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

Migration.ModuleInstance

Version 0.3.0: Revision history at foot of document

Overview

This document provides the specification for the migration of legacy data into the Data Futures ModuleInstance entity and select sub-entities; Migration.StudentCourseSession.ModuleInstance.Module and Migration.Engagement.CostCentre

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
Migration.StudentCourseSession.ModuleInstance	Provides information on how data has been migrated into the ModuleInstance sub-entity of StudentCourseSession.
Migration.StudentCourseSession.ModuleInstance.Module	Provides information on how data has been migrated into the Module entity which can be treated as a sub-entity of ModuleInstance, which will exclude any Modules which are not associated with any of the Student course sessions.
Migration.Engagement.CostCentre	Provides information on how data has been migrated into CostCentre for the Engagement.

Migration.StudentCourseSession.ModuleInstance - This sub-entity of StudentCourseSession represents the module information relating to the Student course session. A student may study multiple modules during the Student course session and modules may span multiple Student course sessions. The module instance start and end dates must fall on or between the Student course session start and end dates, but multiple module instances may relate to the same continuing module.

Migration.StudentCourseSession.ModuleInstance.Module - This is treated as a sub-entity of ModuleInstance and provides the module information relating to the module instance. Each module instance will have a single associated module and hence this information has been combined into the ModuleInstance migrated dataset.

Migration.Engagement.CostCentre – Provides FTE for each cost centre associated with the Engagement.

These entities are only created for students within the migrated population (see Z_POPMIGRATION).

This applies to only the Student record (051). Excludes any modules which were previously returned in error (MODOUT = 9).

Depend upon fields

- Provider.UKPRN
- Student.HUSID
- Instance.NUMHUS
- Instance.INSTANCEKEY
- StudentOnModule.APEL
- StudentOnModule.MODSTAT
- StudentOnModule.LANGPCNT
- StudentOnModule.MODCOUNT
- StudentOnModule.MODID
- StudentOnModule.MODOUT
- Module.CRDTPTS
- Module.CRDTSCM
- Module.FTE
- Module.MODLANG
- Module.LEVLPTS
- Module.MTITLE
- ModuleSubject.COSTCN
- ModuleSubject.MODSBJP
- Migration.Z_COLLECTION_YEAR
- Migration.Z_NONSTANDARD
- Migration.StudentCourseSession.SCSSTARTDATE
- Migration.StudentCourseSession.SCSENDDATE

DRAFT

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 coding manual provides full lists of Data Futures valid entries and coverage of the data: https://codingmanual.hesa.ac.uk/dataDictionary/

Field mapping

Data Futures field	Data Futures field	Legacy field(s)	Applicable to	Migrated?	Direct mapping?
MODINSTID	Module instance identifier	STUDENTONMODULEKEY	051	Y	Y
APEL	APEL Module identifier	APEL	051	Y	N
CONTINUING	Continuing module	MODSTAT	051	Y	N
INACTIVEMOD	Module inactive for the entire student course session marker	MODSTAT	051	Y	N
LANGPCNT	Language percentage	LANGPCNT	051	Y	Y
MIFEEAMOUNT	Module instance fee amount			N	
MODCOUNT	Module count	MODCOUNT	051	Y	N
MODID	Module identifier	MODID	051	Y	Y
MODINSTENDDATE	Module instance end date	SCENDDATE	051	Y	N
MODINSTSTARTDATE	Module instance start date	SCSSTARTDATE	051	Y	N
MODULEOUTCOME	Module outcome	MODOUT	051	Y	N
MODULERESULT	Module result			N	
CRDTPTS	Credit value of module	CRDTPTS	051	Y	Y
CRDTSCM	Credit transfer scheme	CRDTSCM	051	Y	N
FTE	Module FTE	FTE	051	Y	Y
LANGUAGEID	Language identifier	MODLANG	051	Y	Y
LEVELPTS	Level of credit points	LEVLPTS	051	Y	N
MTITLE	Module title	MTITLE	051	Y	Y
COSTCN	Cost centre	COSTCN	051	Y	Y

Module instance information has been taken from the collection year from which the Student course session information has been drawn.

The following illustrates how modules (dark green with light green showing their continuation between Student course sessions) and module instances (dark green) may relate to Student course sessions (blue).

A module will satisfy one of the following scenarios, where the academic year refers to the collection year 1 August Y1 to 31 July Y2:

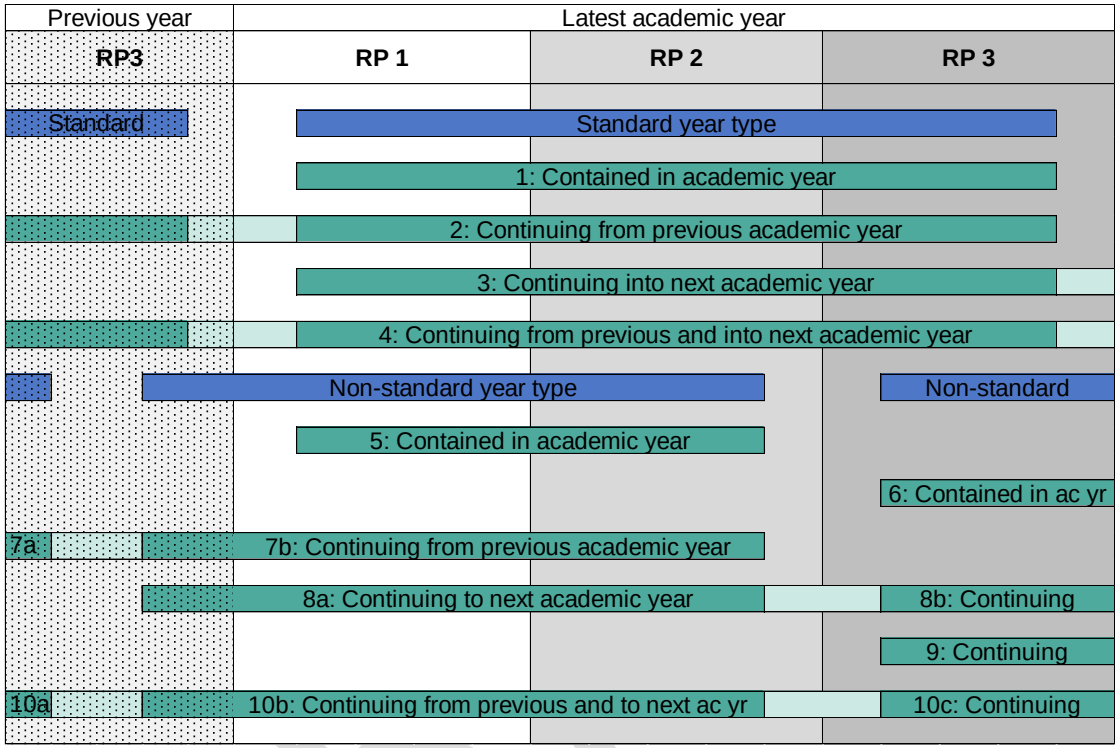
Standard year type

1. Contained within the academic year (MODSTAT = 2).
2. Continuing from previous academic year (MODSTAT = 1).
3. Continuing into next academic year (MODSTAT = 3).
4. Continuing from previous and into next academic year (MODSTAT = 6).

Non-standard year type

5. Student course session starts before 1 August Y1 and module contained within the academic year (MODSTAT = 2).
6. Student course session hasn't ended before 31 July Y2 and module is contained within the academic year (MODSTAT = 2).
7. Student course session starts before 1 August Y1 and module continuing from previous academic year (MODSTAT = 1). Where data is taken from the latest collection year, it is likely that the module is continuing from the previous Student course session if there's a record in the previous collection year for the same module which has MODSTAT = 1 or 6.
8. Student course session starts on or after 1 August Y1 and module continuing to next academic year (MODSTAT = 3). Second Student course session starts on or after 1 August Y1 and module continuing to next academic year (MODSTAT = 3).
9. Student course session starts on or after 1 August Y1 and module continuing to next academic year (MODSTAT = 3).
10. Student course session starts before 1 August Y1 and module continuing from previous and to next academic year (MODSTAT = 6). Second Student course session starts on or after 1 August Y1 and module continuing from previous and to next academic year (MODSTAT = 6). Where data is taken from the latest collection year, It is likely that the module is continuing from the previous Student course session if there's a record in the previous collection for the same module which has MODSTAT = 1 or 6.

Scenarios 2, 4, 8b and 10c represent modules continuing from the previous Student course session. Scenario 7b and 10b represent modules continuing from the previous academic year and may also continue from the previous Student course session. If the data is drawn from the previous collection year with MODSTAT = 1 or 6, we can deduce that scenarios 7b and 10b represent modules continuing from the previous Student course session. Similarly if there is a record in the previous collection year with MODSTAT = 1 or 6 with the same module identifier as one drawn from the latest collection year also with MODSTAT = 1 or 6, we can again deduce that the module continued from a previous Student course session.



The module instance start and end dates align with the Student course session start and end dates respectively, except for the following scenarios. Note that for scenarios 8b, 9 and 10c, where the Student course session end date is NULL, so is MODINSTENDDATE:

- Scenario 5 - start of the academic year (1 August Y1) is used as a proxy for the start of the module instance as we know the module is contained within the academic year.
- Scenario 6 - end of the academic year (31 July Y2) is used as a proxy for the end of the module instance as we know the module is contained within the academic year.

The module start and end dates align with the module instance start and end dates respectively, except for the following scenarios:

- Scenario 2 and 4 – start of the previous Student course session or Engagement start date (whichever is latest) is used as a proxy for the start of the module as we know it is continuing from the previous academic year.
- Scenario 3 and 4 – end of the next session year (Z_PERIODENDDATE) is used as a proxy for the end of the module as we know it's continuing into the following academic year.
- Scenarios 7b and 10b - start of the previous Student course session or Engagement start date (whichever is latest) is used as a proxy for the start of the module only where there is a record in the previous collection year which is continuing from the previous academic year.
- Scenarios 8b and 10c - start of the previous Student course session or Engagement start date (whichever is latest) is used as a proxy for the start of the module. For 10c, this will be updated to a further year back where 10b is updated (as above)
- Scenarios 7b and 10b - start of the previous Student course session or Engagement start date (whichever is latest) is used as a proxy for the start of the module only where there is a record in the previous collection year which is continuing from the previous academic year.
- Scenarios 8b, 9, 10c - end of the session year (Z_PERIODENDDATE) is used as a proxy for the end of the module as we know it's continuing into the following academic year.

DRAFT

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.

For completeness, new codes for Data Futures (056) that have no legacy equivalent are still listed in the tables but marked with ‘N/A’ for legacy code and the 1:1 code mapping. Equally, 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.DEPENDANT. 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’.

APEL credits

StudentOnModule.APEL	Legacy.Label	Migration.SCS.ModuleInstance.APEL	Migration Label	056 Label (where different to Migration)	1:1 code mapping
1	Module taken through APEL	01	Module taken through APEL		Y
2	APEL module	02	APEL module		Y
3	Module not taken/available through APEL	03	Module not taken/available through APEL		Y
Otherwise		NULL	Null value		N

Continuing module

Legacy.MODSTAT (module status) for reference

Legacy.MODSTAT	Legacy.Label
1	Continuing from previous reporting year
2	Contained within reporting year
3	Continuing into next reporting year
4	Taken in a previous reporting year (outcome only)
5	Not coded
6	Continuing from previous reporting year and into next reporting year

Migration. Z_NON STANDARD	Migration. COLLECTION_ YEAR	Migration. SCSSTARTDATE	Legacy StudentOnModule. MODSTAT	Legacy From previous collection: StudentOnModule. MODSTAT	Migration.SCS. ModuleInstance. CONTINUING	Migration.Label
0	2019/20		1, 6		01	Continuing module
1	2019/20	>= 2019-08-01	1, 6			
		< 2019-08-01		3,6		
	2018/19	< 2019-08-01		N/A		
Otherwise					NULL	Null value

Inactive module

Legacy StudentOnModule .MODSTAT	Legacy Label	Migration.SCS.ModuleInstance. INACTIVEMOD	Migration Label	056 Label (where different to Migration)	1:1 code mapping
4	Taken in a previous reporting year (outcome only)	01	Inactive module instance		Y
Otherwise		NULL			N

Module count

Legacy StudentOnModule. MODCOUNT	Legacy.Label	Migration.SCS.ModuleInstance. MODCOUNT	Migration Label	056 Label (where different to Migration)	1:1 code mapping
1	Module not countable	01	Module not countable		Y
2	Module is countable	02	Module is countable		Y

Module Outcome

Legacy.MODOUT (module outcome) codes for reference

Legacy.MODOUT	Legacy.Label
1	Completion - gained full credit
2	Completion - did not gain credit
3	Partial completion (HEFCW HESES Rules)
4	Student did not complete module
5	Module taken on a not-for-credit basis
6	Module outcome not yet known
7	Not coded
9	Module previously returned in error
A	Student did not complete module - gained credit
B	Student did not complete module - deferral
C	Completion - award of credit not known

Module Outcome cont.

Legacy StudentOnModule. MODOUT	Migration. Z_NON STANDARD	Legacy StudentOnModule. MODSTAT	Migration.SCS. ModuleInstance MODULEOUTCOME	Migration Label	056 Label (where different to Migration)	1:1 code mapping
1, 2, C			01	Completion **(includes gained full credit, did not gain credit and award of credit not known)**	Completion	N
3			02	Partial completion (HEFCW HESES Rules)		Y
4, A, B			03	Student did not complete module **(includes did not complete and gained credit)**	Student did not complete module	N
5			04	Module taken on a not-for-credit basis		Y
6	0		06	Will continue Module in a subsequent Student course session		N
6, 7	1	3, 6	06	Will continue Module in a subsequent Student course session		N
6			05	Module result not yet known		N
7			96	Not coded		N
9			N/A	Not mapped		N/A

Module instance start and end dates

Legacy Module Instance. MODSTAT	Migration. COLLECTION_YEAR	Migration. Z_NONSTANDARD	Migration.SCS. SCSSTARTDATE	Migration.SCS. ModuleInstance. MODINSTSTARTDATE	Migration.SCS. ModuleInstance. MODINSTENDDATE
2	2018/19	1		SCSSTARTDATE	2019-07-31
2	2019/20	1	< 2019-08-01	2019-08-01	SCSENDDATE
2	2019/20	1	>= 2019-08-01	SCSSTARTDATE	2020-07-31
Otherwise				SCSSTARTDATE	SCSENDDATE

Credit transfer scheme

Legacy StudentOnModule. CRDTSCM	Legacy.Label	Migration.SCS. ModuleInstance. Module.CRDTSCM	Migration Label	056 Label (where different to Migration)	1:1 code mapping
1	England & Northern Ireland (CATS)	01	A UK Credit Accumulation and Transfer Scheme (CATS)		N
2	Scotland (SCOTCAT)	01	A UK Credit Accumulation and Transfer Scheme (CATS)		N
3	CQFW credit framework	01	A UK Credit Accumulation and Transfer Scheme (CATS)		N
6	LSDA credit framework	01	A UK Credit Accumulation and Transfer Scheme (CATS)		N
N/A		02	European Credit Transfer and Accumulation System (ECTS)	New code	N/A
7	Other scheme	90	Other Scheme		Y
8	Own scheme	03	Provider's own scheme		Y
9	No scheme	91	No Scheme		Y
A	GNVQ units/OCN credits	01	A UK Credit Accumulation and Transfer Scheme (CATS)		N

Language identifier

Legacy StudentOnModule. **MODLANG**	Legacy.**Label**	Migration.SCS.ModuleInstance. Module.**LANGUAGEID**	**Migration Label**	**056 Label (where different to Migration)**	**1:1 code mapping**
1	Welsh	01	Welsh		Y
2	Gaidhlig	02	Gaidhlig		Y
3	Irish	03	Irish		Y
9	Not available through the medium of Welsh, Gaidhlig or Irish	NULL	Null value		Y

Level of credit points

Legacy StudentOnModule. LEVLPTS	Legacy.Label	Migration.SCS. ModuleInstance. Module.LEVLPTS	Migration Label	056 Label (where different to Migration)	1:1 code mapping
1	HE Certificate/NVQ Level 4 or equivalent	20	HE Certificate/NVQ Level 4 or equivalent		Y
2	HE Intermediate	21	HE Intermediate/HE Diploma/Foundation degree		N
3	HE Honours	23	HE Honours		Y
5	Undergraduate unspecified	24	Undergraduate Unspecified		Y
6	HE Masters	25	HE Masters		Y
7	HE Doctorate	26	HE Doctorate		Y
9	Not applicable	96	No associated level		Y
A	NVQ level 1 or equivalent	27	Below HE Level		N
B	NVQ level 2 or equivalent	27	Below HE Level		N
C	NVQ level 3 or equivalent	27	Below HE Level		N
D	HND/Diploma HE	21	HE Intermediate/HE Diploma/Foundation degree		N
E	Ordinary degrees	22	Ordinary degrees		Y

Module start and end dates

Migration.ModuleInstance. CONTINUING	Migration.ModuleInstance. MODULEOUTCOME	Migration. Z_MODSTARTDATE	Migration. Z_MODENDDATE
01		GREATEST(MODINSTSTARTDATE – 1 year, ENGSTARTDATE)	MODINSTENDDATE
	06	MODINSTSTARTDATE	MODINSTENDDATE + 1 Year
01	06	GREATEST(MODINSTSTARTDATE – 1 year, ENGSTARTDATE)	MODINSTENDDATE + 1 Year
Otherwise		MODINSTSTARTDATE	MODINSTENDDATE

In the migrated data for the Student record, each Module associated to an Engagement can have at most two Module instances associated to it (corresponding to at most two Student course sessions). Therefore, there will be one value of Z_MODACTFROM and Z_MODACTTO for each Module. These define the start and end dates of the active parts of the module.

Z_MODACTFROM is the minimum value of Z_MODSTARTDATE across all Module instances with the same module identifier (MODID) and are active (INACTIVEMOD is NULL). Z_MODACTFROM is the maximum value of Z_MODENDDATE across all Module instances with the same module identifier (MODID) and are active (INACTIVEMOD is NULL). If Z_MODACTFROM is on or before 31 July 2019, set as this date to indicate the module started in the previous collection year. If Z_MODACTTO is after 31 July 2020, set to the dummy date of 31 July 9999 to indicate it is continuing into the next collection year.

Migrated data description – StudentCourseSession.ModuleInstance, StudentCourseSession.ModuleInstance.Module

Data Futures derived fields prefixed by Z_: https://codingmanual.hesa.ac.uk/21056/
Student record legacy data fields: https://www.hesa.ac.uk/collection/c20051/index
Student record legacy derived fields: https://www.hesa.ac.uk/collection/c20051/derived/contents
Student Alternative record legacy data fields: https://www.hesa.ac.uk/collection/c20054/index
Student Alternative record legacy derived fields: https://www.hesa.ac.uk/collection/c20054/derived/contents

Entity	Field name	Description
Migration.	Z_COLLECTION_YEAR	Data from 2019/20 (latest) or 2018/19 (previous) collection year
Migration.Provider.	UKPRN	Provider – UKPRN
Migration.Provider.	INSTID	Provider - HESA identifier (key)
Migration.Student.	SID	Student identifier (key)
Migration.Engagement.	NUMHUS	Engagement number (key)
Migration.StudentCourseSession.	SCSESSIONID	Student course session identifier (key)
Migration.StudentCourseSession.	COURSEID	Course identifier
Migration.StudentCourseSession.ModuleInstance	MODINSTID	Module instance identifier (key)
Migration.StudentCourseSession.ModuleInstance	APEL	APEL Module identifier
Migration.StudentCourseSession.ModuleInstance	CONTINUING	Continuing module
Migration.StudentCourseSession.ModuleInstance	INACTIVEMOD	Module inactive for the entire student course session marker
Migration.StudentCourseSession.ModuleInstance	LANGPCNT	Language percentage
Migration.StudentCourseSession.ModuleInstance	MODCOUNT	Module instance fee amount
Migration.StudentCourseSession.ModuleInstance	MODID	Module count

Entity	Field name	Description
Migration.StudentCourseSession.ModuleInstance	MODINSTSTARTDATE	Module identifier
Migration.StudentCourseSession.ModuleInstance	MODINSTENDDATE	Module instance end date
Migration.StudentCourseSession.ModuleInstance	MODULEOUTCOME	Module instance start date
Migration.StudentCourseSession.ModuleInstance.Module	CRDTPTS	Credit value of module
Migration.StudentCourseSession.ModuleInstance.Module	CRDTSCM	Credit transfer scheme
Migration.StudentCourseSession.ModuleInstance.Module	FTE	Module FTE
Migration.StudentCourseSession.ModuleInstance.Module	LANGUAGEID	Language identifier
Migration.StudentCourseSession.ModuleInstance.Module	LEVLPTS	Level of credit points
Migration.StudentCourseSession.ModuleInstance.Module	MTITLE	Module title
Migration.Derived.	Z_NONSTANDARD	Non-standard year type (Student record)
Migration.Derived.	Z_MODSTARTDATE	Module start date
Migration.Derived.	Z_MODENDDATE	Module end date
Migration.Derived.	Z_MODACFROM	Module active from
Migration.Derived.	Z_MODACTTO	Module active to

Migrated data description – Engagement.CostCentre

Entity	Field name	Description
Migration.Provider.	UKPRN	Provider – UKPRN
Migration.Provider.	INSTID	Provider - HESA identifier (key)
Migration.Student.	SID	Student identifier (key)
Migration.Engagement.	NUMHUS	Engagement number (key)
Migration.Engagement.	COSTCN	HESA cost centre (key)
Migration.Engagement.	Z_CCFTE_CYC	FTE - Cost centre FTE for twelve-month cycle (England, Wales, Northern Ireland)
Migration.Engagement.	Z_CCFTESCS_CYC	FTE - Cost centre FTE for all student course sessions in the cycle with an end date (Scotland)

Revision history

2021-10-06 0.0.1 Updated link to coding manual.

2021-12-21 0.1.0 Added code 03 ‘Provider’s own scheme’ to CRDTSCM and code 96 ‘No associated level’ to LEVLPTS.

2022-01-24 0.2.0 MODULEOUTCOME code 99 updated to 96.

2022-02-07 0.2.1 LEVLPTS code 99 updated to 96. 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 0.3.0 Legacy fields removed from migration dataset and links to coding manual updated to 21056, 20051 and 20054.

DRAFT