability	DISABLE	051, 054	Y	N

## Value mapping tables

For those fields which have been migrated but are not a direct mapping (i.e., labels, codes or granularity differ), tables listing the mapping of the valid values are provided below.

Legacy values which have not been mapped are still listed, marked 'Not mapped'.

### 1:1 code mapping

#### 1:1 code mapping is 'Y':

There is only **one** legacy code mapped to **one** 056 code. In some cases, the codes will be the same, e.g. Student.SDEPEND in legacy 01 maps to 01 in Migration.Student.STUDEP. In others they are different codes, Student.ETHNIC '22' maps to '120'. Similarly, the labels may remain the same or differ.

Even for 1:1 code mappings, if the label differs substantially between legacy and 056, the migrated data label will account for both labels to avoid any misinterpretation, particularly for personal characteristic fields e.g. if Legacy 99 'Not known' is a 1:1 mapping to 99 'Not available', then the migrated data label may state 'Not available/ Not known'.

As an exception to this, where 'Information refused' is mapped to 'Prefer not to say', the migration label will state only 'Prefer not to say'.

#### 1:1 code mapping is 'N':

This can be a result of legacy codes either being combined in an 056 code, or split across multiple 056 codes depending on additional fields e.g. some ethnicity codes differ by the provider country and so have been mapped to separate 056 codes accordingly.

As with the 1:1 code mappings, if the label differs substantially between Legacy and 056, the migrated data label will account for both labels to avoid any misinterpretation, particularly for personal characteristic fields.

#### 1:1 code mapping is 'N/A':

This notation is used where the code is new for Data Futures (056) and has no legacy equivalent. These codes are only listed for completeness. Although the new codes should not appear in Migrated data, the 056 code and label values are still listed within the Migrated columns and the 056 label column is marked 'New code'.

#### Dependant

Legacy Student. SDEPEND	Legacy Label	Migration. Student. STUDEP	Migration Label	056 Label (where different to Migration)	1:1 code mapping
01	Young people/children	01	Young people/children		Y
02	Other relatives/friends	02	Other relatives/friends		Y
03	No dependants	03	No dependants		Y
04	Both young people/children & other relatives/friends	04	Both young people/children and relatives/friends		Y
99	Not known	99	Not available/ Not known	Not available	Y
		98	Prefer not to say	New code	N/A

#### Ethnicity

Legacy Student. ETHNIC	Legacy Label	Legacy Provider. XINSTCOU01	Migration. Student. ETHNIC	Migration Label	056 Label (where different to Migration)	1:1 code mapping
10	White	E, W	160	White – English, Scottish, Welsh, Northern Irish or British / White	White - English, Scottish, Welsh, Northern Irish or British	N
10	White	S	161	White – English, Welsh, Northern Irish or Other British / White / White - Scottish	White - English, Welsh, Northern Irish or British	N
10	White	N	162	White – British, Irish, Northern Irish, English, Scottish or Welsh / White	White - British, Irish, Northern Irish, English, Scottish or Welsh	N

## Protective Marking - PUBLIC

Legacy Student. ETHNIC	Legacy Label	Legacy Provider. XINSTCOU01	Migration. Student. ETHNIC	Migration Label	056 Label (where different to Migration)	1:1 code mapping
13	White – Scottish	S	161	White – English, Welsh, Northern Irish or Other British / White / White - Scottish	White - English, Welsh, Northern Irish or British	N
14	Irish Traveller	N	165	White – Irish traveller		Y
15	Gypsy or Traveller	E, W	163	White – Gypsy or Irish Traveller		N
15	Gypsy or Traveller	S, N	164	White – Gypsy or Traveller		N
16	White - Other British	S	161	White – English, Welsh, Northern Irish or Other British / White / White - Scottish	White - English, Welsh, Northern Irish or British	N
		S, N	167	White – Polish	New code	N/A
		S	168	White – Roma	New code	N/A
		S	169	White – Scottish	New code	N/A
		S	170	White – Showman/ Showwoman	New code	N/A
19	Other White background	E, S	179	Any other White background		Y
21	Black or Black British – Caribbean	E, N, S, W	121	Black – Caribbean or Caribbean British		Y
22	Black or Black British – African	E, N, S, W	120	Black – African or African British		Y
29	Other Black background	E, N, S, W	139	Any other Black background		Y
31	Asian or Asian	E, N, S, W	103	Asian – Indian or Indian British		Y

Legacy Student. ETHNIC	Legacy Label	Legacy Provider. XINSTCOU01	Migration. Student. ETHNIC	Migration Label	056 Label (where different to Migration)	1:1 code mapping
	British – Indian					
32	Asian or Asian British – Pakistani	E, N, S, W	104	Asian – Pakistani or Pakistani British		Y
33	Asian or Asian British - Bangladeshi	E, N, S, W	100	Asian – Bangladeshi or Bangladeshi British		Y
34	Chinese	E, N, S, W	101	Asian – Chinese or Chinese British		Y
		N	102	Asian – Filipino	New code	N/A
39	Other Asian background	E, N, S, W	119	Any other Asian background		Y
41	Mixed - White and Black Caribbean	E, N, S, W	142	Mixed or multiple ethnic groups – White or White British and Black Caribbean or Black Caribbean British		Y
42	Mixed - White and Black African	E, N, S, W	141	Mixed or multiple ethnic groups – White or White British and Black African or Black African British		Y
43	Mixed - White and Asian	E, N, S, W	140	Mixed or multiple ethnic groups – White or White British and Asian or Asian British		Y
49	Other mixed background	E, N, S, W	159	Any other Mixed or Multiple ethnic background		Y
50	Arab	E, N, S, W	180	Arab		Y
80	Other ethnic background	E, N, S, W	899	Any other ethnic background		Y

Legacy Student. ETHNIC	Legacy Label	Legacy Provider. XINSTCOU01	Migration. Student. ETHNIC	Migration Label	056 Label (where different to Migration)	1:1 code mapping
90	Not known	E, N, S, W	997	Not known		Y
98	Information refused	E, N, S, W	998	Prefer not to say		Y
		E, N, S, W	999	Not available	New code	N/A

Gender identity

Legacy Student. GENDERID	Legacy Label	Migration. Student. GENDERID	Migration Label	056 Label (where different to Migration)	1:1 code mapping
01	Yes	01	Yes		Y
02	No	02	No		Y
98	Information refused	98	Prefer not to say		Y
		99	Not available	New code	N/A

Nationality

Legacy Student. NATION	Legacy Label	Migration. Student. NATION	Migration Label	056 Label (where different to Migration)	1:1 code mapping
QN	Serbia and Montenegro not otherwise specified	RS	Serbia and Montenegro not otherwise specified	Serbia	N
		ME	Montenegro (Not applicable to migrated data)	Not used	N/A
ZZ	Not known	97	Not known		N
Otherwise		Student.NATION	See valid entries for Student.NATION		Y



## Religion

Legacy Student. RELBLF	Legacy Label	Migration. Student. RELIGION	Migration Label	056 Label (where different to Migration)	1:1 code mapping
01	No religion	20	No religion		Y
02	Buddhist	21	Buddhist		Y
03	Christian	22	Christian		Y
04	Christian - Church of Scotland	23	Christian - Church of Scotland		Y
05	Christian - Roman Catholic	24	Christian - Roman Catholic		Y
06	Christian - Presbyterian Church in Ireland	25	Christian - Presbyterian Church in Ireland		Y
07	Christian - Church of Ireland	26	Christian - Church of Ireland		Y
08	Christian - Methodist Church in Ireland	27	Christian - Methodist Church in Ireland		Y
09	Christian - Other denomination	28	Christian - Other denomination		Y
10	Hindu	29	Hindu		Y
11	Jewish	30	Jewish		Y
12	Muslim	31	Muslim		Y
		32	Pagan	New code	N/A
13	Sikh	33	Sikh		Y
14	Spiritual	80	Any other religion or belief (including Spiritual)	Any other religion or belief	N
80	Any other religion or belief	80	Any other religion or belief (including Spiritual)	Any other religion or belief	N
98	Information refused	98	Prefer not to say		Y

Legacy Student. RELBLF	Legacy Label	Migration. Student. RELIGION	Migration Label	056 Label (where different to Migration)	1:1 code mapping
99	Not known	99	Not available/ Not known	Not available	Y

## Sex

Legacy Student. SEXID	Legacy Label	Migration. Student. SEXID	Migration Label	056 Label (where different to Migration)	1:1 code mapping
1	Male	11	Male		Y
2	Female	10	Female		Y
3	Other	12	Other		Y
		98	Prefer not to say	New code	N/A
		99	Not available	New code	N/A

## Sexual orientation (Optional at providers in Scotland. Required at providers in England, Northern Ireland or Wales)

Legacy Student. SEXORT	Legacy Label	Migration. Student. SEXORT	Migration Label	056 Label (where different to Migration)	1:1 code mapping
01	Bisexual	10	Bisexual		Y
02	Gay man	11	Gay or lesbian		N
03	Gay woman/lesbian	11	Gay or lesbian		N
04	Heterosexual	12	Heterosexual or straight		Y
05	Other	19	Other sexual orientation		Y
98	Information refused	98	Prefer not to say		Y

Legacy Student. SEXORT	Legacy Label	Migration. Student. SEXORT	Migration Label	056 Label (where different to Migration)	1:1 code mapping
		99	Not available	New code	N/A

## Term time accommodation

Legacy Student. TTACCOM	Legacy Label	Migration. Student. TTACCOM	Migration Label	056 Label (where different to Migration)	1:1 code mapping
1	Provider maintained property	01	Provider maintained property		Y
2	Parental/guardian home	02	Parental/guardian home		Y
4	Other	07	Other		Y
5	Not known	08	Not available		N
6	Not in attendance at the provider	08	Not available		N
7	Own residence	04	Own residence		Y
8	Other rented accommodation	05	Other rented accommodation		Y
9	Private-sector halls	06	Private-sector halls		Y

## Disability

Legacy Student. DISABLE	Legacy Label	Migration. Student. DISABILITY	Migration Label	056 Label (where different to Migration)	1:1 code mapping
00	No known disability	95	No known impairment, health condition or learning difference		Y

Legacy Student. DISABLE	Legacy Label	Migration. Student. DISABILITY	Migration Label	056 Label (where different to Migration)	1:1 code mapping
08	Two or more impairments and/or disabling medical conditions	96	A disability, impairment or medical condition that is not listed or two or more impairments and/or disabling medical conditions	An impairment, health condition or learning differences not listed above	N
51	A specific learning difficulty such as dyslexia, dyspraxia or AD(H)D	51	Learning difference such as dyslexia, dyspraxia or AD(H)D		Y
53	A social/communication impairment such as Asperger's syndrome/other autistic spectrum disorder	53	Social/communication conditions such as a speech and language impairment or an autistic spectrum condition		Y
54	A long standing illness or health condition such as cancer, HIV, diabetes, chronic heart disease, or epilepsy	54	Long-term illness or health condition such as cancer, HIV, diabetes, chronic heart disease, or epilepsy		Y
55	A mental health condition, such as depression, schizophrenia or anxiety disorder	55	Mental health condition, challenge or disorder, such as depression, schizophrenia or anxiety		Y
56	A physical impairment or mobility issues, such as difficulty using arms or using a wheelchair or crutches	56	Physical impairment (a condition that substantially limits one or more basic physical activities such as walking, climbing stairs, lifting or carrying).		Y
57	Deaf or a serious hearing impairment	57	D/deaf or have a hearing impairment		Y

Legacy Student. DISABLE	Legacy Label	Migration. Student. DISABILITY	Migration Label	056 Label (where different to Migration)	1:1 code mapping
58	Blind or a serious visual impairment uncorrected by glasses	58	Blind or have a visual impairment uncorrected by glasses		Y
		59	Development condition that you have had since childhood which affects motor, cognitive, social and emotional skills, and speech and language (Not applicable to migrated data)	New code	N/A
96	A disability, impairment or medical condition that is not listed above	96	A disability, impairment or medical condition that is not listed or two or more impairments and/or disabling medical conditions	An impairment, health condition or learning differences not listed above	N
		98	Prefer not to say	New code	N/A
NULL	Null value	99	Not available		Y

## Migrated data description – Student, Student.Disability

Entity	Field name	Description
Migration.Provider.	UKPRN	Provider – UKPRN
Migration.Provider.	INSTID	Provider - HESA identifier (key)
Migration.Student.	SID	Student identifier (key)
Migration.Student.	BIRTHDTE	Date of birth
Migration.Student.	CARER	Carer
Migration.Student.	ETHNIC	Ethnicity
Migration.Student.	GENDERID	Gender identity
Migration.Student.	NATION	Nationality
Migration.Student.	OWNSTU	Own student identifier
Migration.Student.	RELIGION	Religion
Migration.Student.	SCN	Scottish Candidate Number
Migration.Student.	SEXID	Sex identifier
Migration.Student.	SEXORT	Sexual orientation
Migration.Student.	SSN	Student support number
Migration.Student.	STUDEP	Dependant (Scotland)
Migration.Student.	TTACCOM	Term time accommodation type
Migration.Student.	TTPCODE	Term time postcode
Migration.Student.	UCASPERID	UCAS personal identifier
Migration.Student.	ULN	Unique learner number
Migration.Student.Disability.	DISABILITY	Disability

**Revision history**

2021-09-27 0.1.0 Mapping for relevant fields updated following the personal characteristics consultation. Remove sentence indicating that disability entity is created only for those students known to have a disability. Remove mapping of NULL data in RELBLF and DISABLE. DISABILITY code 99 not used in migrated data. Updated link to coding manual.

2021-10-11 0.1.1 Added Provider.XINSTCOU01 to depend upon fields.

2021-11-04 0.2.0 Changed mapping of ethnicity code '49' 'Other mixed background' from '112' to '159'. DF field name update: DEPENDANT to STUDEP.

2021-11-08 0.2.1 Added SCN (Scottish Candidate Number) to document (previously in Engagement migration specification). Added TRANS (Transgender) to field mapping.

2021-12-15 0.3.0 Disability code change from 00 to 95 for 'No known impairment, health condition or learning difference'.

2021-12-21 0.4.0 Legacy codes 11 and 12 removed from ETHNICITY mapping as no longer valid; Codes 98 and 99 added as valid codes for SEXID; Mapping for DEPENDANT added.

2022-02-07 0.5.0 Restructure of value mapping tables to remove colour coding, additional content to describe new tables. Some Migration labels altered to make more consistent across fields. References to SR and SAR changed to 051 and 054 respectively.

2022-02-23 0.6.0 National Identity migrated entity removed to reflect removal from the model. Legacy fields removed from migration dataset and links to coding manual updated to 21056, 20051 and 20054. NULL values in DISABLE legacy field mapped to code 99 'Not available' in DISABILITY.

2022-07-19 0.7.0 Introduction text updated to highlight that migrated data supplied in the HESA data platform will on a per entity basis and not combined. Hyperlink and reference to coding manual removed under 'Additional information'. Derived fields removed from migrated data description. NATIOND removed as depend upon field. Nationality mapping updated to map code ZZ to 97 'Not known'.