

HESA Data Futures

Project Overview for PQQ

2016-04-12

Plan description

The purpose of this paper is to provide an overview of the Data Futures programme and of HESA to help potential bidders respond to HESA's Data Future's OJEU notice and Pre-Qualifying Questionnaire.

Contents

| | |
|---|----|
| 1. Introduction..... | 3 |
| 1.1 HESA..... | 3 |
| 1.2 Data Futures programme | 4 |
| 1.3 HE providers | 4 |
| 1.4 Data customers..... | 4 |
| 2. Current situation..... | 6 |
| 2.1 Collection streams | 6 |
| 2.2 Current metrics..... | 8 |
| 2.3 Quality assurance | 9 |
| 3. The need for change..... | 11 |
| 3.1 Introduction..... | 11 |
| 3.2 New HE data landscape vision..... | 11 |
| 3.3 Data Futures strategic drivers | 13 |
| 4. Data Futures | 15 |
| 4.1 Objectives | 15 |
| 4.2 Vision | 15 |
| 4.3 Indicative programme schedule | 16 |
| 4.4 Detailed design activities and deliverables | 17 |
| 5. Business Processes | 26 |
| 5.1 Collection strategy..... | 26 |
| 5.2 Process maps | 29 |
| 6. Technical building blocks..... | 33 |
| 7. Delivery capabilities..... | 37 |
| 8. Business as usual | 40 |
| APPENDIX A – Data item definition and quality rules | 41 |
| APPENDIX B – Data principles..... | 43 |
| APPENDIX 3 – Views of the technical components | 44 |

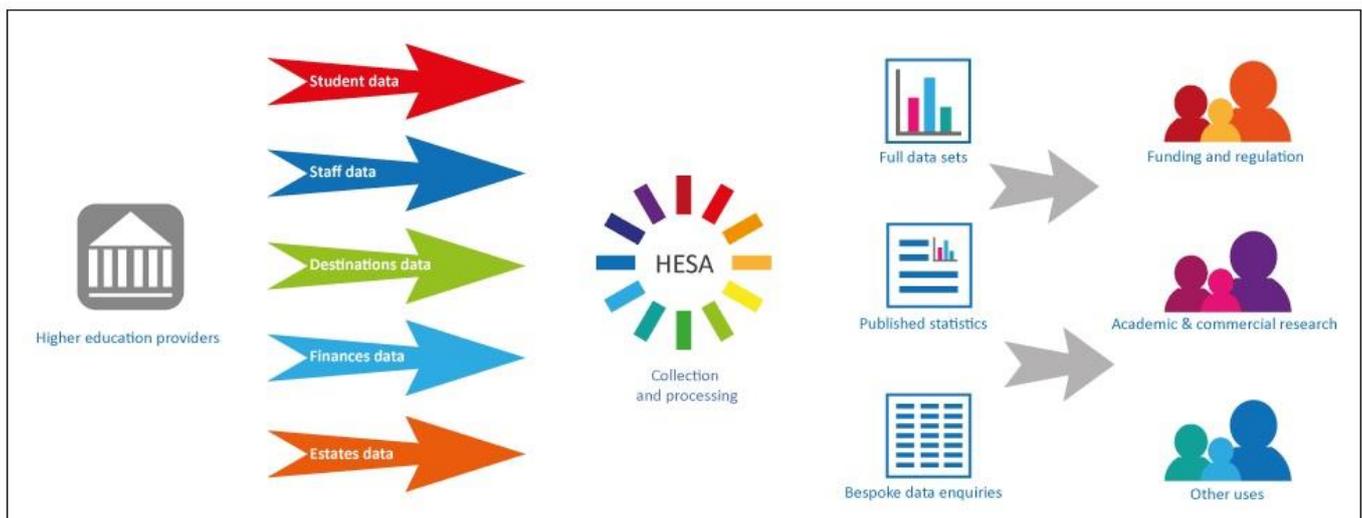
1. Introduction

1.1 HESA

HESA collects a range of data every year UK-wide from universities, higher education colleges and other differently funded providers of higher education. This data is then provided to UK governments and higher education funding bodies to support their work in regulating and funding higher education providers. In addition information derived from the data is published as official statistics and in many accessible formats for use by a wide range of organisations and individuals for a variety of purposes, including HE providers, academic researchers, students, prospective students, private companies, professional bodies and the press and media.

HESA is a charitable company which is funded by the subscriptions of the HE providers from whom it collects data and it also provides an information service for the procurement of expertise, analysis and bespoke information.

Figure 1 HESA overview



For more information on HESA please refer to the website www.hesa.ac.uk

The content of the data that HESA collects is determined by government and HE funding bodies, in consultation with HE providers and other key data users.

The data collection streams are:

- Student record - information about students, courses and qualifications at HEIs
- Alternative Provider (AP) Student record - information about students, courses and qualifications at alternative providers.
- Initial Teacher Training (ITT) in-year record - administrative data about those undertaking teacher training
- Staff record - information about staff employed by HEIs
- Finance record - income and expenditure of HEIs
- Destinations of Leavers from Higher Education - survey of graduate activities six months after leaving HE
- DLHE Longitudinal Survey - sample survey of leavers 3.5 years after graduation
- Aggregate offshore record - count of students studying wholly overseas for UK HE qualifications
- HE Business and Community Interaction survey - information about interactions between HEIs and business and the wider community
- Estates management record - buildings, estates and environmental information about HEIs
- Institution profile record - Information about HEI campuses and departments
- Key Information Set - Data about undergraduate courses, published on Unistats

Further information on these collection streams including manuals, specifications, data models, data items etc can be found on the HESA website (www.hesa.ac.uk) under the 'Data Collection' tab.

1.2 Data Futures programme

Introduction

HESA is embarking on a programme of change that will help to deliver a new higher education information landscape.

The UK's collection and use of HE data is valued by the sector and respected around the world. To achieve this, HESA has worked with HE providers over the last 22 years, but the context in which HESA operates is rapidly becoming more complex and the system used to collect data is no longer fit for purpose.

HESA's vision is for a modernised and more efficient approach to collecting data that is able to respond to a wide range of data users. HESA wants to create more relevant, reliable, comprehensive and timely information about higher education for the benefit of the sector, and to deliver this in more efficient ways which reduce the burden on providers.

For more information on the programme please refer to the website www.hesa.ac.uk/datafutures

Scope

The scope of the procurement includes the following collection streams:

- Student record
- Alternative Provider (AP) Student record
- Initial Teacher Training (ITT) in-year record

1.3 HE providers

Publicly funded higher education providers are required to subscribe to HESA as a full subscriber in accordance with the requirements of their respective public funding bodies and submit records to the Student record collection.

Alternative providers with courses designated by Department for Business Innovation and Skills (BIS) are required to submit records to the AP student record collection.

Higher education providers (HEPs) in England which provide courses (Initial Teacher Training (ITT)) leading to Qualified Teacher Status, and Early Years ITT (EYITT) courses leading to Early Years Teacher Status (EYTS) are required to submit records to the ITT in-year record collection.

1.4 Data customers

Customers that HESA disseminates student data to include:

- Department for Business, Innovation and Skills
- Welsh Government
- Scottish Government
- Department for Employment and Learning, Northern Ireland
- Higher Education Funding Council for England

- Higher Education Funding Council for Wales
- Scottish Further and Higher Education Funding Council
- Department for Education
- Research Councils
- Skills Funding Agency
- National College for Teaching and Leadership
- National Health Service (including Health Education England)
- General Medical Council
- Office For Fair Access
- Quality Assurance Agency for Higher Education

2. Current situation

2.1 Collection streams

2.1.1 Student record

The HESA Student record has been collected since 1994/95 from subscribing Higher Education Providers (HEPs) throughout the devolved administrations of the United Kingdom. The data collected as part of the Student record is used extensively by various stakeholders and is fundamental in the formulation of:

- Funding
- Performance Indicators
- Publications (including UNISTATS)
- League tables

The Student record is collected in respect of all students registered at the reporting provider who follow courses that lead to the award of a qualification or provider credit. The emphasis of the coverage of the Student record is those who are (or were) actively following a course at some time during the HESA reporting period. Full details of the coverage of this record can be found in the coverage statement of the coding manual.

The Student record collects individualised data about students active during the reporting period. A wide range of data items are collected, including: the student's entry profile and personal characteristics, module and course level data, funding information and qualifications awarded. The detail of the different items collected can be found in the Data Items document located on the [coding manual landing page](#).

There are differing requirements for different types of student, and for some students it is possible to make a reduced return, submitting only a subset of the fields of the Student record.

Students who are studying overseas or who come to the UK for a period of less than 8 consecutive weeks during their programme of study are not included in the Student record but should instead be included in the Aggregate offshore record. This is a complementary annual return which provides a headcount of students studying through the provider's provision overseas.

All HESA records are collected on the basis of the HESA reporting period that determines the time period that the data being returned relates to. This ensures consistency across the data streams collected. The reporting period is from 01 August year 1 to 31 July year 2, for example, the 2000/2001 Student record was collected in respect of the activity which took place between 01 August 2000 and 31 July 2001.

2.1.2 Alternative Provider (AP) Student record

In general, the Alternative provider (AP) student record is collected in respect of all students registered at the reporting provider who are studying on undergraduate courses, with a small number of exceptions.

All HESA records are collected on the basis of the HESA reporting period that determines the time period that the data being returned relates to. This ensures consistency across the data collected. The reporting period is from 01 August year 1 to 31 July year 2, for example, the 2015/16 AP student record is collected in respect of the activity taking place between 01 August 2015 and 31 July 2016.

The AP student record is collected to support two functions.

1. Course designation

The primary function of the record is to support course designation and the data returned must include all students on designated courses, regardless of course, mode and level, excluding postgraduate courses only designated for Disabled Students Allowance (DSA).

The AP student record is collected by HESA on behalf of the Department for Business Innovation and Skills (BIS) and the Higher Education Funding Council (HEFCE). Alternative providers with courses designated by BIS are required to submit the AP student record to HESA.

2. Unistats

The second function of the AP student record is to collect data for publication on the Unistats website to aid prospective higher education students in making a decision about what and where to study. From 2015/16, the reporting of data for inclusion in Unistats is compulsory.

Providers must submit data for students on designated and non-designated courses in order for the data to be published on the Unistats website. The data requirements for students on non-designated courses are different to those of students on designated courses and this is explicit in the coverage statement of each data item.

To fulfil the requirement to return data for publication on Unistats providers will additionally need to complete the Destination of Leavers from Higher Education (DLHE) and Key Information Set (KIS) records. Details of these returns can be found on the HESA website. In addition these providers will be included in the National Student Survey (NSS).

The AP student record is intrinsic to both the DLHE and KIS returns and is also used to identify the National Student Survey population

The detail of the different items collected can be found in the Data Items document located on the [coding manual landing page](#).

2.1.3 Initial Teacher Training (ITT) in-year record

The HESA Initial Teacher Training In-Year record has been collected from 2008/09 onwards from constituent higher education providers (HEPs) in England which provide courses (Initial Teacher Training (ITT)) leading to Qualified Teacher Status, and Early Years ITT (EYITT) courses leading to Early Years Teacher Status (EYTS).

This data is collected for census and funding purposes for the National College for Teaching and Leadership (NCTL), and also used by the NCTL in the allocation of Teacher Reference Numbers (TRNs).

The Initial Teacher Training In-Year record collects data about the personal characteristics of new and continuing ITT students and the details of the course(s) that they are enrolled on. A full list of the data items collected can be found in the Data items document on the coding manual page.

Data for the record is collected in-year, therefore reporting throughout the period on activity which is taking place in the current academic year. The HESA Data Collection System opens at the beginning of September, and providers are required to have sent a full and verified set of data to HESA by the deadline for submission in October, although updates can be made to data throughout the year. There are several deadlines which providers must achieve as part of the submission process and these are detailed in the data collection schedule document on the coding manual. It is important that providers work to the set deadlines in order to complete the return on time and ensure data quality.

All HESA records are collected on the basis of the HESA reporting period, which determines the period of time the data being returned relates to. This ensures consistency across the data streams collected. The reporting period is from 01 August in year 1 to 31 July in year 2. For example, the 2015/16 Initial Teacher Training In-Year record is collected in respect of the activity which has taken place between 01 August 2015 and 31 July 2016.

The detail of the different items collected can be found in the Data Items document located on the [coding manual landing page](#)

2.2 Current metrics

This section is intended to provide a brief overview of current indicative volumes. Approximations are shown below because the record is subject to ongoing review and development year on year.

2.2.1 Student Record

HESA collects data on approximately 3 million students each year (more details can be found at www.hesa.ac.uk/stats). Currently, on average HE providers submit their data 20 times before it is accepted therefore HESA processes approx. 60 million student records each year in the collection process.

Approximately 165 HE providers submit data to the Student Record collection.

The student collection includes approx. 214 data items which are structured into 19 groups, many of these groups repeat for example the group 'course subject' can repeat up to 3 times for each 'course'.

Approx. 130 data items have coding frames. A coding frame is a list of codes with associated rules, definitions and descriptions. Some coding frames are used by more than one data item for example the coding frame for 'subject' is used for 'subject of course', 'subject of module', 'PGCE subject of undergraduate degree' etc.

(www.hesa.ac.uk/index.php?option=com_studrec&task=show_file&mnf=15051&href=A^ ^SBJCA.html)

Most coding frames are developed and maintained by HESA however some are sourced externally for example 'occupation code' is based on SOC2000 from the Office for National Statistics.

An example of a data item definition ('term time postcode') is shown in the Appendix

2.2.2 Alternative provider collection

The AP Student Record is collected annually for approx. 60,000 students. Approximately 100 HE providers submit data to the AP collection.

The AP collection includes approx. 60 data items which are structured into 10 groups, many of these groups repeat for example the group 'course subject' can repeat up to 3 times for each 'course'.

Some data items are common to both AP and the Student Collection for example 'postcode' appears in both. However 'term time postcode' which only appears in the student collection has a different definition and quality rules compared to 'post code' because it is used in a specific context.

2.2.3 Initial Teacher Training (ITT)

Data was collected on approx. 45,000 students as part of the ITT collection.

Approximately 70 HE providers submit data to the ITT collection.

The ITT collection includes approx. 49 data items which are structured into 3 groups, groups can repeat for example the group 'course subject' can repeat up to 3 times for each 'course'.

2.3 Quality assurance

Quality assurance of the student data collected by HESA is achieved through a number of processes given below.

XML Schema Validation

- The student data are submitted as XML data. These are validated against published XSD. This validation is achieved using HESA's [Validation Kit](#). The validation kit is a free Windows GUI application available publicly for download. It is used by higher education providers to check their data locally before submission to HESA. The kit detects the type of data being validated and downloads the appropriate XSD to validate it. Schema validation takes from a few seconds to tens of minutes depending on size of provider data.
- XML Schema validation is repeated by HESA when any data is submitted. The same reports generated by the Windows application are made available within the HESA data collection portal.

Once data has passed schema rules that ensure data meets the fundamental structural rules, providers then move on to the next stage of validation, a suite of quality rules that ensure the data meets the current record specification, with correct coverage and is not duplicated. A subset of rules focusses on checking consistency of students between years, reporting potential anomalies where, for example, a student may have been returned to HESA in the previous year but appears to be missing from the current year.

XML Business Rules

- Following schema validation approximately 800 business rules are run against the data. For examples please see those rules with Stage = Business Rules at <https://www.hesa.ac.uk/QualityRules/Index.html?col=C15051>.
- This process also takes place within the HESA Validation Kit and is again repeated when any data is submitted. The same reports generated by the Windows application are made available within the HESA data collection portal.

Exception Rules

- Following XML QA processes the data is loaded to a database and a further set of business rules (around 200) are applied. For examples please see those rules with Stage = Exception at the same link above.
- This process is not part of the HESA Validation Kit because some of these rules rely on previous years' student data or large lookup data that for privacy or technical reasons HESA has not found a way to incorporate into the Validation Kit.
- These reports are made available within the HESA data collection portal.

Continuity Rules

- These are a distinct set of exception rules (around 70) that focus on checking the continuity of student records from year to year. E.g. a record for a student who was in their 1st year last year is expected this year.

Credibility Reporting

- This is a suite of reports that tabulate different dimensions of submitted data and make comparisons to the previous year's data. The goal is to check that changes to scale and distribution of data are proportionate and consistent. The reports are accompanied by a set of rules that automatically identify such potential issues.
- Any data quality issues identified through credibility checking are raised with the HEP who must respond to each issue raised, to either confirm that apparent anomalies are genuine, giving an appropriate explanation, or to correct and re-submit the data. The explanations provided by the HEP serve as valuable intelligence in subsequent years as they are used for reference to refine the questions asked.
- There is significant scope and expectation that future QA automates much more of this activity which is currently dependent on HESA analysts checking these reports and raising issues manually.

3. The need for change

3.1 Introduction

The Department for Business, Innovation and Skills (BIS) White Paper, 'Students at the Heart of the System', called for specific improvements to the collection of higher education data. These improvements include ensuring data meets the needs of wider groups of users, reducing duplication in data collection and driving for timelier and more relevant data. As the official agency responsible for data collection, HESA has a key role to play in ensuring the HE sector can achieve these objectives.

As part of the sector's response to this White Paper, a report was commissioned by the Higher Education Data and Information Improvement Programme (HEDIIP) to assess the HE data landscape. The resulting HEDIIP New Data Landscape report recommended that HESA takes on more responsibility for data collection and data governance across the sector with a view to reducing the number of data demands that are made directly of Higher Education Providers (HEPs).

The HE sector as a whole recognises that a data collection approach is required, that results in relevant, reliable, and comprehensive data - one that makes timely HE information available to those who need it, to ensure critical HE policy and planning decisions are appropriately informed and the burden on HEPs to provide this data is as low as possible, to maximise the efficiency of the data return process.

Building on the above government requirement, HESA developed a Vision 2020, which is:

"to have in place, from the perspective of all UK stakeholders, a modern and efficient approach to collecting UK wide data that is able to respond to the multiple needs of professional, statutory and regulatory bodies, higher education providers and users who wish to understand and engage with higher education. The reformed collection will result in relevant, reliable, comprehensive and more specifically, timely in-year information."

In order to achieve this vision and respond to the HEDIIP New Data Landscape recommendation, HESA, together with the HE sector, recognise that the way in which data is collected must fundamentally change. Data Futures is the programme through which this change will be brought about.

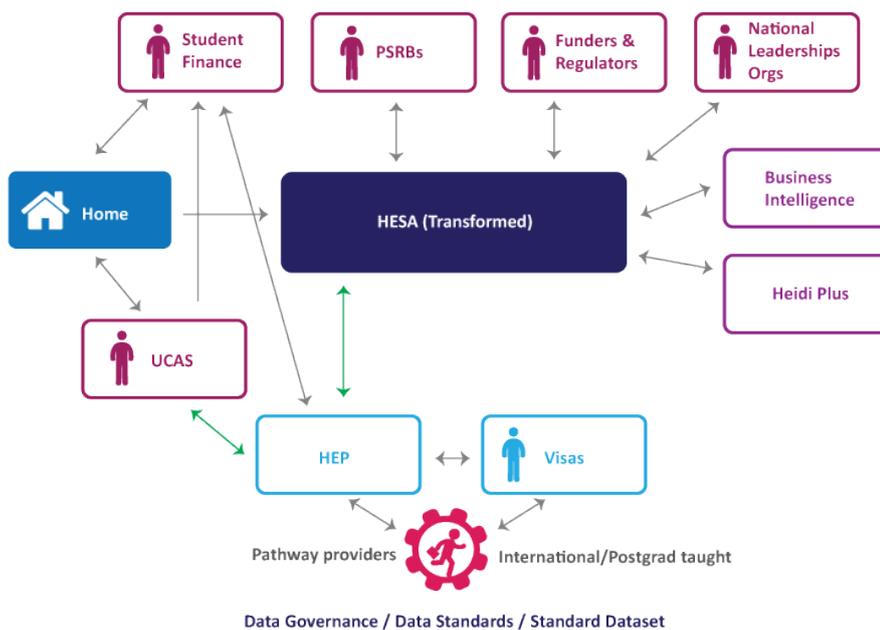
3.2 New HE data landscape vision

HEDIIP identified a wide range of issues with the current administration and collection of data across the sector:

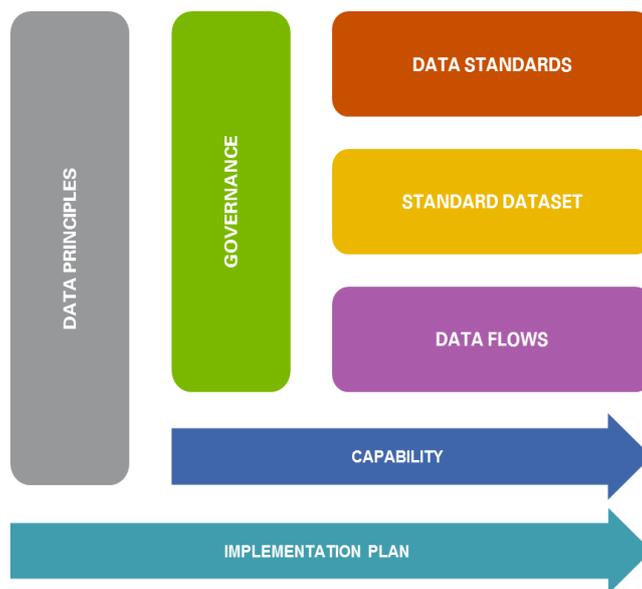
- There are 97 different Data Collectors and over 520 different data returns that HEPs may be required to submit;
- There is no requirement for the data collectors to collaborate and coordinate their data needs from HEPs;
- There is duplication in the data requested by the different data collectors, creating unnecessary burdens for HEPs;
- The plethora of data requests has led to multiple data definitions for the same or similar data field that creates complexity for HEPs without any real gain;
- HEPs have varying levels of sophistication and capability in their data processes; and
- The need for certain data collections and the use of the data is not commonly understood by HEPs.

These matters contribute to a view that the burden of data collection is greater than it should be. HEDIIP identified that principally four things influence the perception of burden: 1) the physical volumes of data collected; 2) the complexity and multiplicity of data definitions; 3) the methods and number of organisations collecting data from HEPs; and 4) the extent to which the data collected is used.

In response to these issues HEDIIP developed a vision for a new HE data landscape; shown in the diagram below (for more information refer to the HEDIIP website www.hediip.ac.uk/transformed-hesa-collection).



This vision is underpinned by the core building blocks described below.



| | |
|------------------------|---|
| Data Principles | Will be used to provide a shared understanding of the ambitions of the landscape and an agreed framework for the data collectors and HEPs across the landscape to adhere to (See Appendix). |
| Governance | <p>Effective governance is central to the development of the landscape. It will provide the administration of the data standards, the inventory of data collections and Data Collectors. It will also provide oversight of the adherence to the Principles.</p> <p>HESA and HEDIIP are working to setup collective governance arrangements in the form of the HE Data Landscape Steering Group and Advisory Panel; these groups will have a role to play in reviewing key deliverables from the Data Futures programme.</p> |

| | |
|----------------------------|---|
| Data Standards | <p>Would be provided by the development and publication of the common data specifications for the data being collected within the HE sector. The publication of the data standards would provide all data collectors with an opportunity to ensure that the data that they require is not already being collected before they embark on additional data collection, thus minimising duplication. It will also enable retrospective data collection to cease.</p> <p>The HEDIIP Data Language project is tasked with producing a logical data model and definition by July 2016.</p> |
| Standard Dataset | This would be the collective name for the data items that would be part of the landscape |
| Data Flows | <p>Describes an optimised approach for the exchange of data between data collectors and HEPs. The development of the data flows in the New Landscape will be strongly influenced by the data collection approaches that are adopted. The potential change of the HESA return to an in-year data collection will enhance the opportunities for rationalising the data flows.</p> <p>Data Futures is the implementation of this building block.</p> |
| Capability | <p>The need to raise data management capability across data collectors and HEPs.</p> <p>HEDIIP and HESA are working on a Data Capability initiative to raise the data capability of the HE sector. Further details of this initiative are shown on the website www.hediip.ac.uk/improved-data-capability. The toolkit to support HE providers and collectors improve their data management capability is available on the website hediip.ac.uk/dc_toolkit.</p> |
| Implementation plan | The implementation plan is described in the next section. |

3.3 Data Futures strategic drivers

The current approach to collection of HE data is increasingly viewed as unsustainable by many in the sector. In particular, the once per year retrospective collection of student data by HESA means that organisations who need access to in-year data have to collect it themselves. This has resulted in unstructured growth of the number of data collections that HEPs must respond to.

Because of the large number of different data collectors that request similar data from HEPs, with slightly different definitions and data sets, HEPs have stated repeatedly that they compile data for each data collection separately, often manually, focusing on responding to each individual return. This conflicts with their need to produce higher quality internal data and analysis, meaning their leadership teams do not have the highest quality data and analysis on which to take decisions.

This leads to variable data quality overall across the sector, and results in there being no 'single version of the truth' for student numbers, progression and achievement, making it harder to take appropriate, timely and strategic decisions based on data that is hard to reconcile or link effectively and share. This is also potentially true for other data returns made to HESA such as finance and staff.

In response to these issues, HEDIIP conducted a review of the HE data landscape and recommended that HESA take on an enhanced role within the HE sector for both data collection and data governance. In order to respond to this recommendation, HESA must prepare for significant change to the way in which data is collected, as well as the role HESA will play in data governance for the sector. Data Futures is a key vehicle through which the HEDIIP recommendations will be delivered, by preparing HESA to take on this enhanced responsibility.

From discussions with HE funding bodies, government departments, HEPs and other relevant stakeholders, it has become clear that the overarching benefit of delivering Data Futures would be better informed strategic decision-making and intelligence for the HE sector, requiring less overall effort in data collection. In the longer term, HEPs have indicated that the reduction in the number of data collections would drive efficiencies in their data collection and reporting activities.

The strategic drivers for the Data Futures programme are:



1. Enable timely data collection, analysis and dissemination

Expectations of data and business intelligence (BI) have changed. Data providers and users now expect and require data faster and HESA must be able to collect, analyse and distribute data more frequently.



2. Reduce the burden of submission on providers

The programme must alleviate the burden of data submissions on data providers.



3. Drive improved value for money and greater return on investment

HESA must be able to deliver more services and drive greater value from its activities in the future. In order to provide greater value for money, HESA must become more efficient at data collection and dissemination to release capacity to drive additional value from the data collected.



4. Deliver data that is fit for purpose

The data collected must satisfy the requirements of data users. HESA must be able to respond to changes in data requirements, to ensure the data collected is 'fit for purpose', whilst minimising the impact that these changes have on data providers.



5. Support changes to the new data landscape (as recommended by the Higher Education Data and Information Improvement Programme)

HESA must be able to adapt to its role in the new data landscape, which will require HESA to collect data on behalf of data users who currently collect their own data directly from HEPs.

4. Data Futures

4.1 Objectives

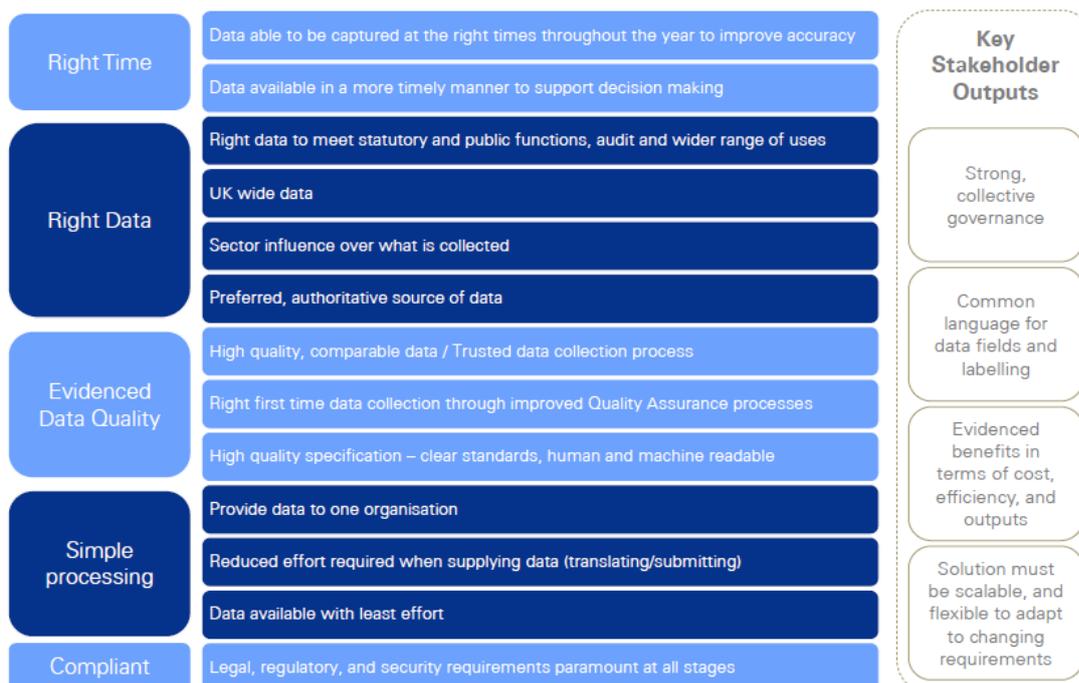
The objectives of the programme:

- To implement an in year HE student data collection and dissemination stream that meets the requirements of HESA’s customers and HE providers.
- To implement a new HE student data model that meets the requirements of HESA’s customers and HE providers; including the migration of historical data into the new model to support time series reporting.
- To transition HESA’s IT student data collection and dissemination systems to a secure cloud based platform.
- To design, build and implement a scalable and flexible student data collection and dissemination stream that HESA can extend to its other collection streams.
- To transition the new HE student data collection and dissemination stream to business as usual.
- To work with HE providers, HESA customers and student information system providers to ensure their readiness for implementation.
- To communicate to all Data Futures stakeholders to ensure they understand the programme, the impact it will have on them and how they can engage with the programme.
- To work with the HE Data Landscape Steering Group and Advisory Panel to ensure that the design of the new HE student data collection and dissemination stream is supported by the stakeholder community.

The desired outcomes:

- Timely data collection, analysis and dissemination of HE data
- Reduced burden of submission on HE providers
- Improved value for money and greater return on investment on HE data collection
- Data that is fit for purpose
- Implementation of the new data landscape (as recommended by the Higher Education Data and Information Improvement Programme)

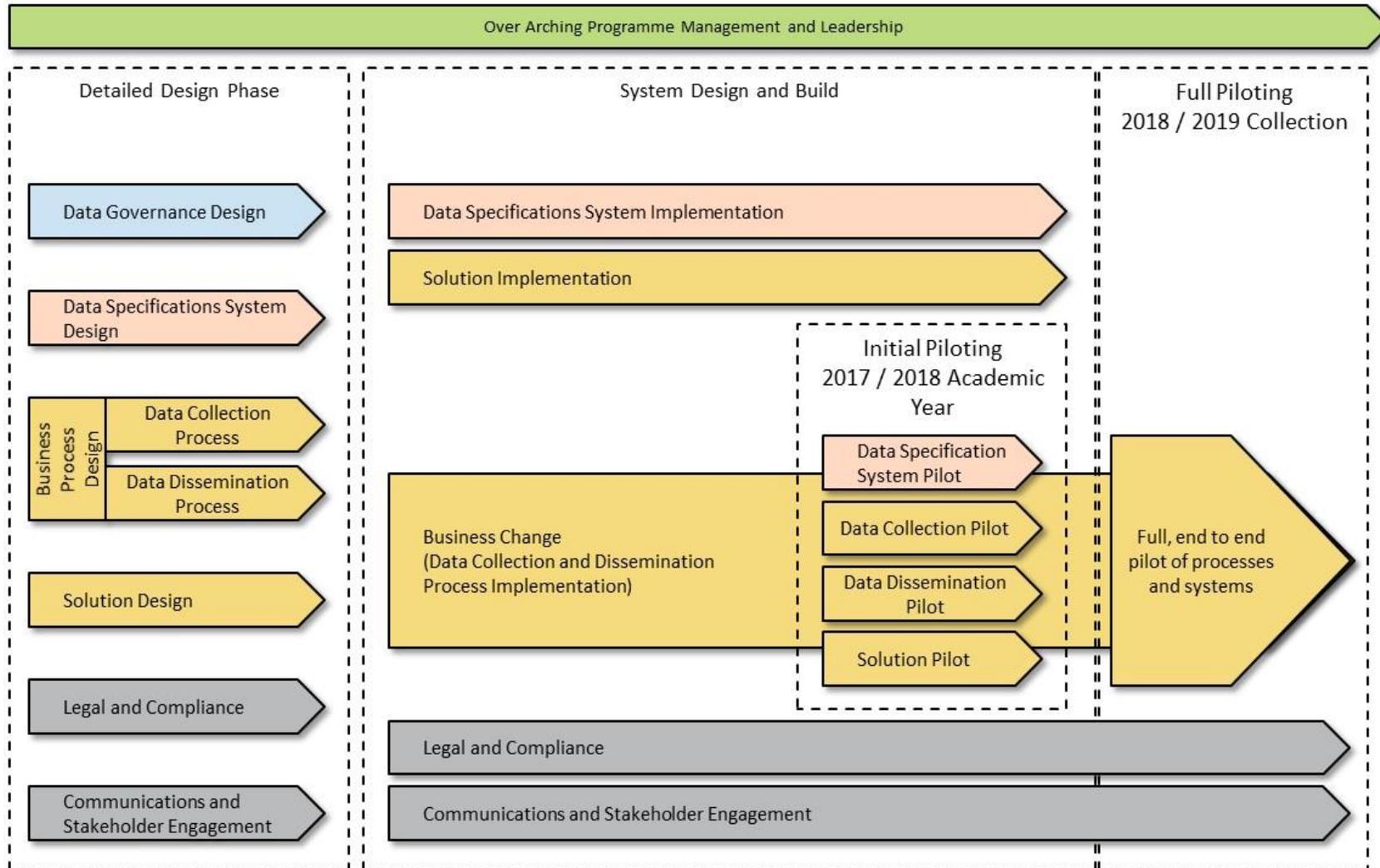
4.2 Vision



4.3 Indicative programme schedule

HESA has identified an outline implementation approach shown below. HESA welcomes supplier’s alternative approaches especially use of agile. Suppliers will need to bear in mind that the specification of the HE provider interface/s must be frozen at least 12 months before the start of the full pilots in the academic year 2018/19 (i.e. by August 2017) to enable HE student information system providers to update, test and deploy any changes required to HE systems.

Figure 2 Indicative programme schedule

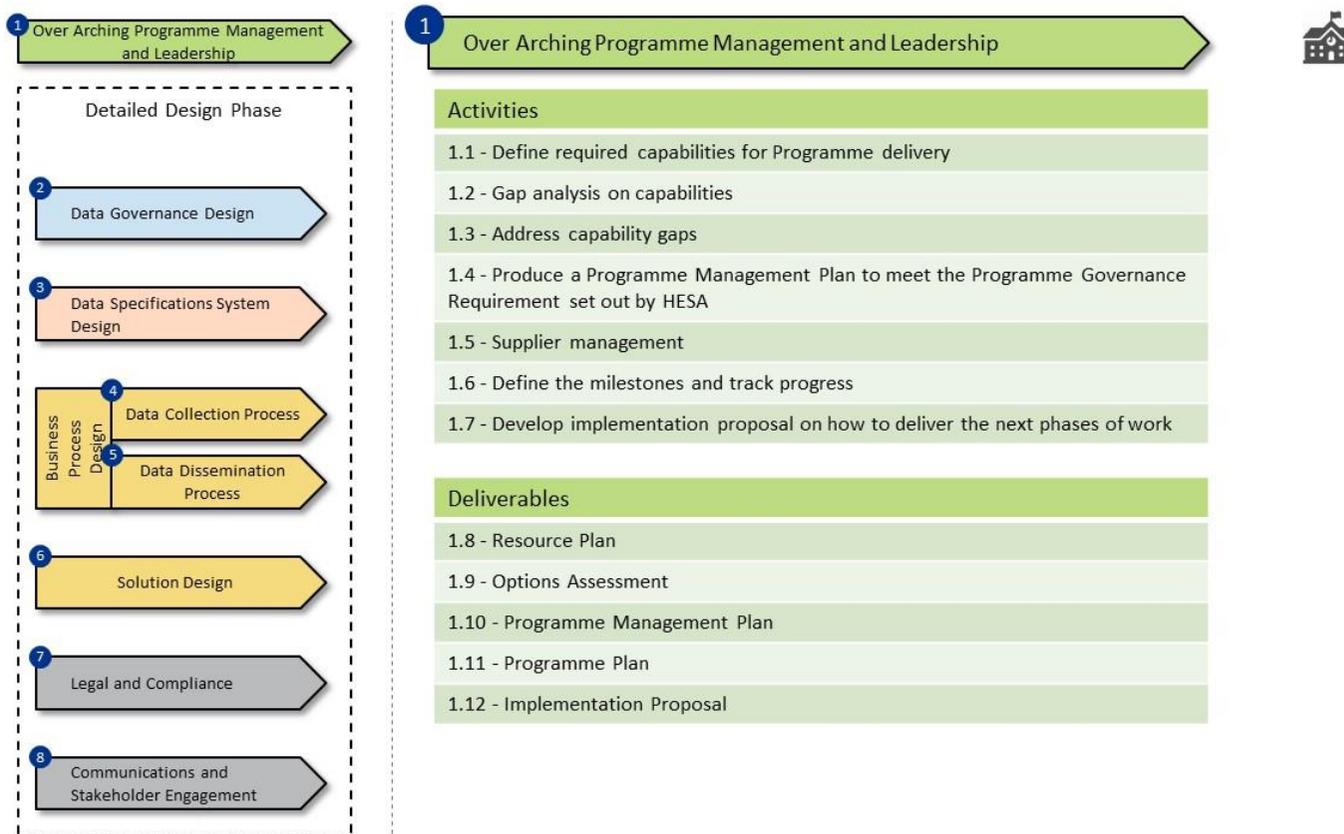


4.4 Detailed design activities and deliverables

This section describes the outcomes of the Detailed Design phase shown in the indicative programme schedule above. The building icons shown on the right indicate the level of external stakeholder engagement on a scale of high, medium and low. The activities and deliverables shown in this section are indicative and may be considered as Sprint Zero activities to initiate the project and create the product backlog.

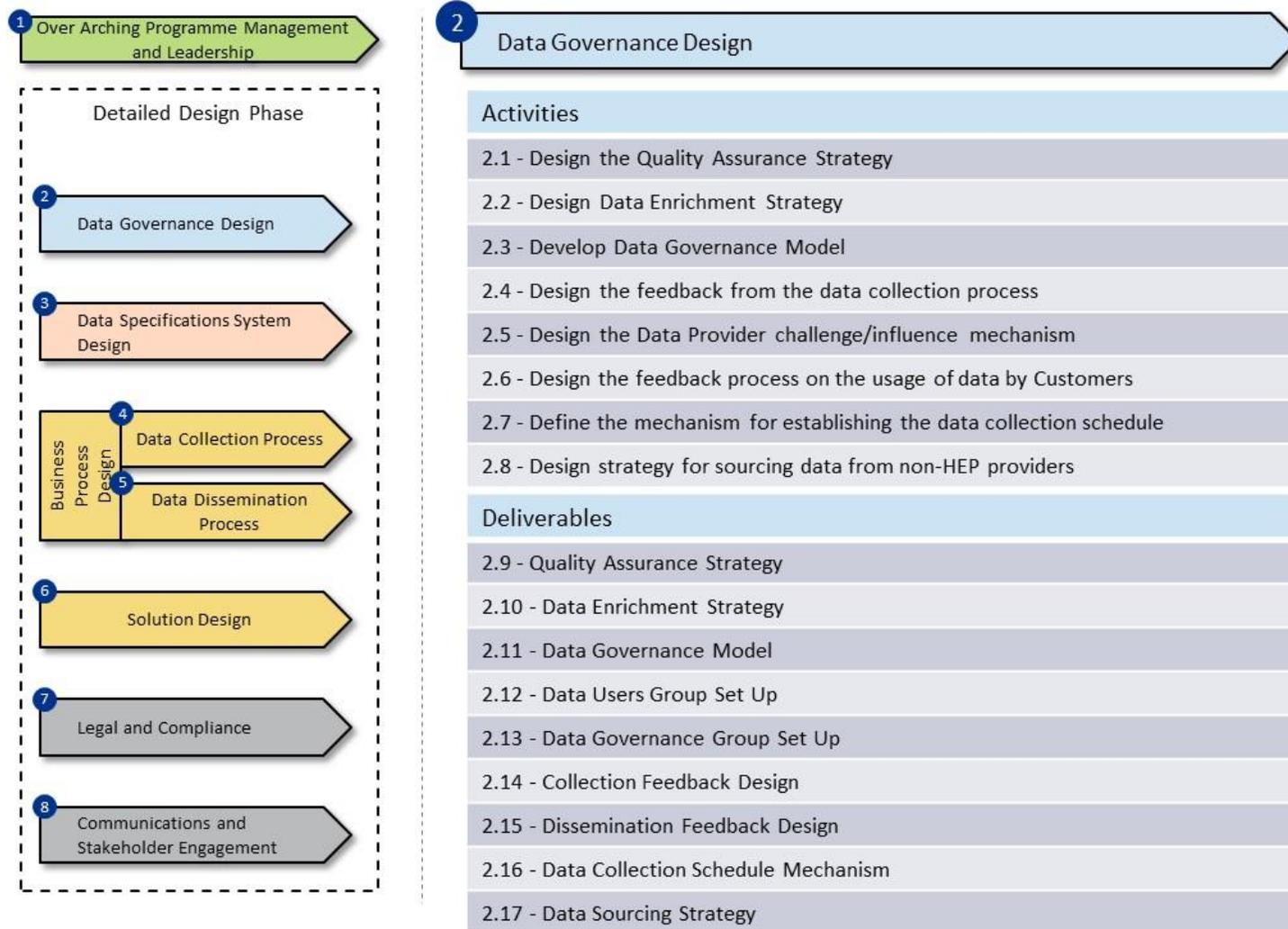
A Senior Responsible Owner appointed by HESA will lead this activity supported by a Programme Manager from the supplier.

Figure 3 Programme leadership



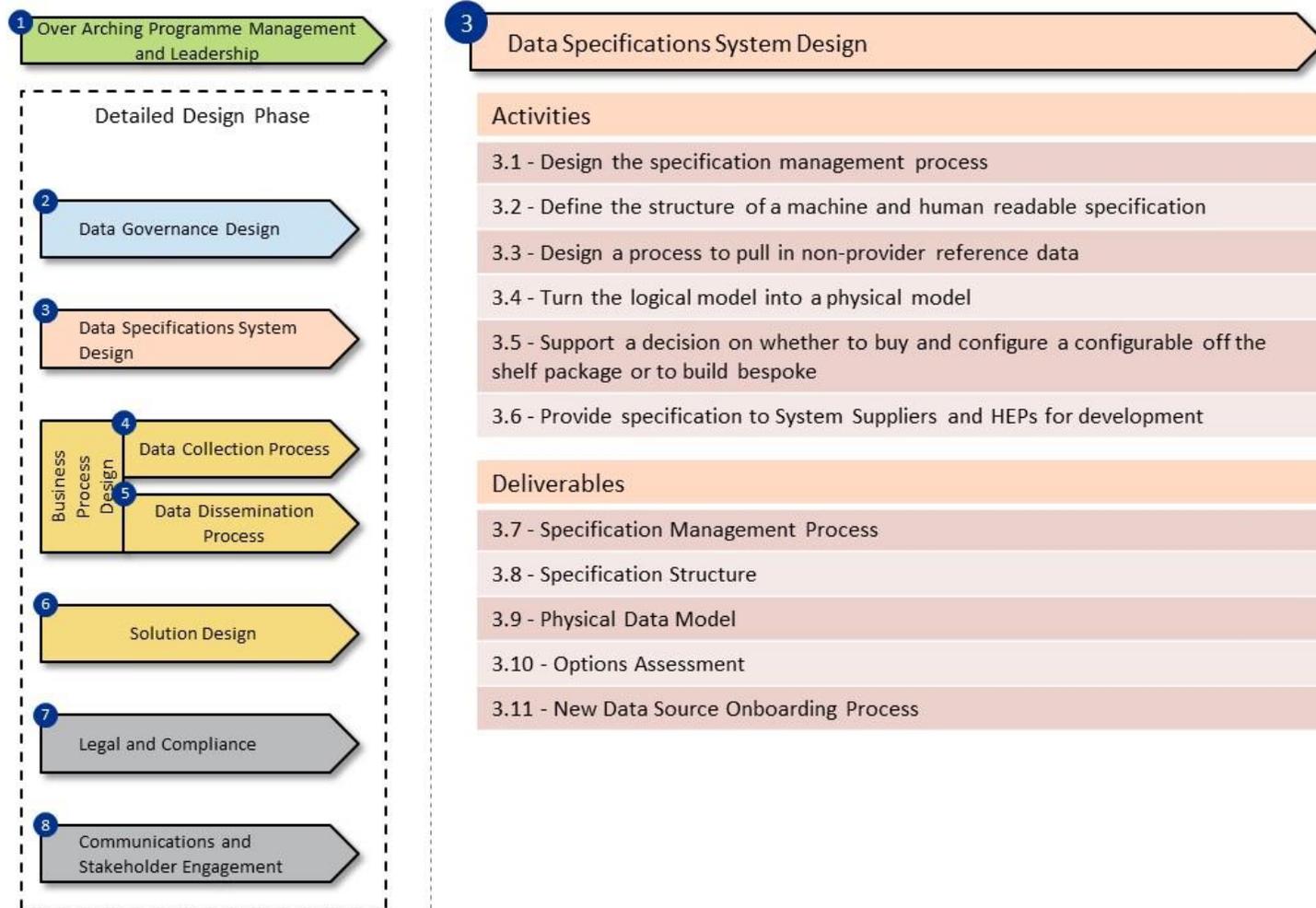
It is envisaged that HESA will lead this work stream supported by the supplier.

Figure 4 Data governance



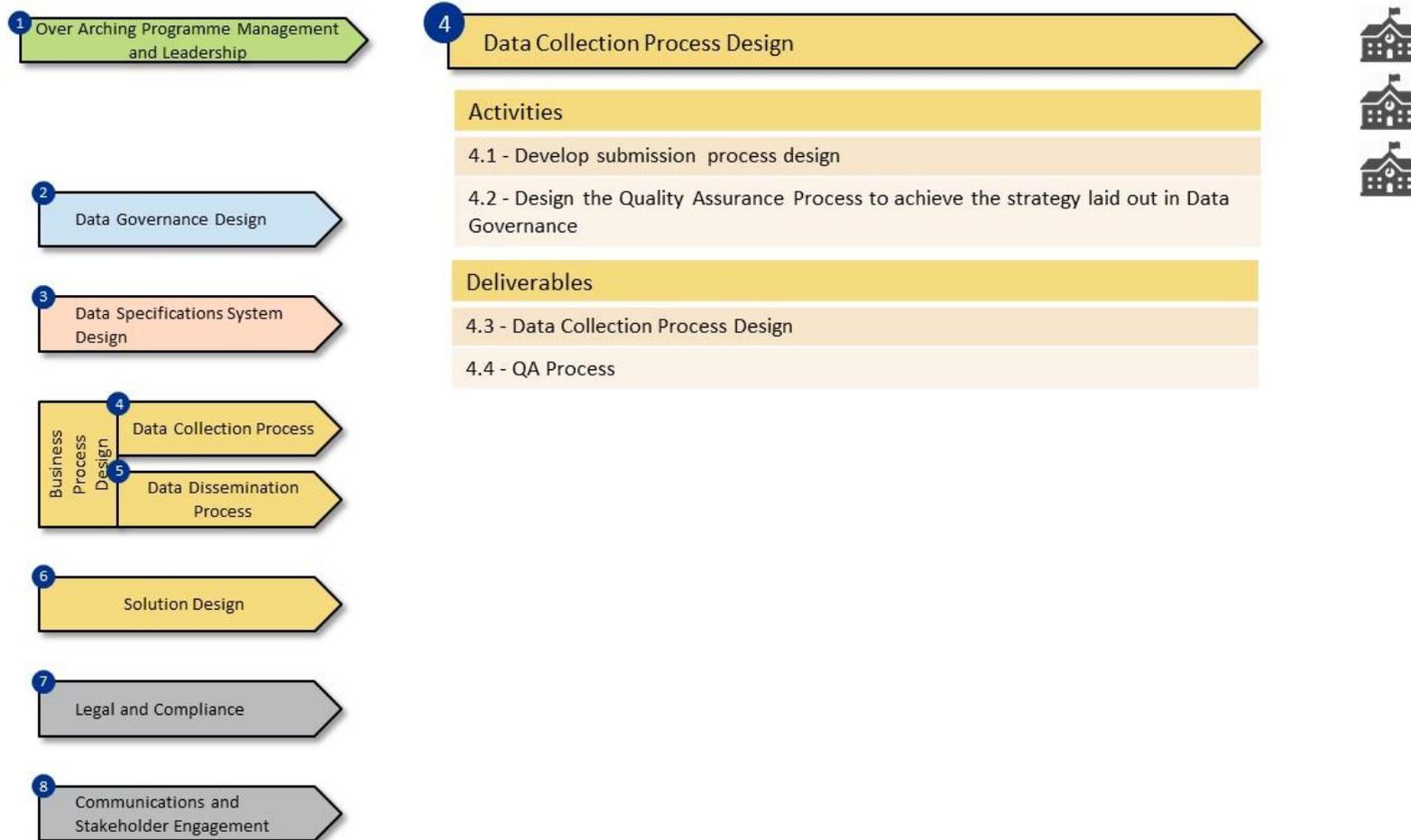
It is envisaged that the supplier will lead this workstream

Figure 5 Data specification & system design



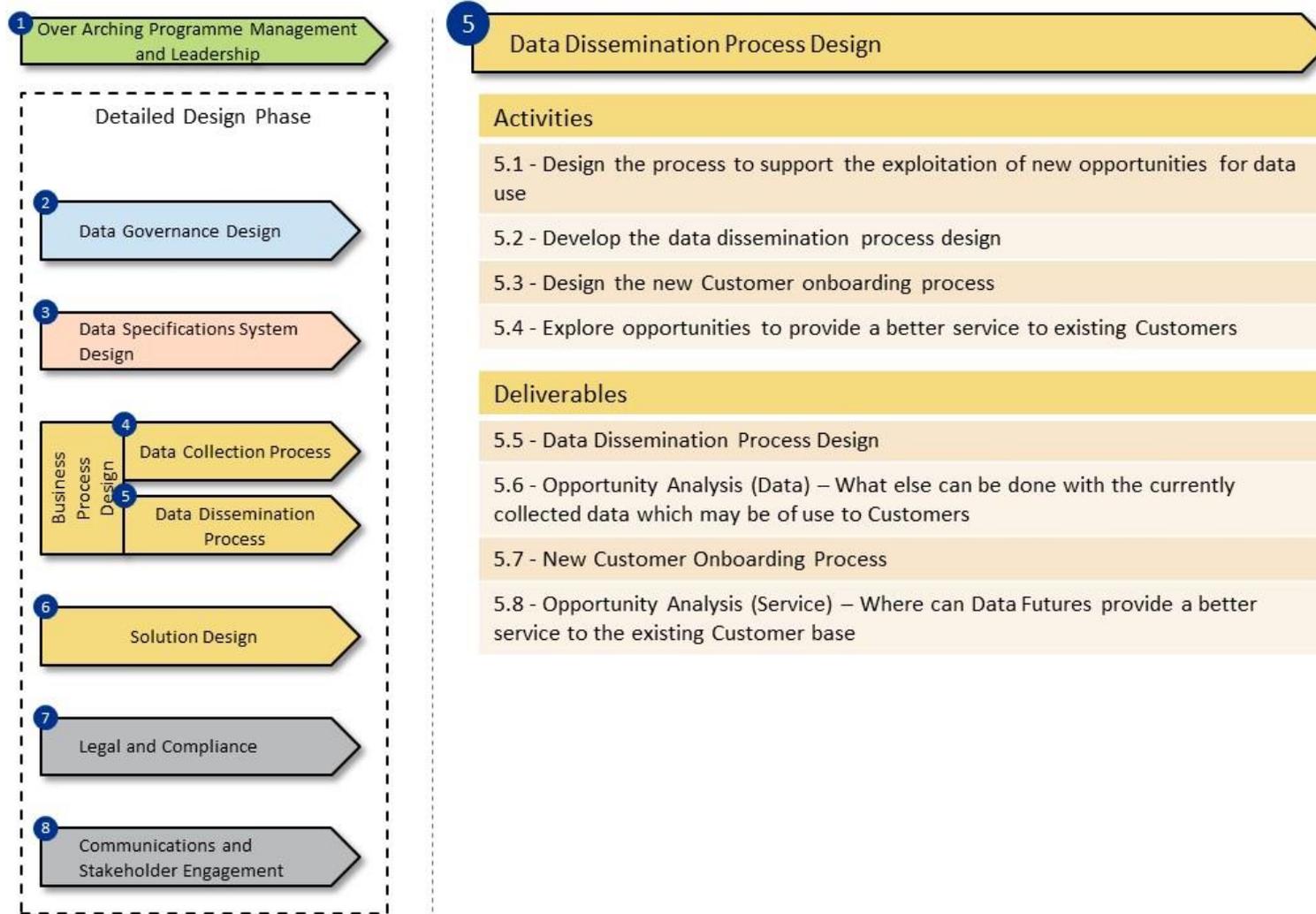
It is envisaged that the supplier will lead this workstream

Figure 6 Data collection process design



It is envisaged that the supplier will lead this workstream

Figure 7 Data dissemination process



It is envisaged that the supplier will lead this workstream

Figure 8 Solution design - activities

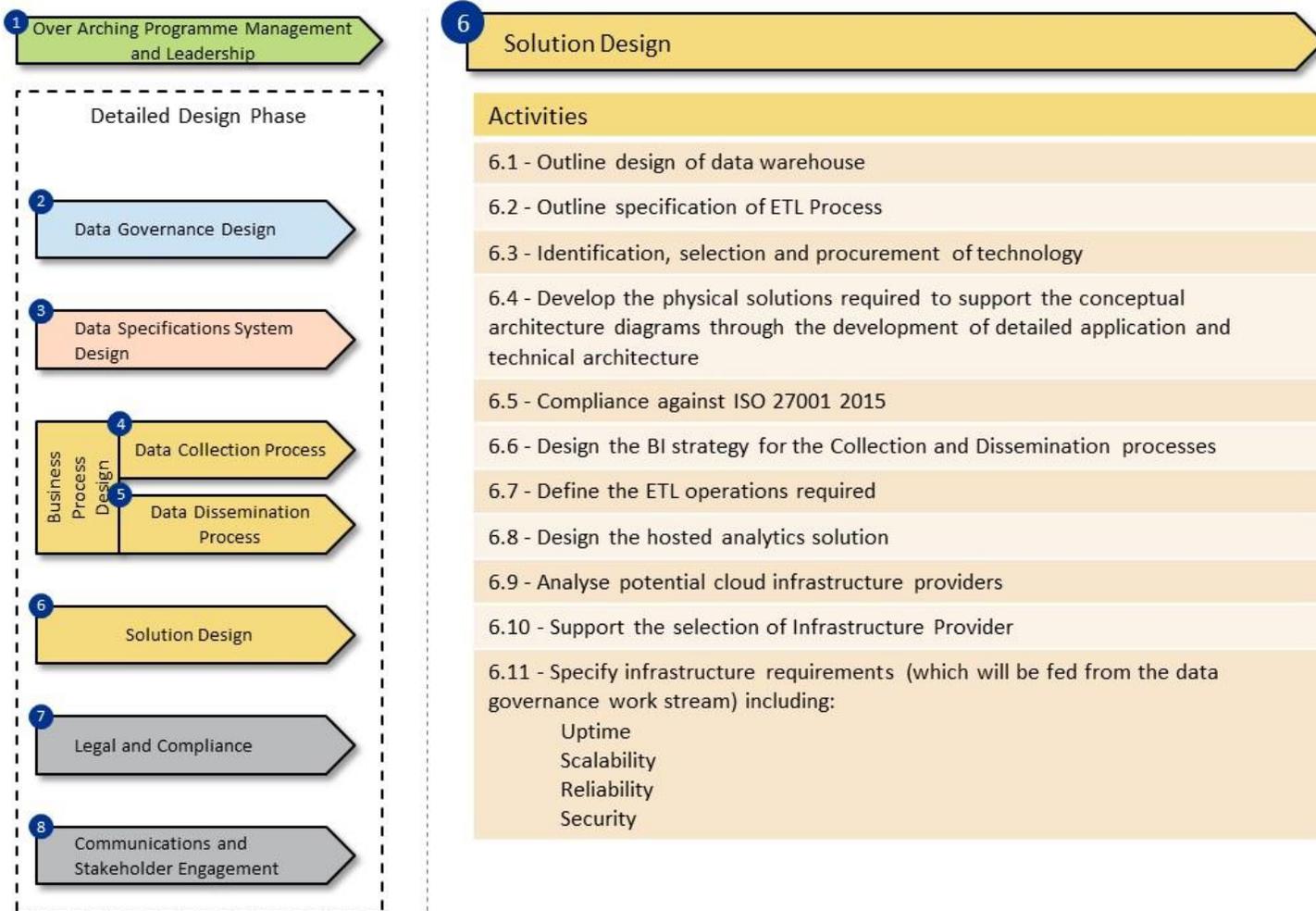
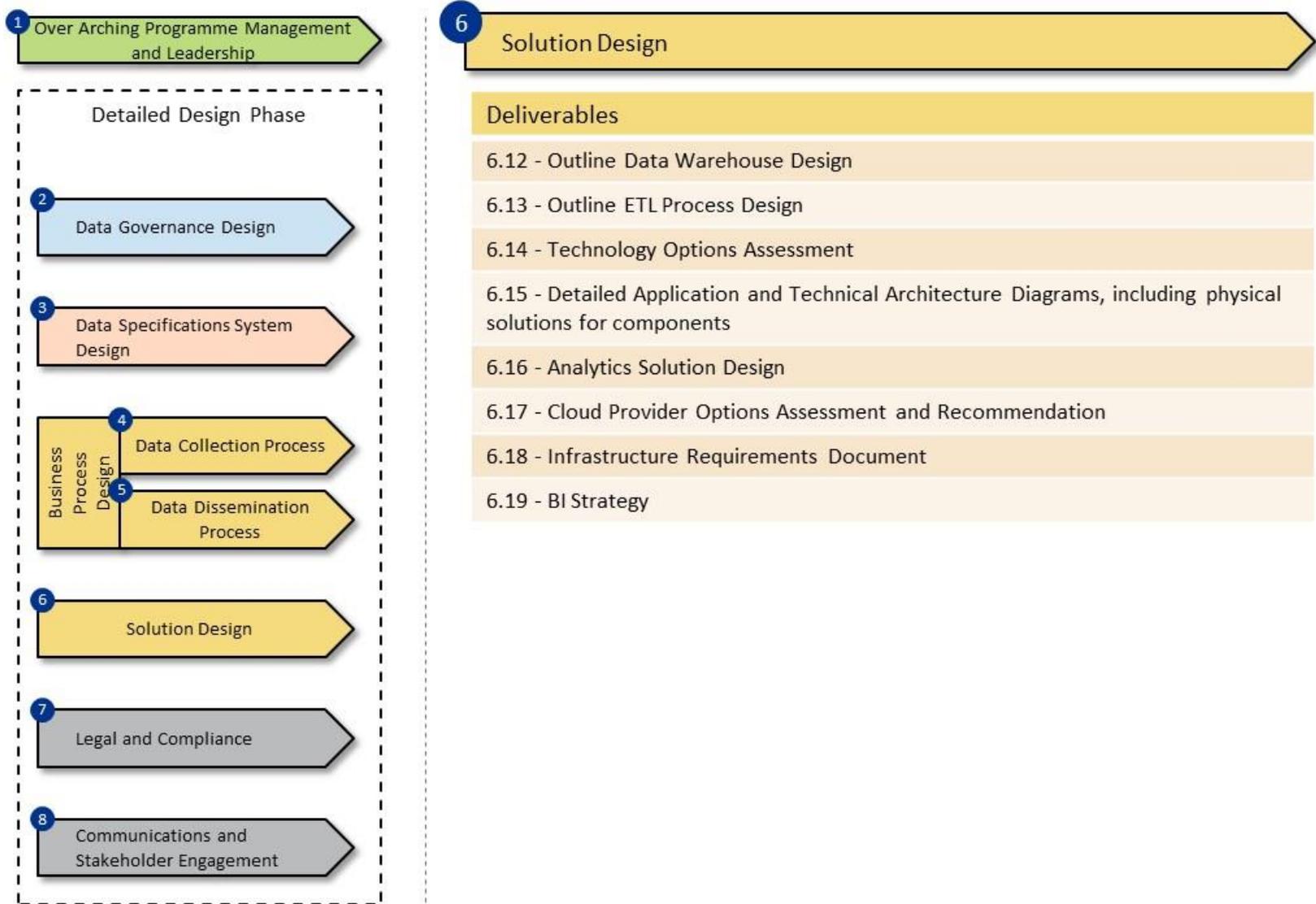
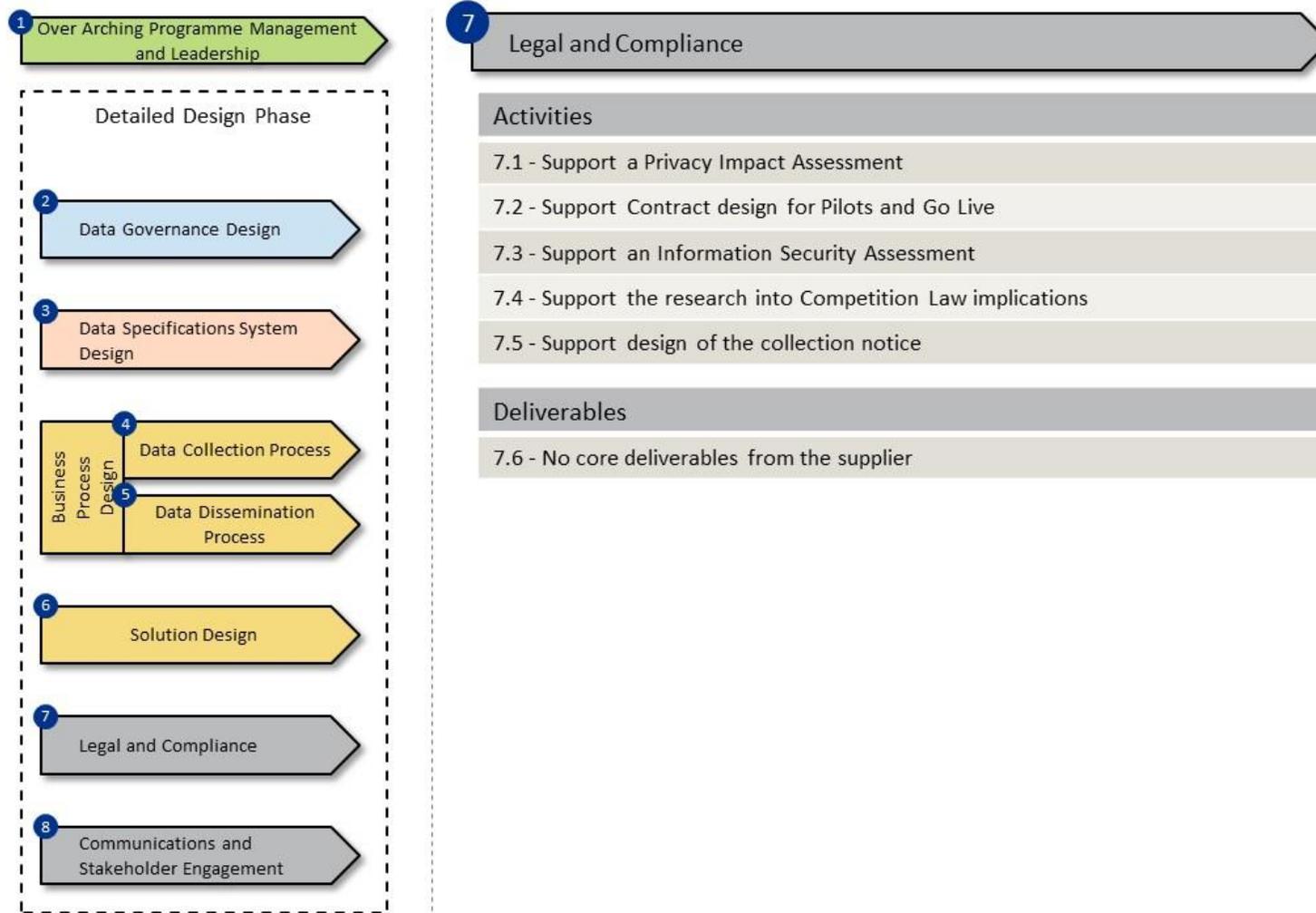


Figure 9 Solution design – deliverables



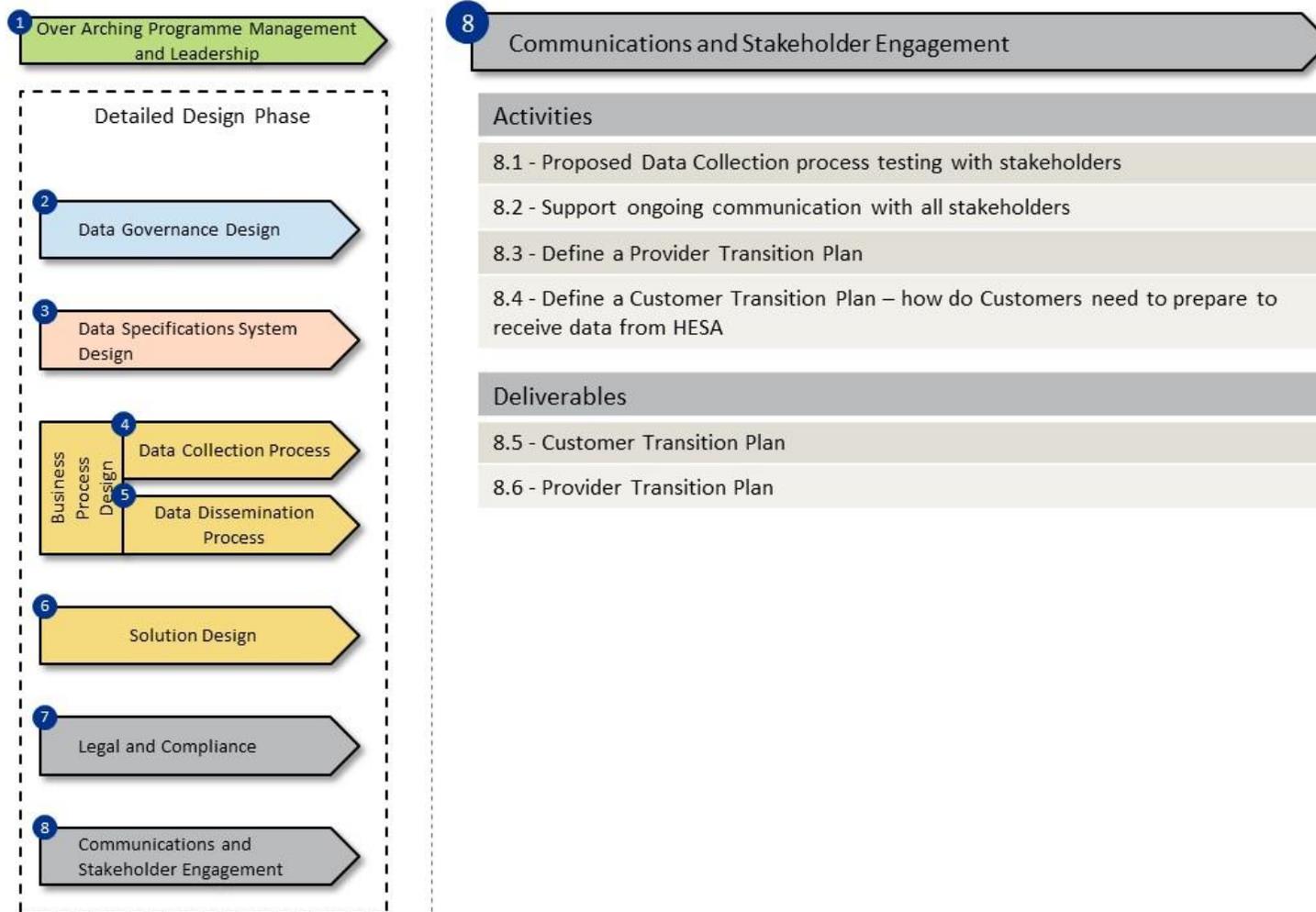
It is envisaged that the HESA will undertake this workstream with support from the supplier

Figure 10 Legal and compliance



It is envisaged that the HESA will undertake this workstream with support from the supplier

Figure 11 Stakeholder engagement



5. Business Processes

HESA has mapped the 'To Be' processes at a high level, these are summarised in this section.

5.1 Collection strategy

The diagram below shows HESA's target collection schedule, it shows users (i.e. HE providers) submitting data which is then being made available at specified points during the year either for 'Operational' or 'Statutory' purposes. HESA is currently engaging with the sector to specify this schedule in more detail particularly when the points will occur, what data will be collected at each point, who it will be made available to and how it will be used. This work will be completed before the supplier is appointed. However the schedule will be subject to change management and may develop over the lifetime of the project.

Figure 12 Conceptual collection and dissemination schedule

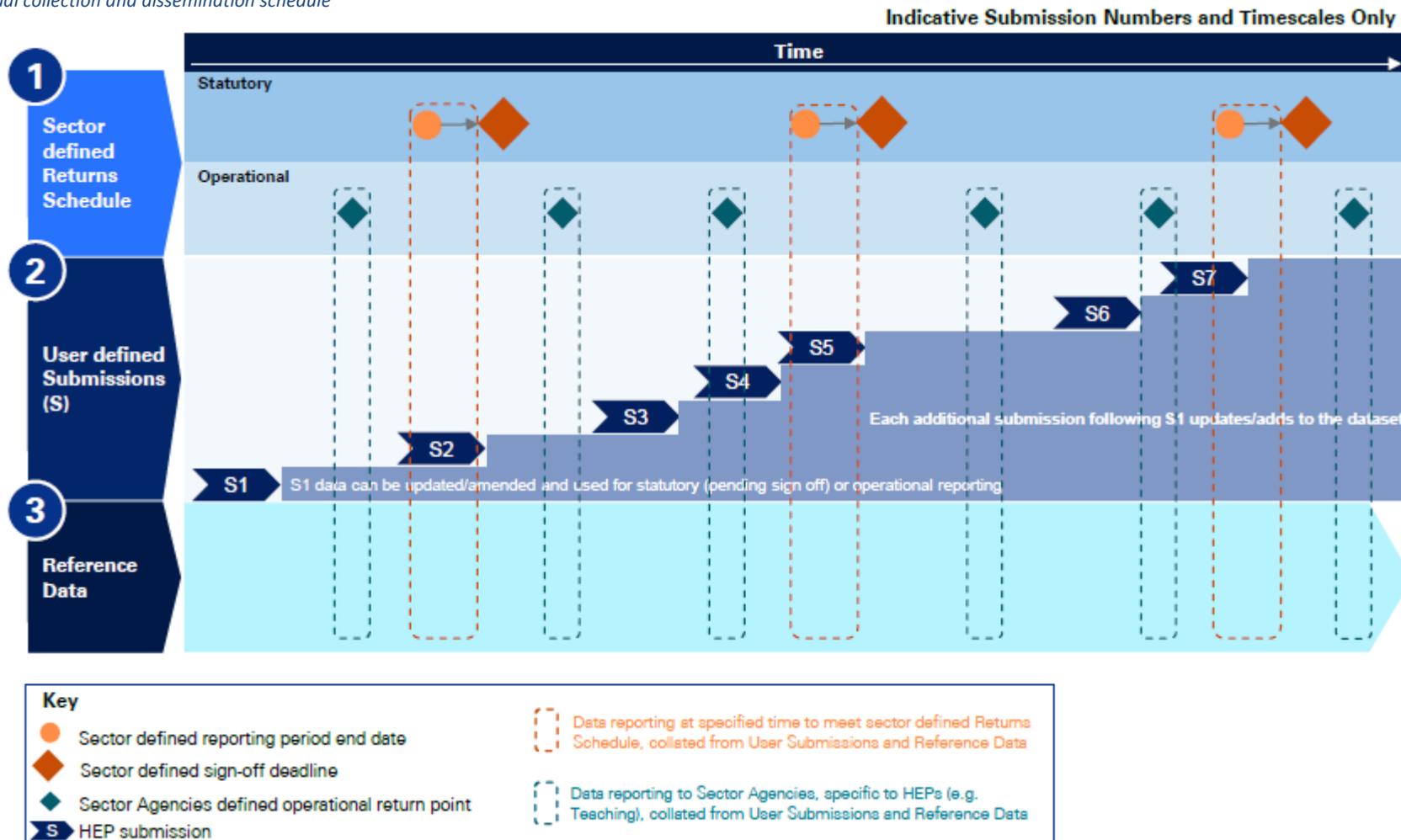


Figure 13 Statutory collection and dissemination

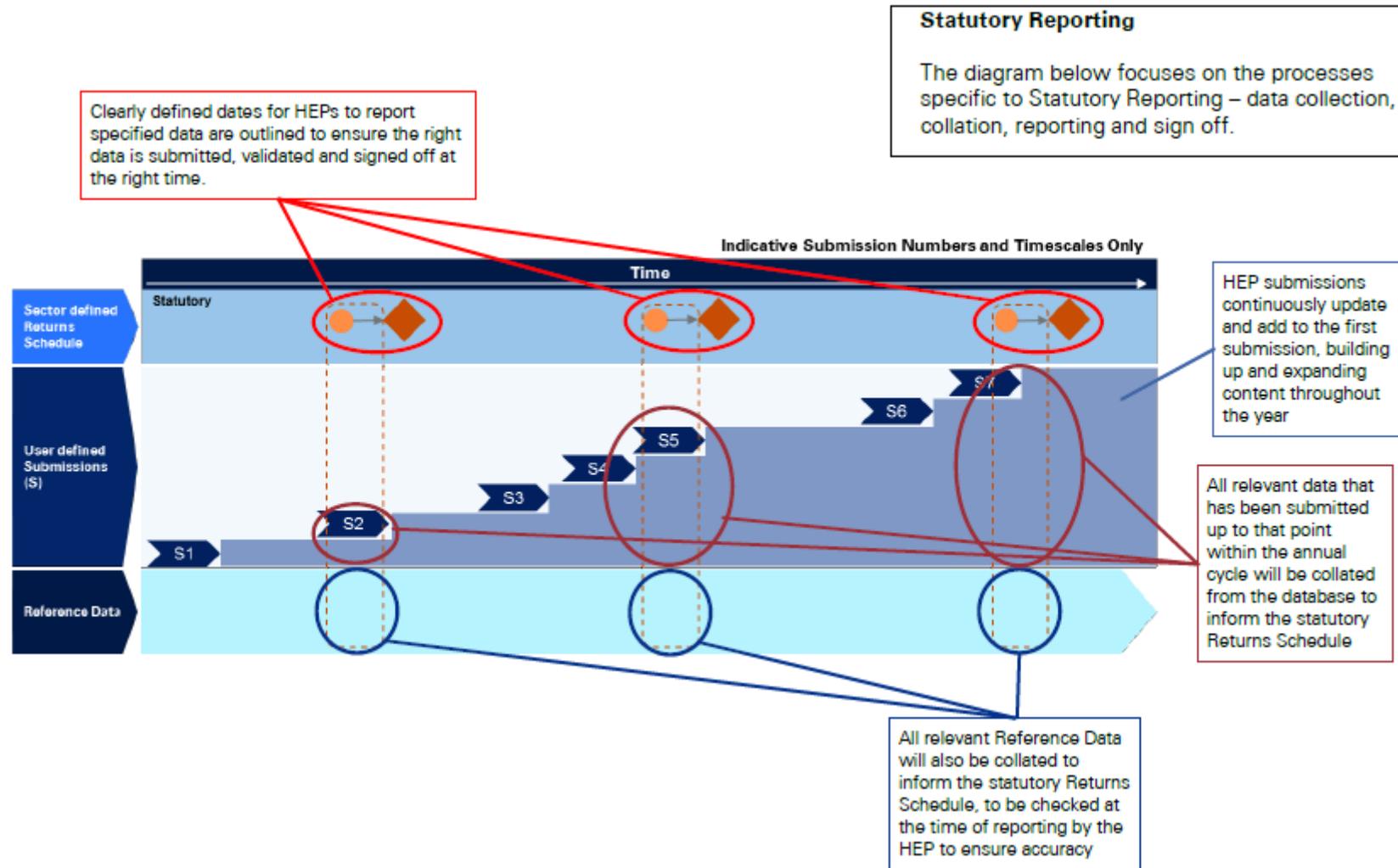
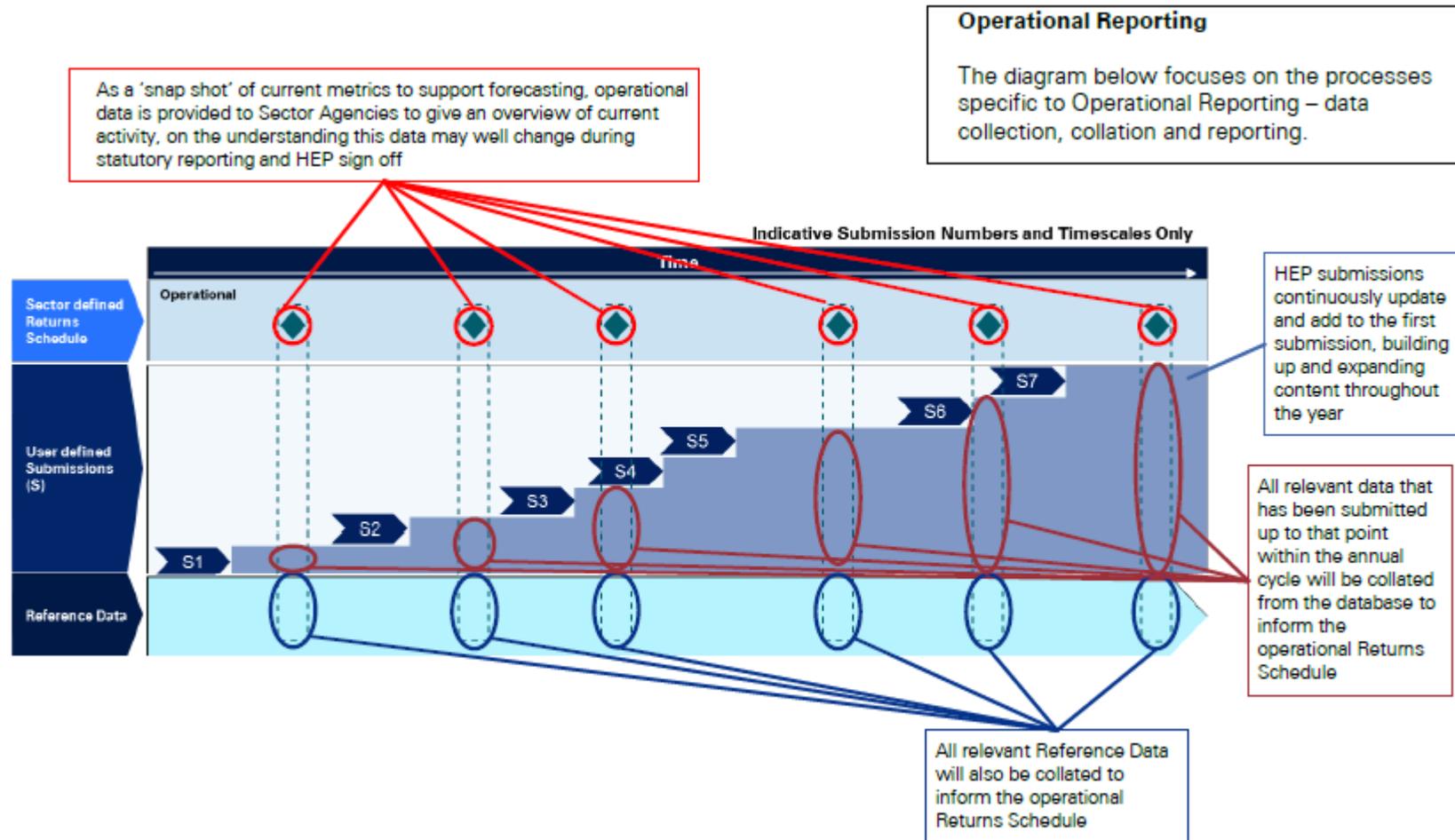


Figure 14 Operational collection and dissemination



5.2 Process maps

HESA has developed high level 'To Be' process models shown below. Working with HESA the supplier will develop the processes to the activity level and define the operating procedures.

Figure 15 Identify data required

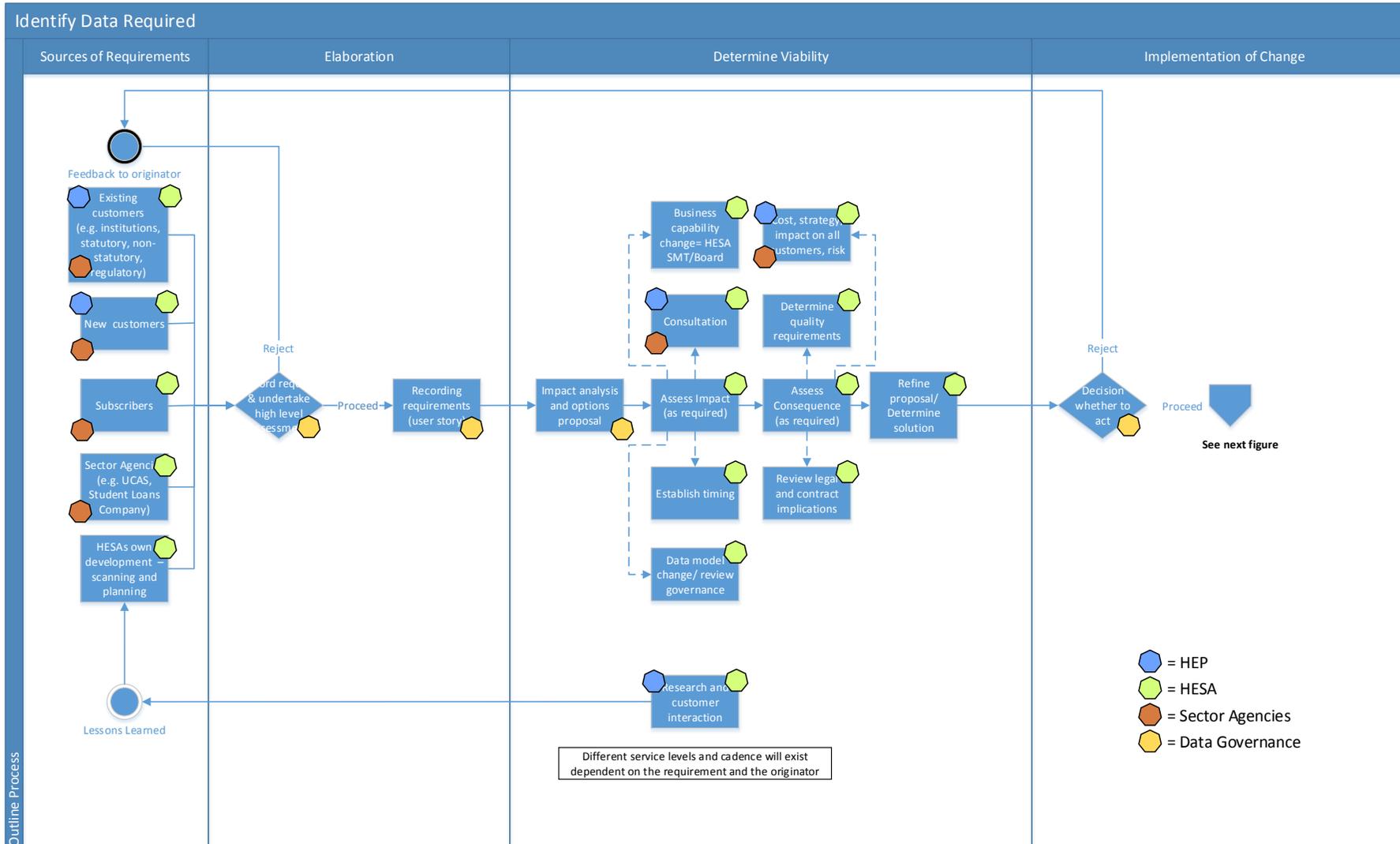


Figure 16 Implement change

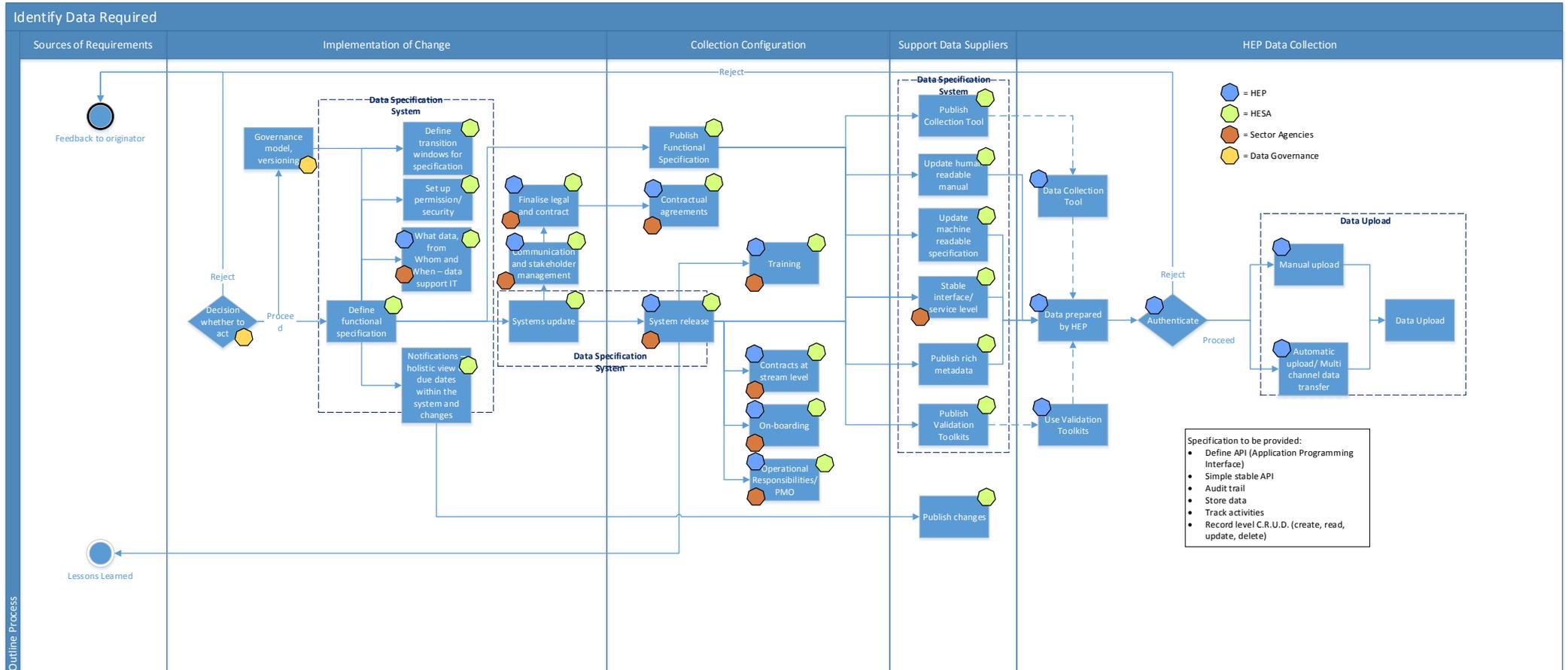


Figure 17 Check data, validate and sign-off

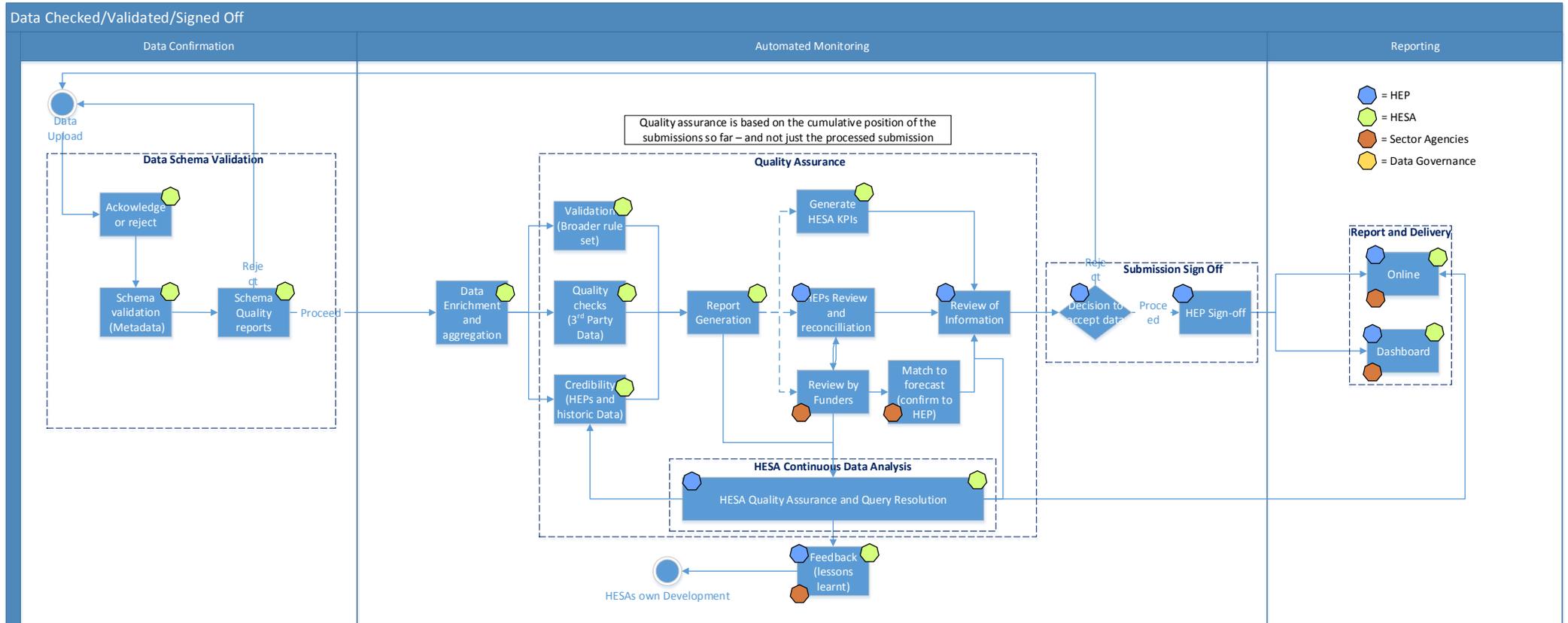
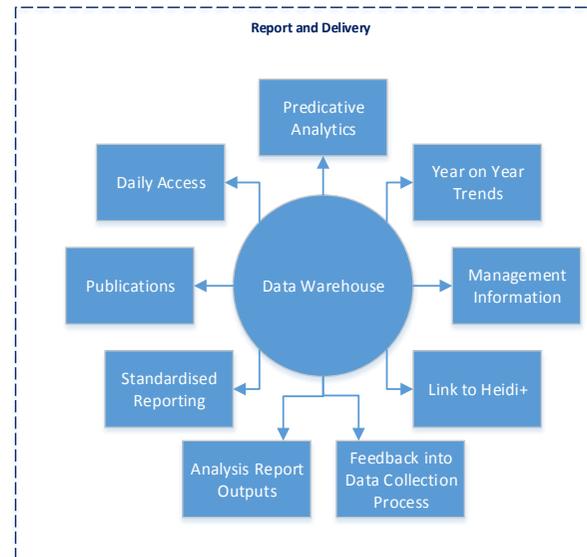


Figure 18 Report and share data

Data is Reported/Shared

The final data must satisfy the requirements of users/statutory bodies/regulatory bodies as well as meeting the HESA Vision 2020. As a minimum, the data must meet and address the following requirements:

- Daily access outside of data warehouse
- Manual/bulk, automated/API feed into data warehouse
- Data warehouse (internal)
- Front end to query data from data warehouse – query As-Is also needed
- Derived field built into data warehouse
- Historical data to be uploaded (data warehouse)
- Management information statistics/ dashboard also feeds to SCs
- A sense of the quality
- Ability to link to old data sets (data warehouse)
- Autofill statistical first release tables at any point
- Link data output to CRM
- External data to map/link to HESA data mechanism
- Audit trail to track changes/identify changes made/rational notes
- Ability to pull snapshot of auto filled template at any time
- Security for data trans feeds
- HEPs getting own data back (data supply file)
- Data manager, IT , various roles, permission access/ field level
- Previewing back to HEP key outputs like league tables
- Ability to generate As-Is/ when published time stamping version history



6. Technical building blocks

In order to help potential bidders understand the scope of the technical solution required, HESA has developed a model of the logical building blocks required.

Figure 19 Logical building blocks

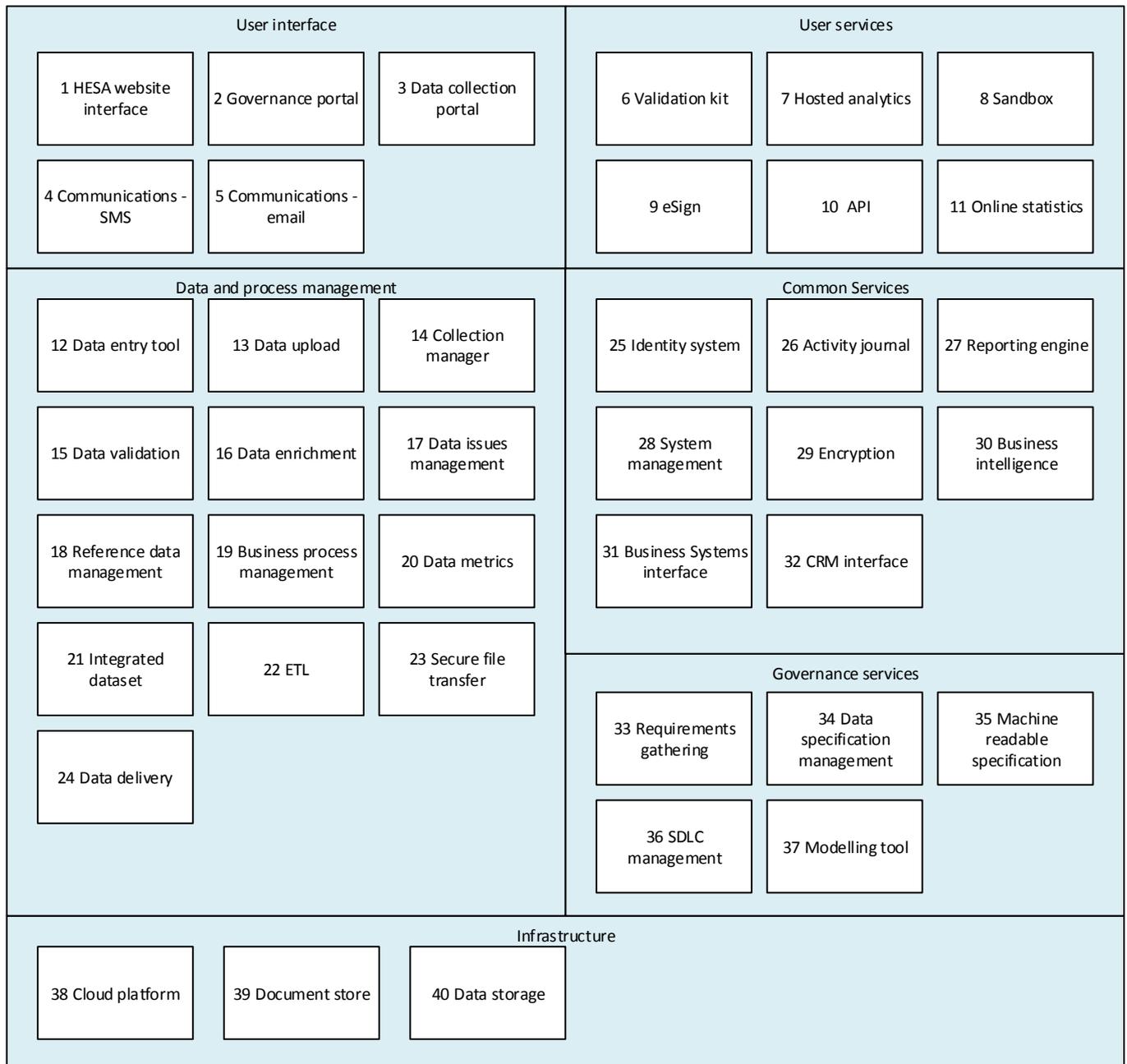


Table 1 Logical building block description

| ID | Name | Description | Exists ¹ |
|----|------------------------|--|---------------------|
| 1 | HESA website interface | The ability to interface to and present information to HESA's corporate web site and the ability for users with appropriate authorisation to access the data collection and governance portals. | Y |
| 2 | Governance portal | A portal to govern the management of meta and reference data and present that data to users | |
| 3 | Data collection portal | A portal to enable data providers to access data collection, quality assurance and messaging tools related to the data collection process. The portal must also provide HE providers with the HUSID (HESA Unique Student Identifier) look up service. | U |
| 4 | Communications - SMS | Ability to send mobile messages e.g. for the purpose of 2 factor authentication | |
| 5 | Communications - email | Ability to send messages and alerts by emails and to track email delivery and receipt | U |
| 6 | Validation kit | Downloadable tools to be used by HE providers to valid their data before submission to HESA | U ² |
| 7 | Hosted analytics | The capability for HESA to provide self-service analytics and provide access to tools to analyse data | |
| 8 | Sandbox | A sandbox environment to enable HE providers and data customers to test their processes and systems | |
| 9 | eSign | Ability for HE providers and data customers to electronically sign documents (e.g. data submission sign-off) | |
| 10 | API | The data collection and dissemination system must be accessible through API's from third party systems such as data providers and customers | U |
| 11 | Online statistics | The capability to formulate statistical reports of the student data collected and publish them to the HESA web site | U |
| 12 | Data entry tool | A tool that enables HE providers (typically smaller HE providers that do not use student information systems to submit their data) to enter their data and validate it prior to uploading to HESA. | |
| 13 | Data upload | The data upload service used by data providers, to submit their data to HESA. The service will provide a user interface and options for uploading data, the service will accept multiple data types and user defined submissions (see Figure 12 Conceptual collection and dissemination schedule) and initiate data validation. As well as HE providers other data sources may include for example the capability to capture qualification information from Awarding Bodies. | |
| 14 | Collection manager | The capability for users to specify, view and manage the data collection process; a user interface accessed by HESA through the data collection portal | |

¹ **KEY Exists:** Y (Yes) indicates the building block exists and U (Update) indicates the building block exists but may require update or replacement

² Currently exists but we require replacement because of significant changes to the collection specification.

| ID | Name | Description | Exists ¹ |
|----|-----------------------------|---|---------------------|
| 15 | Data validation | A key component used to check and feedback on the quality of the data when uploaded includes: schema validation, business rules, exception rules, continuity and credibility checks and aggregation routines. The data validation component must be easily configurable and flexible to enable data validations to be defined and edited and enable advanced data validations e.g. using contextual intelligence. | U |
| 16 | Data enrichment | The capability to derive values and calculate the contents of fields based on other fields, reference data, or by data matching and/or linking with third party sourced data | U |
| 17 | Data issues management | The capability for the data collection system, HESA and specified data customers to raise data issues (i.e. on data collected by HESA prior to it being signed-off by HE providers) and enable the issues and responses to be routed (to the appropriate users), viewed, tracked and managed. | |
| 18 | Reference data management | Capability to manage transactional reference data (e.g. institutional, course and module data) so it can be used in the data collection and dissemination process; particularly the quality assurance process. | U |
| 19 | Business process management | The capability to automatically track and manage the workflow of transactions throughout the data collection and dissemination process. | U |
| 20 | Data metrics | Defined metrics to measure: the quality of the data, how the data is accessed and consumed. | |
| 21 | Integrated dataset | A data warehouse or an alternative data storage method for managing an integrated data set | |
| 22 | ETL | Extract Transform Load (ETL) capability to automate the transfer and integration of data from ingestion into HESA, making data available to internal systems and data warehousing | U |
| 23 | Secure File Transfer | Secure File Transfer – the capability to send and receive data files securely to and from data providers and HESA customers | U |
| 24 | Data delivery | The capability to disseminate data to HESA’s customers using secure file transfers | U |
| 25 | Identity system | Capability to authenticate a user accessing the HESA system and control their access rights | |
| 26 | Activity journal | The ability to record data processed and changes to data e.g. to enable reporting of data provenance | |
| 27 | Reporting engine | A reporting engine which is flexible, easy to maintain and has the ability to query multiple data types and sources. The engine must have the capability to automatically generate reports, at specified points in the data collection and dissemination process and provide reports to HESA, HE provider and customer users | U |
| 28 | System management | Capability to manage the performance of the system identifying and alerting errors and exceptions that occur in the system | |
| 29 | Encryption | The capability for data encryption at rest and in transit, management of the encryption process and keys | |

| ID | Name | Description | Exists ¹ |
|----|--------------------------------|---|---------------------|
| 30 | Business intelligence | A flexible and scalable Business Intelligence system that can be used to meet internal and external user requirements including analysis of data processing metrics, data usage metrics and analysis of the data collected through the student data collection process. | |
| 31 | Business systems interface | Capability to interact with HESA's Business Systems (e.g. Enterprise Resource Planning systems) including the ability to interface data into and out of the business system | |
| 32 | CRM interface | Capability to interact with HESA's Customer Relationship Management system including the ability to interface data into and out of the CRM system and to query customer information (e.g. to look up the contract status of a HESA customer) | |
| 33 | Requirements gathering | The capability to efficiently collect and manage requirements from HE providers and customers about their data and reporting needs; including a consultation platform to enable stakeholders to provide feedback on proposed specifications and changes. | U |
| 34 | Data specification management | Capability to manage human readable data specifications i.e.: meta data (including versioning), quality assurance rules and coding frames (i.e. allowed values) and standards management. | U |
| 35 | Machine readable specification | Capability to enable data provider systems to automatically manage changes to data specifications and APIs (e.g. to automate the population of changes to coding frames and field definitions to HE provider systems) | |
| 36 | SDLC management | Capability for HESA to manage the end to end Software Development Life Cycle including release management and new API's | U |
| 37 | Modelling tool | Tools for process and data modelling to support the SDLC and data governance | |
| 38 | Cloud platform | A cloud platform service hosted and supported in a secure, scalable cloud environment | |
| 39 | Document store | An area to store files that are uploaded to the collection system (on the cloud platform) | |
| 40 | Data storage | Physical storage of data to support the collection and dissemination processes, analytical and information services (on the cloud platform) | U ³ |

³ Data storage currently exists on-site

7. Delivery capabilities

Figure 20 Supplier capabilities required

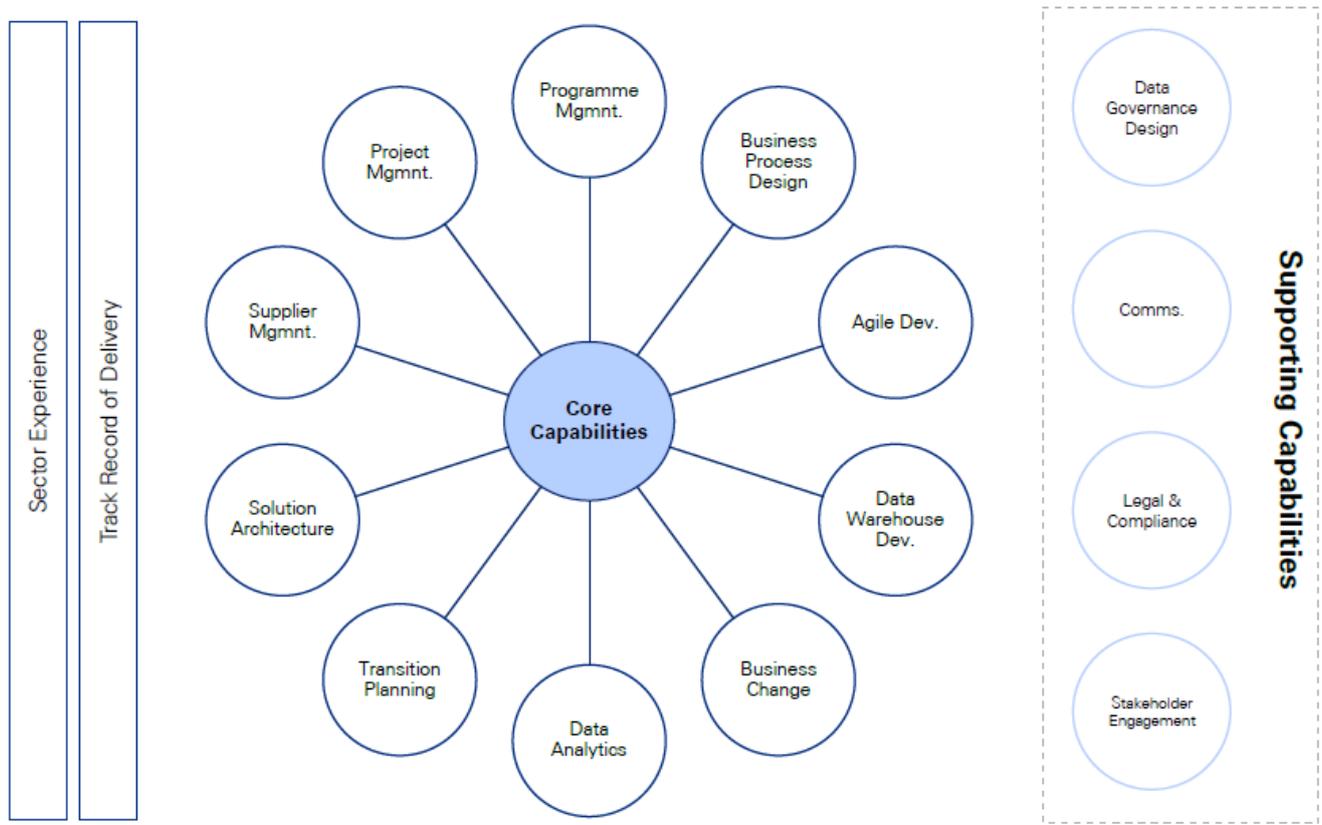


Table 2 Capability description

| Capability | Supplier Role | Capability Description | Services Required by the Supplier |
|-------------------------|-----------------------------------|--|--|
| Programme Management | Owned and staffed by the Supplier | Planning, coordination and management of the multiple different work streams and projects which make up the Data Futures Programme | Programme Management Office support |
| Business Process Design | Owned and staffed by the Supplier | Mapping of current business processes, as well as future state processes, to the required level of detail | Business Process Mapping software licenses |

| Capability | Supplier Role | Capability Description | Services Required by the Supplier |
|----------------------------|---|--|--|
| Agile Development | Owned by the Supplier and jointly staffed with HESA | Experience of the chosen technologies to enable application level design and development of solutions, as well as Agile development methods including User Centred Design | Tools and software licences required to design and develop using the chosen technology |
| Data Warehouse Development | Owned by the Supplier and jointly staffed with HESA | Experience and knowledge of the chosen technologies to enable design and development of the Data Warehouse | Tools and software licences required to design and develop using the chosen technology |
| Business Change | Jointly owned and staffed by the Supplier and HESA | Enabling the HESA behavioural change required to deliver business benefits, including Involvement strategies, Organizational alignment, Training and Measuring benefit. Supplier leading on training. | Training environment |
| Data Analytics | Owned by the Supplier and jointly staffed with HESA | Design, planning, development around producing reports and ensuring reporting requirements and formats are defined and continuously reviewed | Tableau licenses |
| Transition Planning | Owned and staffed by the Supplier | Coordinating the activities required to efficiently move from the current state to new systems and processes | N/A |
| Solution Architecture | Owned and staffed by the Supplier | Technical design and planning of the solutions required for the Data Futures Programme, and the evaluation of potential new IT Application solutions | N/A |

| Capability | Supplier Role | Capability Description | Services Required by the Supplier |
|---------------------|-----------------------------------|---|-----------------------------------|
| Supplier Management | Owned and staffed by the Supplier | Managing mutually beneficial relations with suppliers, outsourcers, and other third-parties to ensure consistent with business expectations and contractual commitments | N/A |

Figure 21 HESA capabilities

| Capability | Supplier Role | Capability Description | Services Required by the Supplier |
|------------------------|--|--|-----------------------------------|
| Data Governance Design | Owned by HESA and jointly staffed with the Supplier | Designing Data Stewardship, governance frameworks that provide control and mitigate risks, and supporting decision making | N/A |
| Communications | Owned by HESA with support from the Supplier | Develop and disseminate detailed communication plans and delivery work for internal and external communication from HESA | N/A |
| Legal & Compliance | Owned by HESA with support from the Supplier | Identify, assess and mitigate risks, while ensuring internal controls meet compliance requirements Support the production of Legal documentation | N/A |
| Stakeholder Engagement | Owned by HESA and jointly staffed with the Supplier | Engaging with internal and external stakeholders to ensure they remain informed and participate | N/A |
| Project Management | Supplier project managing the Supplier work streams and HESA managing the HESA work streams. | Planning, coordination and management of projects within the Data Futures Programme incorporating well defined scope, commitment of resources, and required milestones | N/A |

8. Business as usual

At the end of the project the following will be transitioned to HESA from the supplier:

- Knowledge and documentation related to the business, application, data, technical and solution architectures to enable HESA to maintain and develop them to meet future requirements.
- The developed end to end solution with appropriate training and documentation to enable HESA to operate, maintain and develop the solution.
- Licences and support agreements related to any COTS products used in the solution.
- Contracts related to the ongoing use of the cloud infrastructure platform
- Source code and specifications for any bespoke developments and configuration
- Process maps, operating procedures and training materials related to the operation and maintenance of the business processes and systems.
- Data specifications, models and governance processes

HESA will require the solution to be fully developed for a full business pilot for the 2018/19 academic year. The supplier will be required to start transitioning to HESA in the 2018/19 academic year and for this to be completed by the end of the academic year 2019/20. A transition plan should be developed during the Detailed Design phase described earlier.

APPENDIX A – Data item definition and quality rules

Example of a data item definition ‘term time postcode’ (Student.TTPCODE)

This example can be viewed from the following link

www.hesa.ac.uk/index.php?option=com_studrec&task=show_file&mntl=15051&href=A^ ^TTPCODE.html

This field is required for all students except those studying by distance learning and those on placements.

This can be a postcode from any point during term-time, but should not be a correspondence or summer contact address. It is expected that in many cases the postcode will change between years. Where students spend all or part of the year in a hall of residence the provider will be able to insert a generic postcode for that hall into all of the relevant records.

Although completion of the field will be compulsory for all students, 'Unknown' values will be acceptable. If no valid postcode (full or outward only) can be ascertained an empty element should be returned, i.e:

```
<TTPCODE></TTPCODE>
```

It will be possible to provide only the outward part of the postcode if this is all that is known, although that is not expected to arise commonly.

Postcodes returned must pass schema validation of format. Individual postcodes will be validated against the Royal Mail list but only a warning rather than an error will be generated if validation fails. This is intended to assist providers that attach importance to the accuracy of their contact information but who may not be in a position to validate postcodes themselves.

British Forces Post Office (BFPO) postcodes

Where BFPO codes relate to UK locations, providers will need to locate the Royal Mail equivalent postcode for the base.

There is further guidance on valid postcode formats at:

<http://webarhive.nationalarchives.gov.uk/+http://www.cabinetoffice.gov.uk/govtalk/schemasstandards/e-gif/datastandards.aspx>. The business rule structure validation will be slightly less strict than that described in UK Government Data Standards Catalogue, as not all parsers will support the full GDSC definition. All postcodes are validated against valid postcodes at exception.

Quality Rules

| Rule ID | Description | Stage | Tolerance | Status |
|-------------------------------------|---|---------------|-----------|-----------------|
| QR.C15051 .Student.TT PCODE.1 | <i>Student.TTPCODE must exist where any Instance.LOCSDY does not equal 6, 9, D, S or T and the corresponding Instance.REDUCEDI = 00 or 07 unless institution in Wales and Course.COURSEAIM begins with P, Q, R, S or X.</i> | Business Rule | Error | Carried Forward |
| QR.C15051 .Student.TT PCODE.2 | <i>Where exists Student.TTPCODE must not equal EntryProfile.POSTCODE where Student.TTACCOM = 1 or 9 and Course.COURSEAIM begins with H, I, J, C, P, Q, R, S or X or is M22, M26 or M28.</i> | Business Rule | Error | Carried Forward |

| <i>Rule ID</i> | <i>Description</i> | <i>Stage</i> | <i>Tolerance</i> | <i>Status</i> |
|--------------------------------------|--|---------------|------------------|-----------------|
| QR.C15051 .Student.TT PCODE.5 | Where exists Student.TTPCODE must contain: a valid full postcode structure or a valid outward postcode structure. | Business Rule | Error | Carried Forward |
| QR.C15051 .Student.TT PCODE.6 | Student.TTPCODE must not be null when Student.TTACCOM = 1. | Business Rule | Error | Carried Forward |
| QR.C15051 .Student.TT PCODE.7 | Student.TTPCODE should not be null when Student.TTACCOM = 9. | Business Rule | Warning | Carried Forward |
| QR.C15051 .Student.TT PCODE.8 | Where exists Student.TTPCODE should not equal EntryProfile.POSTCODE where Student.TTACCOM = 8. | Business Rule | Warning | Carried Forward |
| QR.C15051 .Student.TT PCODE.9 | More than 50 young (between 18 and 21), full-time, undergraduate students, at parental home (Student.TTACCOM = 2), have Student.TTPCODE (where exists) not equal to EntryProfile.POSTCODE | Exception | Warning | Carried Forward |
| QR.C15051 .Student.TT PCODE.10 | More than 150 students have the same Student.TTPCODE (where exists) and do not have Student.TTACCOM =1 or 9. | Exception | Warning | Carried Forward |
| QR.C15051 .Student.TT PCODE.11 | More than 600 students have the same Student.TTPCODE (where exists) and do not have Student.TTACCOM =1 or 9. | Exception | Error | Carried Forward |
| QR.C15051 .Student.TT PCODE.12 | Where exists Student.TTPCODE must contain a valid full postcode or a valid outward part of a postcode. | Exception | Warning | Carried Forward |
| QR.C15051 .Student.TT PCODE.13 | Where exists Student.TTPCODE must not equal EntryProfile.POSTCODE where Student.TTACCOM = 1 or 9 and Course.COURSEAIM begins with H, I, J, C, P, Q, R, S or X or is M22, M26 or M28. | Exception | Error | Carried Forward |
| QR.C15051 .Student.TT PCODE.14 | Where exists Student.TTPCODE should not equal EntryProfile.POSTCODE where Student.TTACCOM = 8. | Exception | Warning | Carried Forward |
| QR.C15051 .Student.TT PCODE.15 | Student.TTPCODE must not be in Northern Ireland, Channel Islands or Isle of Man where institution is in England, Scotland or Wales and Instance.MODE is in (01,12,23,24,73) and Instance.LOCSY is not coded 6, 9, D, S or T. | Exception | Error | Carried Forward |
| QR.C15051 .Student.TT PCODE.16 | More than 50% of students on courses where Course.TTCID = F have Student.TTPCODE coded as null (unknown) | Exception | Error | Carried Forward |

APPENDIX B – Data principles

The vision for the New Data Landscape is ‘A data and information landscape for Higher Education in the UK that has effective governance and leadership, promotes data standards, rationalises data flows and maximises the value of technology and enables improved data capability’.

To underpin the vision and provide a common basis for all Data Collectors and HEPs to operate the following data principles have been developed:

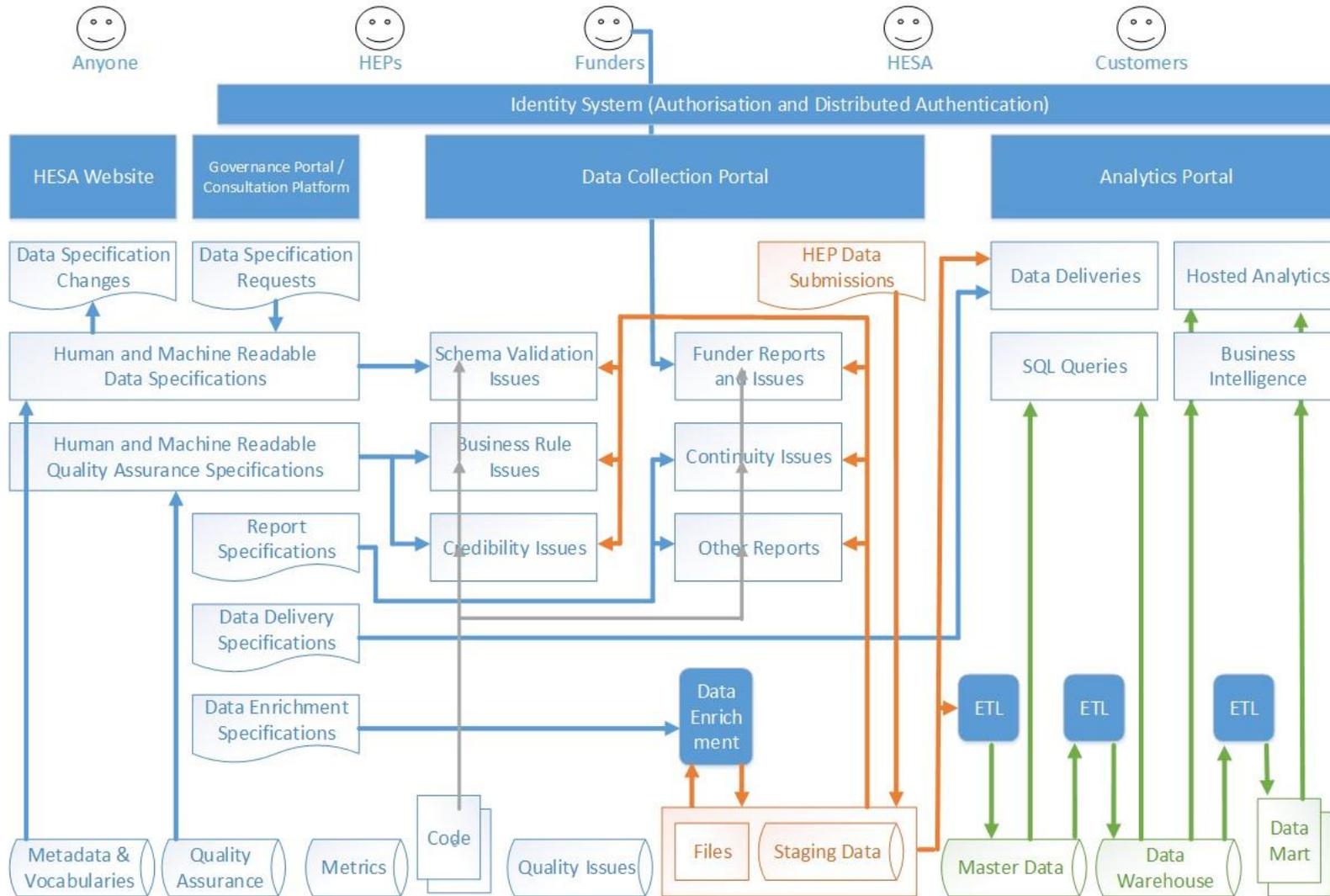
- **Effective Governance:** Data collection requirements in the stakeholders and Higher Education Providers need collective governance that is free from bias to allow for effective implementation and delivery of a New Data Landscape. This governance should be delivered by representatives from HEPs and Data Collectors from across the sector, and represent sector views, with student views represented by the NUS;
- **Adherence to Principles:** These principles of information management apply to all bodies exchanging data with HE sector organisations;
- **Maximum Benefit to HE Sector as a Whole:** Information management decisions are made to provide maximum benefit to the HE sector as a whole including students;
- **Information Management is Everybody’s Business:** All key stakeholders within the HE Sector participate in information management decisions needed to accomplish business objectives;
- **Compliance with the Law:** HE information management processes comply with all relevant laws, policies, and regulations, including competition law;
- **Data is an Asset:** Data is an asset that has value to the HE Stakeholders and is managed accordingly;
- **Data is Shared:** Users have access to the data necessary to perform their duties or answer their query; therefore, data is shared across Stakeholders, where permissible and appropriate;
- **Data Trustee:** Each data element has an owner and a trustee accountable for data quality;
- **Common vocabulary and Data Definitions:** Data is defined consistently throughout the HE sector, and the definitions are understandable and available to all users;
- **Data Security:** Data is protected from unauthorised use and disclosure. In addition to the traditional aspects of national security classification, this includes, but is not limited to, protection of sensitive and proprietary information;
- **Technology Independence:** Applications are independent of specific technology choices and therefore can operate on a variety of technology platforms;
- **Responsive Change Management:** Changes to the HE information landscape are implemented in a timely manner;
- **Interoperability:** Software and hardware should conform to defined standards that promote interoperability for data, applications, and technology; and
- **Protection of Competitive Position:** Data that is viewed as compromising the competitive position of HEPs will still be collected in year, but will not be available publically, nor to other collectors (save for that data required for them to fulfil their requirements) until it is no longer viewed as commercially sensitive.

The data principles will be owned and governed by the new HE Data Governance Body.

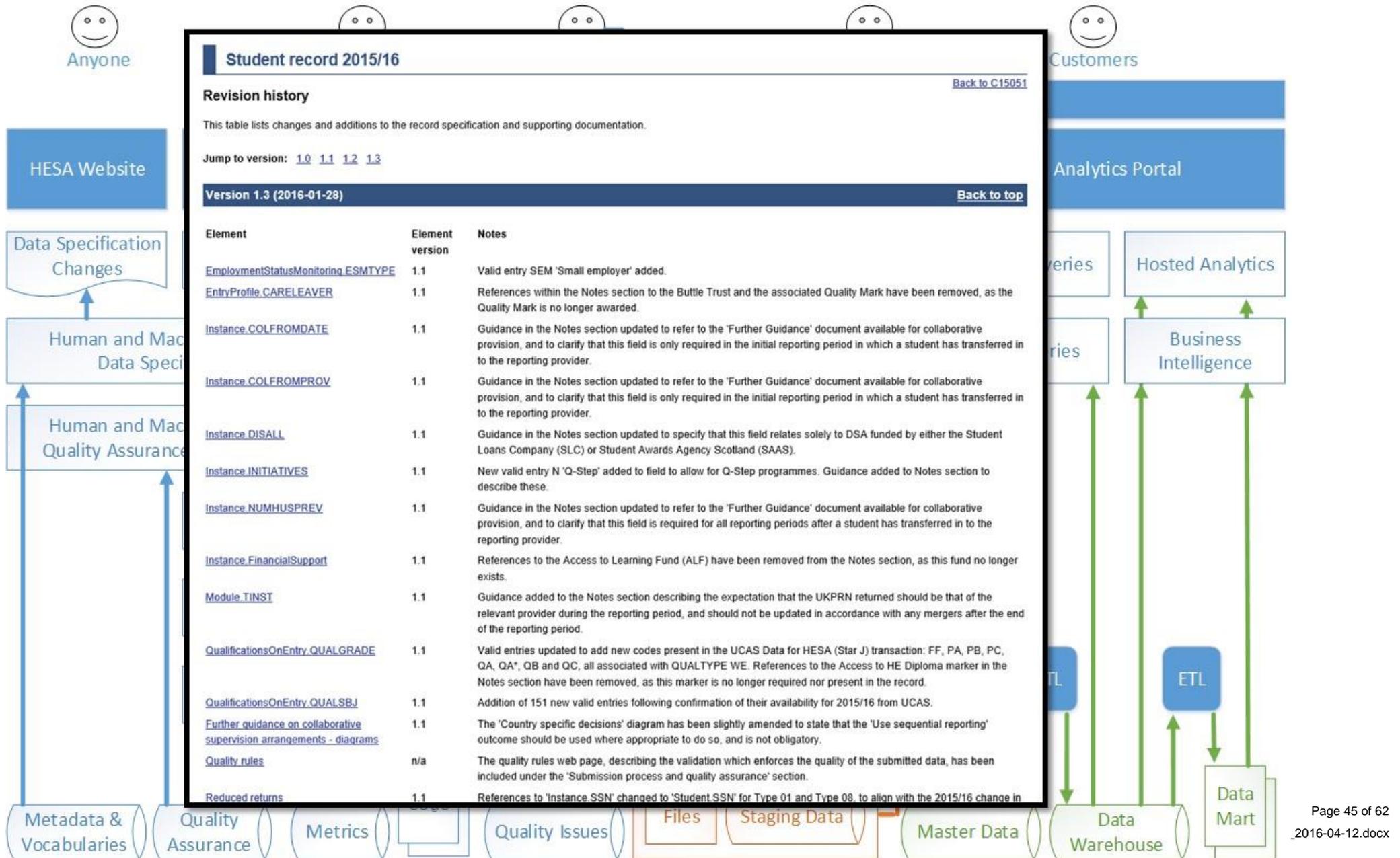
APPENDIX 3 – Views of the technical components

This section gives a logical view of the required technical components and gives examples of existing specifications, quality rules and reports etc.

Logical View



Data Specification Changes



Human Readable Specifications

Student record 2015/16

[Back to C15051](#)

All fields [return to field list](#) [valid entries](#)

Nationality

| Type | field | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--------------------------|--|------|-------|----|-------------|----|---------|----|---------|----|----------------|----|---------|----|--------|----|----------|----|---------------------|----|-----------|----|---------|----|-----------|----|---------|
| Short name | NATION | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Description | This field defines the country of legal nationality. | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Applicable to | England Northern Ireland Scotland Wales | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Coverage | Compulsory for all students where any Instance_REDUCED1 = 00 or 03 and where any Instance_COMDATE is on or after 2007-08-01. Optional but strongly encouraged for all other students. This field is optional for all students at providers in Northern Ireland. | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Valid entries and labels | Download the valid entries in csv format | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table border="1"> <thead> <tr> <th>Code</th> <th>Label</th> </tr> </thead> <tbody> <tr><td>AF</td><td>Afghanistan</td></tr> <tr><td>AL</td><td>Albania</td></tr> <tr><td>DZ</td><td>Algeria</td></tr> <tr><td>AS</td><td>American Samoa</td></tr> <tr><td>AD</td><td>Andorra</td></tr> <tr><td>AO</td><td>Angola</td></tr> <tr><td>AI</td><td>Anguilla</td></tr> <tr><td>AG</td><td>Antigua and Barbuda</td></tr> <tr><td>AR</td><td>Argentina</td></tr> <tr><td>AM</td><td>Armenia</td></tr> <tr><td>AU</td><td>Australia</td></tr> <tr><td>AT</td><td>Austria</td></tr> </tbody> </table> | Code | Label | AF | Afghanistan | AL | Albania | DZ | Algeria | AS | American Samoa | AD | Andorra | AO | Angola | AI | Anguilla | AG | Antigua and Barbuda | AR | Argentina | AM | Armenia | AU | Australia | AT | Austria |
| Code | Label | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AF | Afghanistan | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AL | Albania | | | | | | | | | | | | | | | | | | | | | | | | | | |
| DZ | Algeria | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AS | American Samoa | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AD | Andorra | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AO | Angola | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AI | Anguilla | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AG | Antigua and Barbuda | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AR | Argentina | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AM | Armenia | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AU | Australia | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AT | Austria | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Notes | <p>For students entering through UCAS this information will be available from UCAS via the complete new code list for entrants in 2009. Therefore in 2007/08 (and 2008/09) codes as provided in the *_J, however, providers can recode if they choose. The details by UCAS can be found in the *_J specification.</p> <p>Where a student has dual nationality including British, they should be coded as United Kingdom (UK) not including British, but including non-UK EU country then use relevant EU country code, otherwise if not including British, but including non-UK EU country then use relevant EU country code as either nationality.</p> <p>HESA Guidance</p> <p>This coding frame is determined by the National Statistics Country Classification 2006 (NSCC) 3166-1 Codes for the Representation of Names of Countries and their Subdivisions, Country or Area Codes for Statistical Use. In a small number of cases, the NSCC uses codes on which the NSCC is based. Where that happens, the ISO 3166 usage is given in brackets to facilitate cross-reference to the international standard. For example for code KR the description is Korea (South). The text in brackets, [Korea, Republic of], is the ISO 3166 description. Information in braces is provided by HESA to clarify usage, and is not derived either from NSCC or ISO 3166 lists of short country names, for example Spain (includes Ceuta, Melilla).</p> | | | | | | | | | | | | | | | | | | | | | | | | | | |

Where HESA publishes domicile and nationality data it does so using the National Statistics Country Classification 2006 (NSCC) 3166-1 Codes for the Representation of Names of Countries and their Subdivisions, Country or Area Codes for Statistical Use. HESA includes the ISO 3166-1 description (text in brackets) and additional information (text in braces) to ensure that labeling standards, both UK and international, are available only for cross-reference to the international standard. HESA does not insist that HEPs display these additional descriptions in their own output.

Codes for countries with no settled inhabitants are referred to in the guidance but omitted from the list, for example Antarctica.

United Kingdom

Codes for the United Kingdom are UK, XJ, XK: United Kingdom

Where domicile is known to be United Kingdom but nationality is not known, XK must be used. The code GB is invalid for domicile. For nationality the code GB is invalid for nationality. The codes XF, XG, XH, XI, XK are invalid for nationality. The code XJ is invalid for both domicile and nationality. The code UK is not on the NSCC list and is invalid for both domicile and nationality. Students with domicile in the United Kingdom should be coded GB (there are no separate nationality codes for British National (Overseas) and other nationalities).

Guernsey and Jersey (Channel Islands)

Codes for Guernsey and Jersey are not part of the United Kingdom and not part of the European Union. United Kingdom codes must not be used for domicile or nationality. The Bailiwicks of Guernsey (which includes the smaller islands of Alderney and Jersey) must be treated separately, and the codes GG and JE must be used for both domicile and nationality. The code XL may be used for either domicile or nationality, but only for student instances from 2007/08 onwards (already current in 2006/07, to avoid the need for re-coding).

Islands not part of the United Kingdom and not part of the European Union. United Kingdom codes must not be used for domicile or nationality. The code IM must be used for both domicile and nationality.

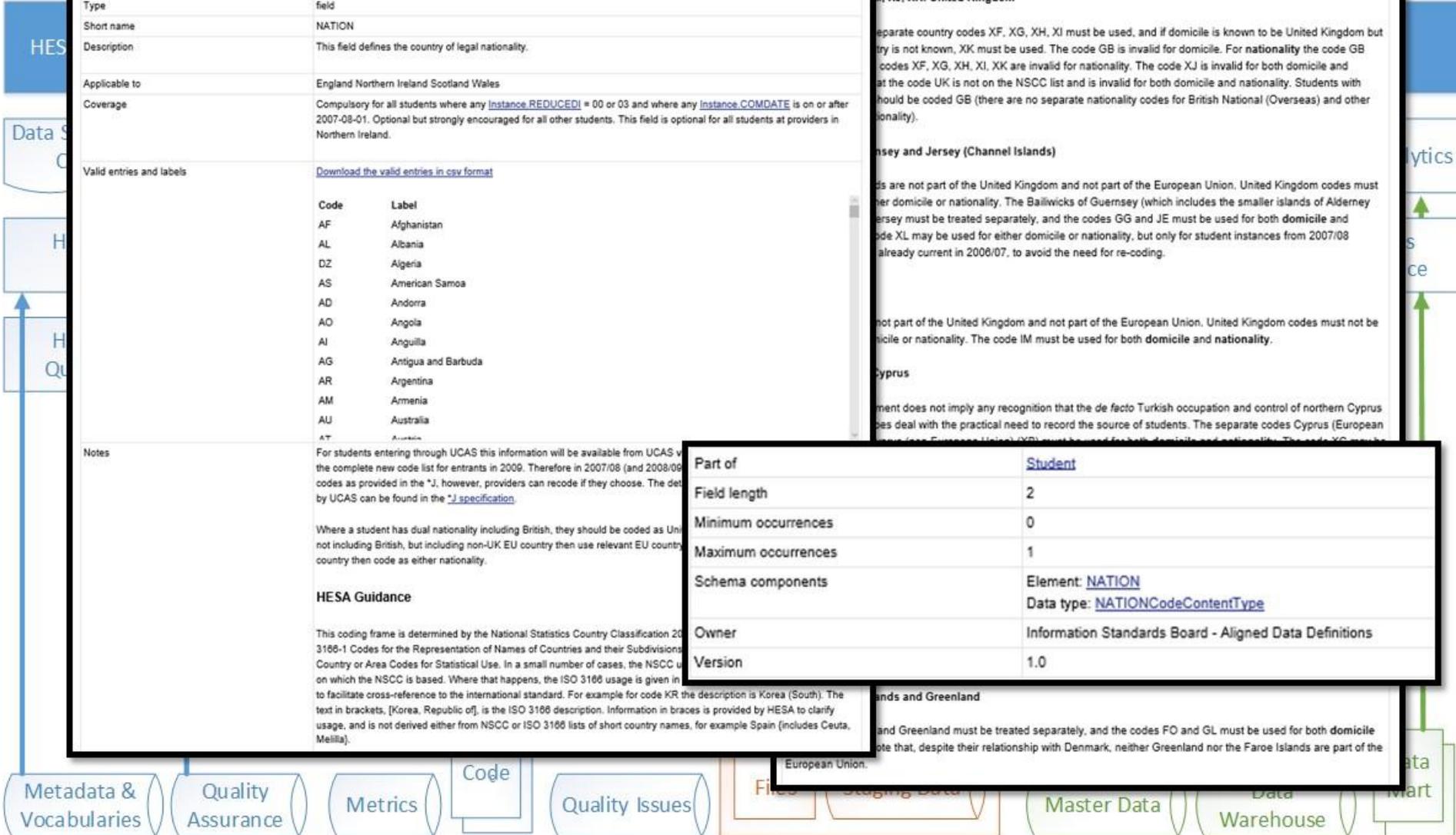
Cyprus

The separate code for Cyprus (European Union) is CY. The code IM must be used for both domicile and nationality. The code XK must be used for nationality. The code XJ must be used for domicile.

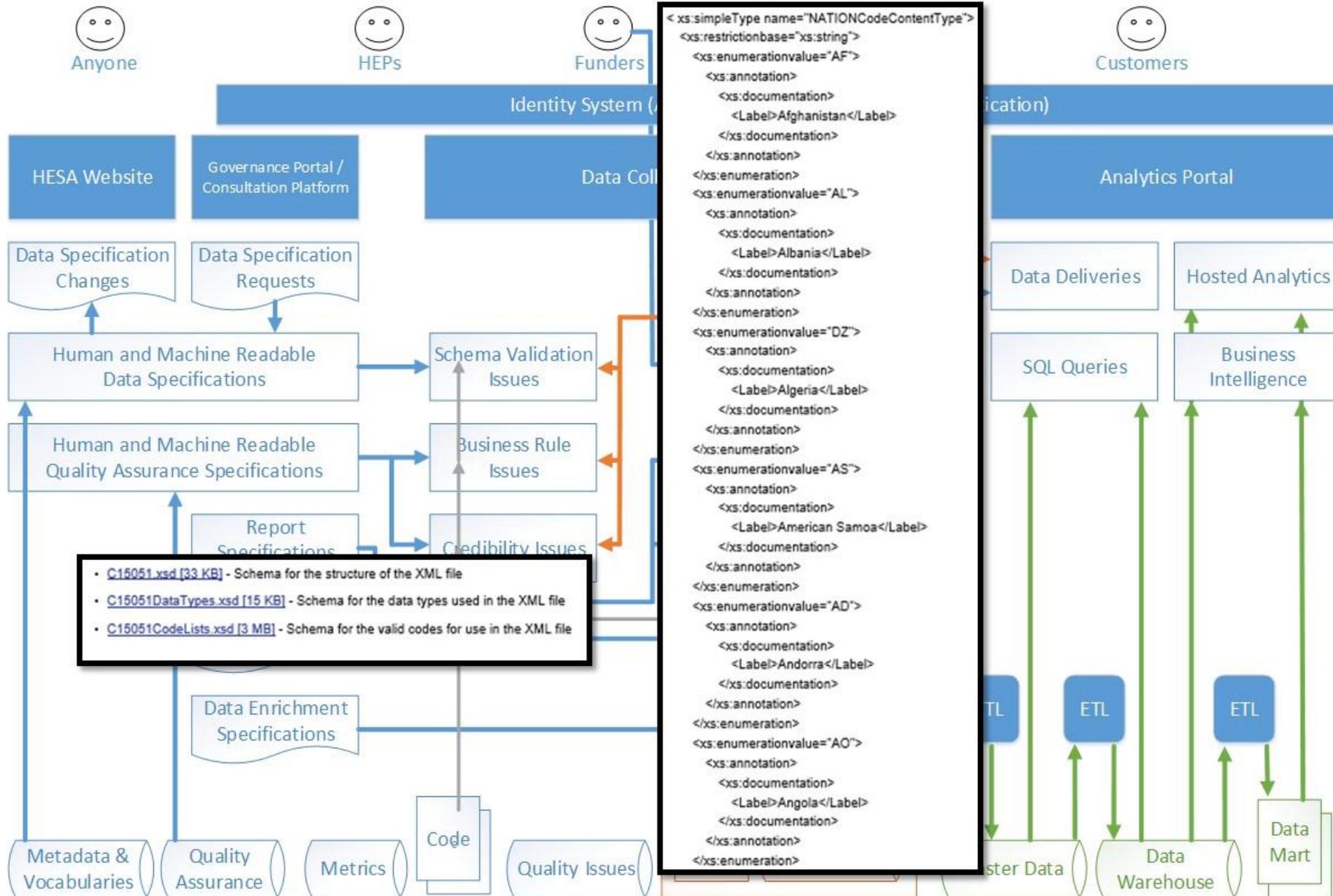
| | |
|---------------------|---|
| Part of | Student |
| Field length | 2 |
| Minimum occurrences | 0 |
| Maximum occurrences | 1 |
| Schema components | Element: NATION Data type: NATIONCodeContentType |
| Owner | Information Standards Board - Aligned Data Definitions |
| Version | 1.0 |

Faroe Islands and Greenland

Codes for the Faroe Islands and Greenland must be treated separately, and the codes FO and GL must be used for both domicile and nationality. Note that, despite their relationship with Denmark, neither Greenland nor the Faroe Islands are part of the European Union.



Machine Readable Specifications



```

<xs:simpleType name="NATIONCodeContentType">
  <xs:restriction base="xs:string">
    <xs:enumeration value="AF">
      <xs:annotation>
        <documentation>
          <Label>Afghanistan</Label>
        </documentation>
      </xs:annotation>
    </xs:enumeration>
    <xs:enumeration value="AL">
      <xs:annotation>
        <documentation>
          <Label>Albania</Label>
        </documentation>
      </xs:annotation>
    </xs:enumeration>
    <xs:enumeration value="DZ">
      <xs:annotation>
        <documentation>
          <Label>Algeria</Label>
        </documentation>
      </xs:annotation>
    </xs:enumeration>
    <xs:enumeration value="AS">
      <xs:annotation>
        <documentation>
          <Label>American Samoa</Label>
        </documentation>
      </xs:annotation>
    </xs:enumeration>
    <xs:enumeration value="AD">
      <xs:annotation>
        <documentation>
          <Label>Andorra</Label>
        </documentation>
      </xs:annotation>
    </xs:enumeration>
    <xs:enumeration value="AO">
      <xs:annotation>
        <documentation>
          <Label>Angola</Label>
        </documentation>
      </xs:annotation>
    </xs:enumeration>
  </xs:restriction>
</xs:simpleType>
  
```

Quality Assurance Specifications



Student record 2015/16

Quality rules

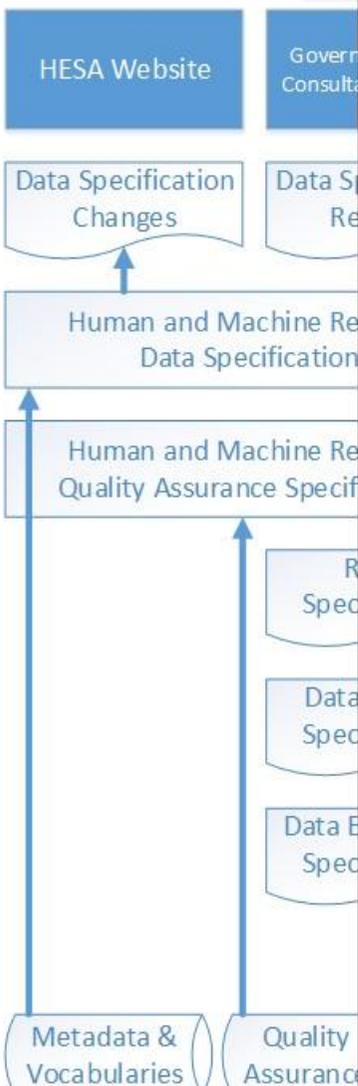
C15051.Student.NATION

| Rule ID | Rule description | Stage | Tolerance | Status |
|-----------------------------|---|---------------|-----------|-----------------|
| QR.C15051.Student.NATION.1 | Student.NATION must exist for institutions in England, Wales or Scotland where any Instance.REDUCEDI = 00 or 03 and Instance.COMDATE is on or after 2007-08-01. | Business Rule | Error | Carried Forward |
| QR.C15051.Student.NATION.2 | Student.NATION code XL must have Instance.COMDATE before 2007-08-01. | Business Rule | Error | Carried Forward |
| QR.C15051.Student.NATION.3 | Student.NATION code XC must have Instance.COMDATE before 2007-08-01 unless (EntryProfile entity exists and EntryProfile.UCASAPPID exists) or EntryProfile entity does not exist. | Business Rule | Error | Carried Forward |
| QR.C15051.Student.NATION.4 | Student.NATION code XC must have Instance.COMDATE before 2007-08-01 unless EntryProfile.UCASAPPID exists. | Exception | Error | Carried Forward |
| QR.C15051.Student.NATION.5 | Student.NATION is unknown (ZZ), but has a known value in the UCAS (*J) data. It is expected that a known value in the UCAS data will be passed through to the HESA return. | Exception | Warning | Carried Forward |
| QR.C15051.Student.NATIOND.1 | Student.NATIOND must exist for institutions in Wales where any (EntryProfile.DOMICILE exists and is coded XI) and the corresponding Instance.REDUCEDI = 00. | Business Rule | Error | Carried Forward |
| QR.C15051.Student.NATIOND.2 | Student.NATIOND must not exist for institutions in England, Northern Ireland or Scotland. | Business Rule | Error | Carried Forward |
| QR.C15051.Student.NATIOND.3 | Where there are two occurrences of Student.NATIOND they must not be identical. | Business Rule | Error | Carried Forward |
| QR.C15051.Student.NATIOND.4 | Where Student.NATIOND is coded R or U there must be only one occurrence of Student.NATIOND | Business Rule | Error | Carried Forward |
| QR.C15051.Student.NATIOND.5 | Student.NATIOND must exist where Welsh institution and any EntryProfile.DOMICILE (including those previously sent for the student) = XI and the corresponding Instance.REDUCEDI = 00. | Exception | Error | Carried Forward |

Showing 1 to 10 of 11 rows

<< < 1 2 > >>

[Download Excel copy of all rules](#)



Quality Assurance Code



Student record 2015/16

Quality rules

C15051.Student.NATION

| Rule ID | Rule description | Stage | Tolerance | Status |
|----------------------------|--|---------------|-----------|-----------------|
| QR.C15051.Student.NATION.1 | Student.NATION must exist for institutions in England, Wales or Scotland where any Instance.REDUCEDI = 00 or 03 and Instance.COMDATE is on or after 2007-08-01. | Business Rule | Error | Carried Forward |
| QR.C15051.Student.NATION.2 | Student.NATION code XL must have Instance.COMDATE before 2007-08-01. | Business Rule | Error | Carried Forward |
| QR.C15051.Student.NATION.3 | Student.NATION code XC must have Instance.COMDATE before 2007-08-01 unless (EntryProfile entity exists and EntryProfile.UCASAPPID exists) or EntryProfile entity does not exist. | Business Rule | Error | Carried Forward |
| QR.C15051.Student.NATION.4 | Student.NATION code XC must have Instance.COMDATE before 2007-08-01 unless EntryProfile.UCASAPPID exists. | Exception | Error | Carried Forward |

| | | | | | | | | | | |
|----------------------------|--|---------|-------|-------|-------|-----------------|-----------------|-----------------|-----------------|-----------------|
| QR.C15051.Student.NATION.4 | <?xml version="1.0" encoding="utf-8"?> <Rule xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="rule.xsd"> <Specification> <Id>QR.C15051.Student.NATION.4</Id> <Stage>Exception</Stage> <Status>Carried Forward</Status> <SpecificationDescription>Student.NATION code XC must have Instance.COMDATE before 2007-08-01 unless EntryProfile.UCASAPPID exists.</SpecificationDescription> <Tolerance>0</Tolerance> <StructuralRule>>false</StructuralRule> <PlainEnglishDescription></PlainEnglishDescription> <PreviousNameOfRule>QR.C14051.Student.NATION.4</PreviousNameOfRule> <AdditionalFieldToDisplay>Student.HUSID</AdditionalFieldToDisplay> <AdditionalFieldToDisplay>Student.OWNSTU</AdditionalFieldToDisplay> </Specification> <Implementation> <Context /> <Test> <Commands> <Select>select <f_fields> from t_<f_inst>;amal a WHERE a.f_nation = 'XC' AND a.f_comdate >= to_date('2007-08-01', 'YYYY-MM-DD') AND nvl(length(a.f_ucasappid), 0)=0</Select> </Commands> </Test> <AdditionalFieldToDisplay name="HUSID">a.f_husid</AdditionalFieldToDisplay> <AdditionalFieldToDisplay name="OWNSTU">a.f_ownstu</AdditionalFieldToDisplay> <AdditionalFieldToDisplay name="NATION">a.f_nation</AdditionalFieldToDisplay> <AdditionalFieldToDisplay name="COMDATE">a.f_comdate</AdditionalFieldToDisplay> <OrderBy>F_1</OrderBy> </Implementation> </Rule> | Warning | Error | Error | Error | Carried Forward |
|----------------------------|--|---------|-------|-------|-------|-----------------|-----------------|-----------------|-----------------|-----------------|

Showing 1 to 10 of 11

Download Excel copy

HESA Website

Govern
Consulta

Data Specification
Changes

Data S
Re

Human and Machine Re
Data Specification

Human and Machine Re
Quality Assurance Specif

R
Spec

Data
Spec

Data E
Spec

Metadata &
Vocabularies

Quality
Assurance

Analytics

ss
nce

Data
part

Report Specifications

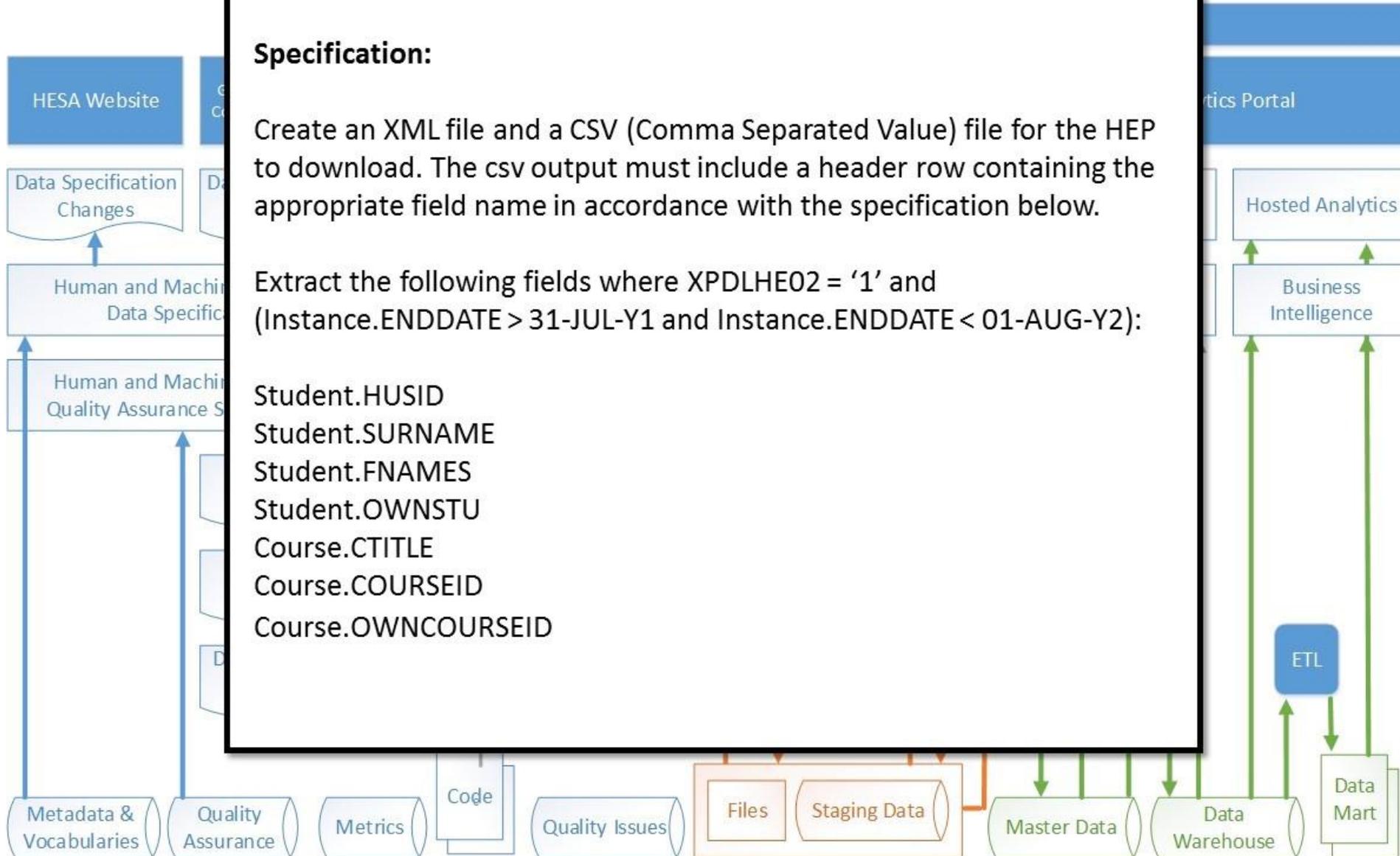


Specification:

Create an XML file and a CSV (Comma Separated Value) file for the HEP to download. The csv output must include a header row containing the appropriate field name in accordance with the specification below.

Extract the following fields where XPDLE02 = '1' and (Instance.ENDDATE > 31-JUL-Y1 and Instance.ENDDATE < 01-AUG-Y2):

- Student.HUSID
- Student.SURNAME
- Student.FNAMES
- Student.OWNSTU
- Course.CTITLE
- Course.COURSEID
- Course.OWNCOURSEID



Data Delivery Specifications

Data collection: Cyy051

Process/ Function: For Statutory Custome

Purpose: To define the structure of the file

Written by:

Revisions:

```
<DeliveryFile table="EntryProfile" content="Data" extension="Txt" label="EntryProfile entity">
```

```
<Fields>
```

```
<Field>EntryProfile.EntryProfileKEY</Field>
<Field>EntryProfile.InstanceKEY</Field>
<Field>Institution.UKPRN</Field>
<Field>EntryProfile.ACCESS</Field>
<Field>EntryProfile.ARTICLN</Field>
<Field>EntryProfile.CARELEAVER</Field>
<Field>EntryProfile.DOMICILE</Field>
<Field>EntryProfile.MARSTAT</Field>
<Field>EntryProfile.NIDEPEND</Field>
<Field>EntryProfile.PARED</Field>
<Field>EntryProfile.PGCECLSS</Field>
<Field>EntryProfile.PGCESBJ1</Field>
<Field>EntryProfile.PGCESBJ2</Field>
<Field>EntryProfile.PGCESBJ3</Field>
<Field>EntryProfile.POSTCODE</Field>
<Field>EntryProfile.PREINST</Field>
<Field>EntryProfile.QUALENT3</Field>
<Field>EntryProfile.RELIGION</Field>
<Field>EntryProfile.SEC</Field>
<Field>EntryProfile.SOC2000</Field>
<Field>EntryProfile.SOC2010</Field>
<Field>EntryProfile.UCASAPPID</Field>
<Field>EntryProfile.WELBACC</Field>
<Field>EntryProfile.YRLLINST</Field>
```

```
</Fields>
```

```
<Variants>
```

```
<Variant Id="SC" Subset="All"/>
<Variant Id="EFA" Subset="EFA">
```

```
<Restrictions>
```

```
<And>
```

```
<Condition>
```

```
<FieldName1>Derive.XINSTCOU01</FieldName1>
```

```
<Operator>=</Operator>
```

```
<Expression2>'E'</Expression2>
```

```
</Condition>
```

```
<Condition>
```

```
<FieldName1>Instance.FESTUMK</FieldName1>
```

```
<Operator>In</Operator>
```

```
<Expression2>('1', '3', '4')</Expression2>
```

```
</Variants>
```

H

Dat

Metadata & Vocabularies

Quality Assurance

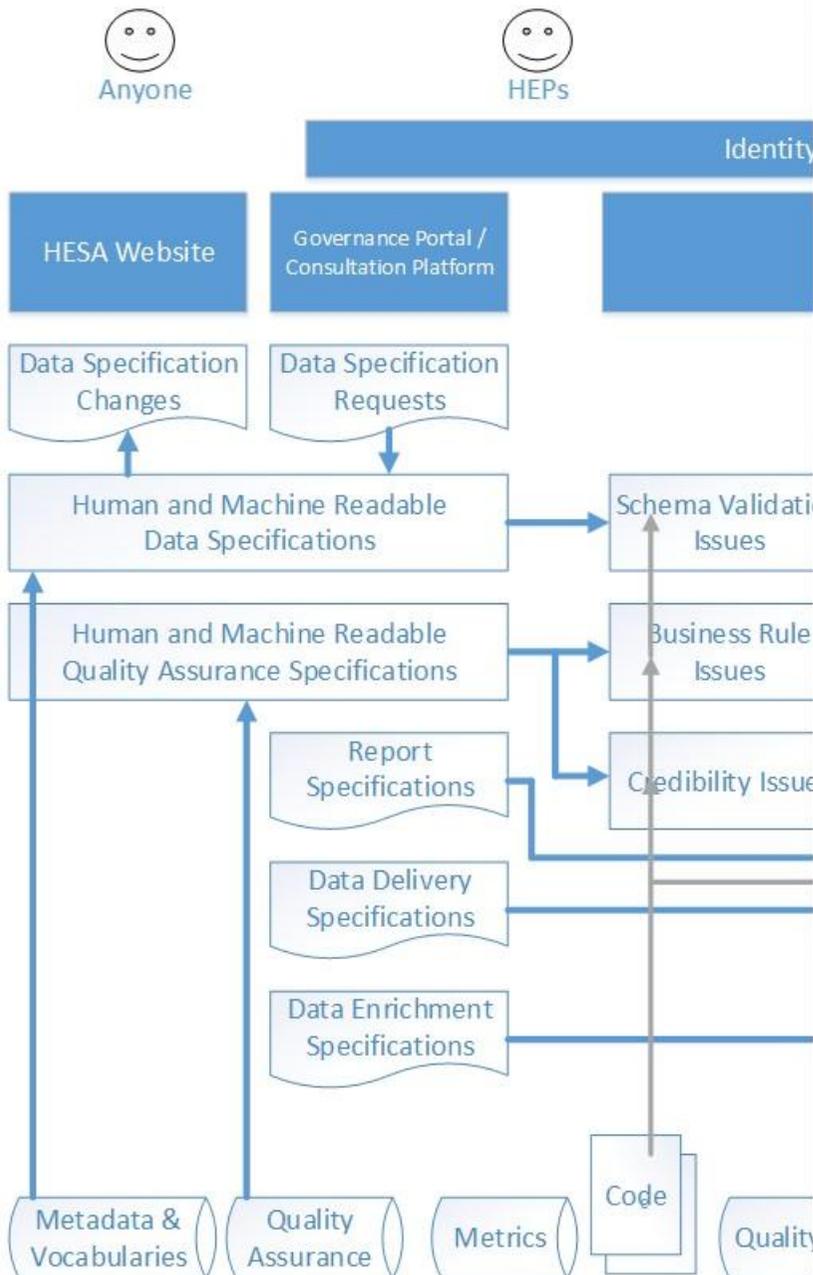
Metric

Analytics

Business Intelligence

Data Mart

Data Enrichment Specifications



| Field Description | Field Abbreviation | Field Length | Field Type | Field Version |
|---|---|--------------|-------------------|---------------|
| Level of qualification in DLHE - 3 way split | XDLEV301 | 1 | Char | 1.5.1 |
| VALID ENTRIES | | | | |
| 1 | Postgraduate | | | |
| 2 | Undergraduate | | | |
| \$ | Qualification not required for the Destinations of Leavers from Higher Education Survey | | | |
| - | No qualification awarded or no link to Student or AP student record (as applicable) | | | |
| ADDITIONAL INFORMATION | | | | |
| Postgraduate courses are those leading to higher degrees, diplomas and certificates (including Postgraduate Certificate in Education (PGCE at level M) and professional qualifications) which usually require a first degree as an entry qualification (i.e. already qualified at level H). | | | | |
| Undergraduate courses are programmes of study at level H, I, J and C including, but not limited to, first degrees (including eligibility to register to practice with a health or social care or veterinary statutory regulatory body/Architects Registration Board), first degrees with Qualified Teacher Status (QTS)/leading to Early Years Teacher Status (EYTS)/registration with a General Teaching Council (GTC), postgraduate bachelors degrees at level H, enhanced first degrees (including those leading towards obtaining eligibility to register to practice with a health or social care or veterinary statutory regulatory body), first degrees obtained concurrently with a diploma and intercalated first degrees, Professional Graduate Certificate in Education (PGCE at level H), foundation degrees, diplomas in HE with eligibility to register to practice with a health or social care regulatory body, Higher National Diploma (HND), Higher National Certificate (HNC), Diploma of Higher Education (DipHE), Certificate of Higher Education (CertHE), qualifications leading to Early Years Teacher Status (EYTS) only, foundation courses at HE level, National qualifications framework levels 4 and 5, post-degree diplomas and certificates at undergraduate level, professional qualifications at undergraduate level and other undergraduate diplomas and certificates including pre- and post-registration health and social care and veterinary courses and courses in Teaching in the Lifelong Learning Sector. Entrants to these programmes of study do not usually require an HE qualification. | | | | |
| Where an instance had no QualificationsAwarded entity associated with it, this field will be a single underscore. | | | | |
| DEPENDENT FIELDS | | | | |
| none | | | | |
| TECHNICAL SPECIFICATION | | | | |
| XOBTND01 (CHAR 3) | | | XDLEV301 (CHAR 1) | |
| D00, D01, E00, E40, E43, L00, L80, M00, M01, M02, M10, M11, M16, M40, M41, M42, M43, M44, M45, M50, M70, M71, M78, M79, M80, M86, M88 | | 1 | | |
| M22, M26, M28, H00, H11, H12, H16, H18, H22, H23, H41, H42, H43, H50, H60, H61, H62, H70, H71, H78, H79, H80, H81, H88, I00, I11, I12, I16, I60, I61, I70, I71, I73, I74, I79, I80, I81, J10, J16, J20, J26, J30, J41, J42, J43, J45, J80, C20, C30, C41, C42, C43, C77, C78, C80 | | 2 | | |
| \$\$\$ | | | \$ | |
| — | | | — | |

CS

Quality Issues



Anyone



HEPs



Funders



HESA



Customers

Identity System (Authorisation and Distributed Authentication)

Quality Rules

Errors 0 issues, 0 rules

Warnings 2900 issues, 23 rules

Switched Errors 1 issues, 1 rules

Filter by stage: All

Search

| Rule ID | Rule Description | Stage | Tolerance | Count |
|------------------------------------|---|---------------|-----------|-------|
| QR.C14051.StudentOnModule.MODOUT.8 | For institutions in England and Wales, more than 5% of student on module records coded as StudentOnModule.MODOUT = 6 where StudentOnModule.MODSTAT = 2 and Instance.MODE = 01, 02, 23, 24 or 25. | Exception | Warning | 1014 |
| QR.C14051.StudentOnModule.MODOUT.9 | For institutions in England and Wales, more than 15% of student on module records coded as StudentOnModule.MODOUT = 6 where StudentOnModule.MODSTAT = 2 and Instance.MODE = 31. | Exception | Warning | 753 |
| QR.C14051.Instance.NETFEE.6 | Instance.NETFEE = 0 where Instance.MODE = 31, 44 or 74 and Instance.SSN does not exist and (Instance.TYPEYR = 1 or (Instance.TYPEYR > 1 and Instance.ENDDATE is null)) and ((Instance.SPECFEE = 0, 1, 2, 4, 5, 6 or 9 or Instance.SPECFEE does not exist) and (Instance.MSTUFEE = 01-04, 06-26, 32-61, 96, 97 or 99) and (Course.MSFUND = 01-06, 11-29, 32-35, 38-91 and Instance.FUNDCODE = 2, 3, 5 or 7)). Please check it is correct that the net fee is zero for these instances. | Exception | Warning | 585 |
| QR.C14051.Instance.NETFEE.5 | Instance.NETFEE = 0 where Instance.MODE = 01, 02, 23, 24, 25, 43 or 73 and Instance.SSN does not exist and (Instance.TYPEYR = 1 or (Instance.TYPEYR > 1 and Instance.ENDDATE is null)) and ((Instance.SPECFEE = 0, 1, 2, 4, 5, 6 or 9 or Instance.SPECFEE does not exist) and (Instance.MSTUFEE = 01-04, 06-26, 32-61, 96, 97 or 99) and (Course.MSFUND = 01-06, 11-29, 32-35, 38-91 and Instance.FUNDCODE = 2, 3, 5 or 7)). Please check it is correct that the net fee is zero for these instances. | Exception | Warning | 177 |
| QR.C14051.Instance.TYPEYR.7 | Instance.TYPEYR should not normally be coded 1 where Course.COURSEAIM begins D or L. | Business Rule | Warning | 98 |
| QR.C14051.Instance.RFFData.3 | Where the Instance.RFFData entity occurs more than once, there are multiple occurrences of | Business | Warning | 75 |

Hosted Analytics

Business Intelligence

ETL

Data Mart

Metadata & Vocabularies

Quality Assurance

Metrics

Code

Quality Issues

Files

Staging Data

Master Data

Data Warehouse

Quality Issues



Anyone



HEPs



Funders



HESA



Customers

Identity System (Authorisation and Distributed Authentication)

HESA V

QR.C14051.StudentOnModule.MODOUT.8

For institutions in England and Wales, more than 5% of student on module records coded as StudentOnModule.MODOUT = 6 where StudentOnModule.MODSTAT = 2 and Instance.MODE = 01, 02, 23, 24 or 25.

[Coding manual](#)

| HUSID | NUMHUS | MODID | OWNSTU | OWNINST | COURSEID | OWNCOURSEID | MODOUT | MODSTAT | MODE |
|------------|--------|-----------------|------------|------------|----------|--|--------|---------|------|
| [REDACTED] | 03 | N-ASE-GP/A/Y | [REDACTED] | [REDACTED] | MSASEFTC | Astronautics and Space Engineering | 6 | 2 | 01 |
| [REDACTED] | 02 | I-MNU-A1031/A/Y | [REDACTED] | [REDACTED] | MSKMIFTC | Knowledge Management for Innovation | 6 | 2 | 01 |
| [REDACTED] | 02 | N-TPA-BC/A/Y | [REDACTED] | [REDACTED] | MSTPAFTC | Thermal Power | 6 | 2 | 01 |
| [REDACTED] | 03 | N-THP-CFDGT/A/Y | [REDACTED] | [REDACTED] | MSTHPFTC | Thermal Power | 6 | 2 | 01 |
| [REDACTED] | 02 | I-OOT-A1080/A/Y | [REDACTED] | [REDACTED] | MSOSEFTC | Offshore & Ocean Tech with Subsea Eng | 6 | 2 | 01 |
| [REDACTED] | 02 | M-L/THS/A/MT34 | [REDACTED] | [REDACTED] | MSLOSFTC | Logistics and Supply Chain Management | 6 | 2 | 01 |
| [REDACTED] | 02 | M-M/IP1/A/MT34 | [REDACTED] | [REDACTED] | MBFTMFTC | Full-time MBA Programme | 6 | 2 | 01 |
| [REDACTED] | 02 | M-M/LIA/A/MT34 | [REDACTED] | [REDACTED] | MBFTMFTC | Full-time MBA Programme | 6 | 2 | 01 |
| [REDACTED] | 02 | M-M/LIA/A/MT34 | [REDACTED] | [REDACTED] | MBFTMFTC | Full-time MBA Programme | 6 | 2 | 01 |
| [REDACTED] | 02 | I-OOT-A1084/A/Y | [REDACTED] | [REDACTED] | MSOSEFTC | Offshore & Ocean Tech with Subsea Eng | 6 | 2 | 01 |
| [REDACTED] | 02 | I-OOT-A1076/A/Y | [REDACTED] | [REDACTED] | MSOSEFTC | Offshore & Ocean Tech with Subsea Eng | 6 | 2 | 01 |
| [REDACTED] | 02 | N-ASD-AMS/A/Y | [REDACTED] | [REDACTED] | MSASDFTC | Aerospace Dynamics | 6 | 2 | 01 |
| [REDACTED] | 03 | N-ALS-AFEM/A/Y | [REDACTED] | [REDACTED] | MSALSFTC | Advanced Lightweight Structures and Impact | 6 | 2 | 01 |
| [REDACTED] | 03 | N-ALS-SM/A/Y | [REDACTED] | [REDACTED] | MSALSFTC | Advanced Lightweight Structures and Impact | 6 | 2 | 01 |
| [REDACTED] | 03 | N-ALS-GA/A/Y | [REDACTED] | [REDACTED] | MSALSFTC | Advanced Lightweight Structures and Impact | 6 | 2 | 01 |
| [REDACTED] | 03 | N-ALS-ACAS/A/Y | [REDACTED] | [REDACTED] | MSALSFTC | Advanced Lightweight Structures and Impact | 6 | 2 | 01 |
| [REDACTED] | 03 | N-ALS-ID/A/Y | [REDACTED] | [REDACTED] | MSALSFTC | Advanced Lightweight Structures and Impact | 6 | 2 | 01 |

Data Spe
Char

Hum

Hum
Qualit

ed Analytics

Business
Intelligence

ETL

Data
Mart

Metadata &
Vocabularies

Quality
Assurance

Metrics

Code

Quality Issues

Files

Staging Data

Master Data

Data
Warehouse

Continuity Issues



Anyone



HEPs



Funders



HESA



Customers

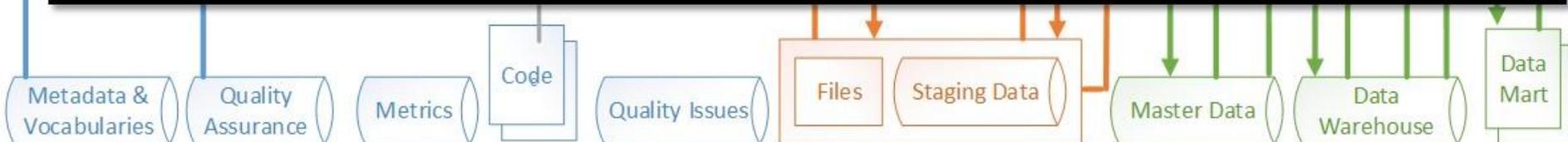
Identity System (Authorisation and Distributed Authentication)

Report A errors

| CheckID | Description | No. of records (00752201) | Percentage of errors (00752201) | Change in No. of records (00751692) | Change in percentage of errors (00751692) |
|---------|------------------------------|------------------------------|------------------------------------|--|--|
| 4.11.e | QR.C14051.Instance.PHDSUB.14 | 1 | 0.02% | +1 (0) | +0.02% (0.00%) |
| 4.50.e | QR.C14051.Student.SURNAME.4 | 3 | 0.07% | =0 (3) | =0.00% (0.07%) |

Report A warnings

| CheckID | Description | No. of records (00752201) | Percentage of warnings (00752201) | Change in No. of records (00751692) | Change in percentage of warnings (00751692) |
|---------|------------------------------------|------------------------------|--------------------------------------|--|--|
| 4.3.w | QR.C14051.Student.BIRTHDTE.11 | 3 | 0.07% | =0 (3) | =0.00% (0.07%) |
| 4.31.w | QR.C14051.Student.SURNAME.3 | 11 | 0.25% | =0 (11) | =0.00% (0.25%) |
| 4.34.w | QR.C14051.Student.ETHNIC.10 | 1 | 0.02% | =0 (1) | =0.00% (0.02%) |
| 4.35.w | QR.C14051.Student.SEXID.3 | 1 | 0.02% | =0 (1) | =0.00% (0.02%) |
| 4.6.w | QR.C14051.Student.SURNAME.2 | 18 | 0.40% | =0 (18) | =0.00% (0.40%) |
| 4.8.w | QR.C14051.EntryProfile.POSTCODE.14 | 1 | 0.02% | =0 (1) | =0.00% (0.02%) |
| 4.9.w | QR.C14051.Course.COURSEID.2 | 907 | 20.21% | =0 (907) | =0.00% (20.21%) |



Credibility Issues



Report Viewer Credibility Checking Student instance profile ▾

Credibility Checking

752201 (2015-10-29)

Student instance profile

This chapter contains information relating to the HE level student instance population and provides a useful starting point for quality assurance.

[View](#)

Subjects of study

This chapter contains information relating to the subject profile of HE level instances.

[View](#)

Qualifications on entry

This chapter contains information about the qualifications on entry held by new entrants to HE in the reporting period.

[View](#)

Student analytics

This chapter contains information relating to student analytics in the reporting period.

Mobility

This chapter contains information relating to mobilities which students take part in during the reporting period.

[View](#)

ITT

This chapter contains information relating to ITT students.

[View](#)

Qualifications awarded

This chapter contains information relating to qualifications awarded to students during the reporting period.

[View](#)

DLHE

This chapter contains information relating to DLHE data.

New data items

This chapter contains information relating to new data items.

FE students

This chapter contains information relating to FE students.

Study years

This chapter contains information relating to study years.

Average

This chapter contains information relating to average values.

Metadata & Vocabularies

Quality Assurance

Metrics

Quality Issues

Files

Staging Data

Master Data

Data Warehouse

Data Mart

Credibility Issues

Students by mode and level (SIN1)

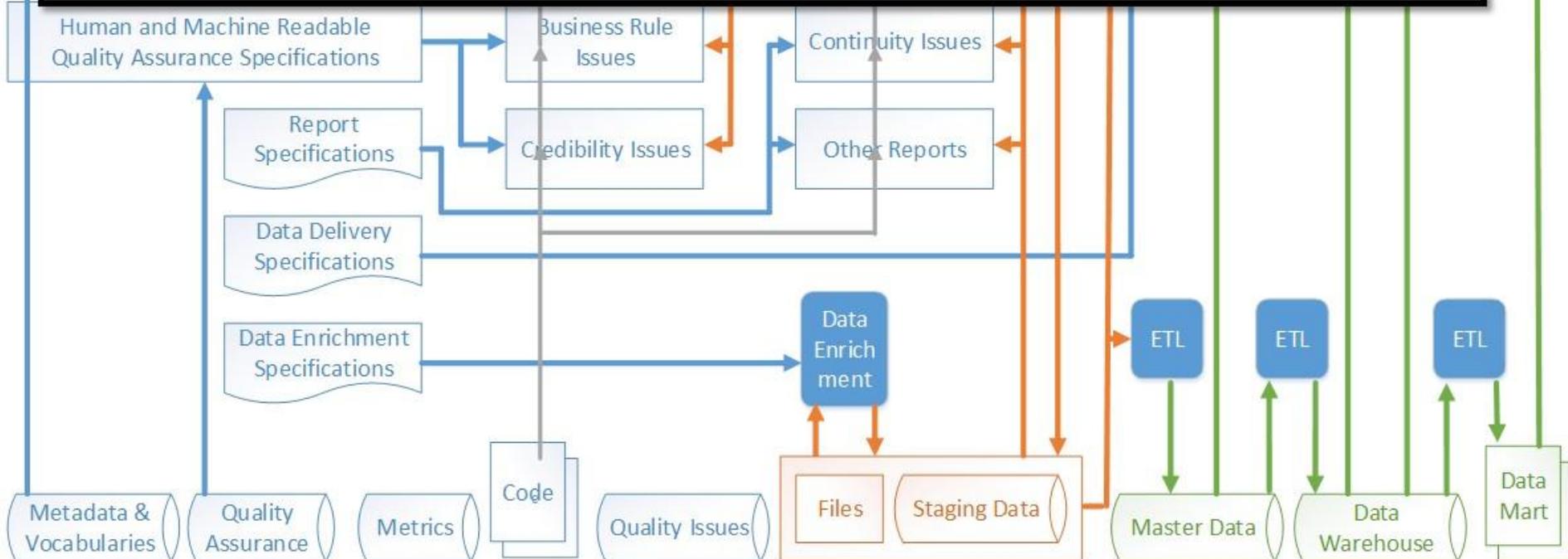
This table details student instance numbers by mode and level. Figures are provided for both the current and previous reporting years to provide a comparison along with percentage and total number differences.

Show Table Information



| | Full-time (including sandwich) | | | | | | Part-time | | | | | | Total |
|-------------|--------------------------------|-------|--------------|------------|-----|-------|-----------|-------|--------------|------------|-----|-------|-------|
| | PGR | PGT | First degree | Foundation | OUG | Total | PGR | PGT | First degree | Foundation | OUG | Total | |
| First years | 237 | 1,761 | 0 | 0 | 0 | 1,998 | 80 | 528 | 0 | 0 | 0 | 608 | 2,606 |
| Other years | 345 | 91 | 0 | 0 | 0 | 436 | 218 | 891 | 0 | 0 | 0 | 1,109 | 1,545 |
| Total | 582 | 1,852 | 0 | 0 | 0 | 2,434 | 298 | 1,419 | 0 | 0 | 0 | 1,717 | 4,151 |

Fields: XFYRSR01, XMODE301, XMODE01, Course.COURSEAIM, XLEV501, Population: XPSR01 = '1'



Credibility Issues

Students by mode and level (SIN1)

This table details student instance numbers by mode and level. Figures are provided for both the current and previous reporting years to provide a comparison along with percentage and total number differences.

Show Table Information

| | Full-time | | Part-time | | | | Total |
|--------------------|-----------|-------|----------------|----------------|------------------|------------------|-------|
| | PGR | PGT | PGR | | PGT | | |
| First years | 237 | 1,761 | 2014/15 237 | 2013/14 224 | 2014/15 1,761 | 2013/14 1,663 | 2,606 |
| Other years | 345 | 91 | NumDiff 13 | %Change 6% | NumDiff 98 | %Change 6% | 1,545 |
| Total | 582 | 1,852 | 2014/15 582 | 2013/14 575 | 2014/15 1,852 | 2013/14 1,675 | 4,151 |
| | | | NumDiff 7 | %Change 1% | NumDiff 177 | %Change 11% | 1,545 |

Human and Machine Readable
Quality Assurance Specifications

Report
Specification

Data Delivery
Specifications

Data Enrichment
Specifications

Total, Full-time (including sandwich), PGT 1,852 from 1,675 (177, 11%)

Metadata &
Vocabularies

Quality
Assurance

Metrics

Code

Quality Issues

Files Staging Data

Master Data

Data
Warehouse

Data
Mart

ETL

Analytics

Business
Intelligence

Credibility Issues

Students by mode and level (SIN1)

This table details student instance numbers by mode and level and total number differences.

Show Table Information

| | Full-time | | | First years | 2 |
|-------------|-----------|-------|---------|-------------|---|
| | PGR | PGT | First c | | |
| First years | 237 | 1,761 | | | |
| Other years | 345 | 91 | | | |
| Total | 582 | 1,852 | | | |

Fields: XFYRSR01, XMODE301, XMODE01

Human and Machine Readable Quality Assurance Specifications

Report Specification

Data Delivery Specifications

Data Enrichment Specifications

Metadata & Vocabularies

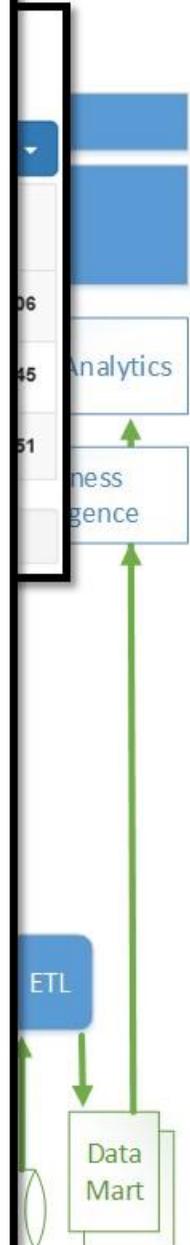
Quality Assurance

Metrics

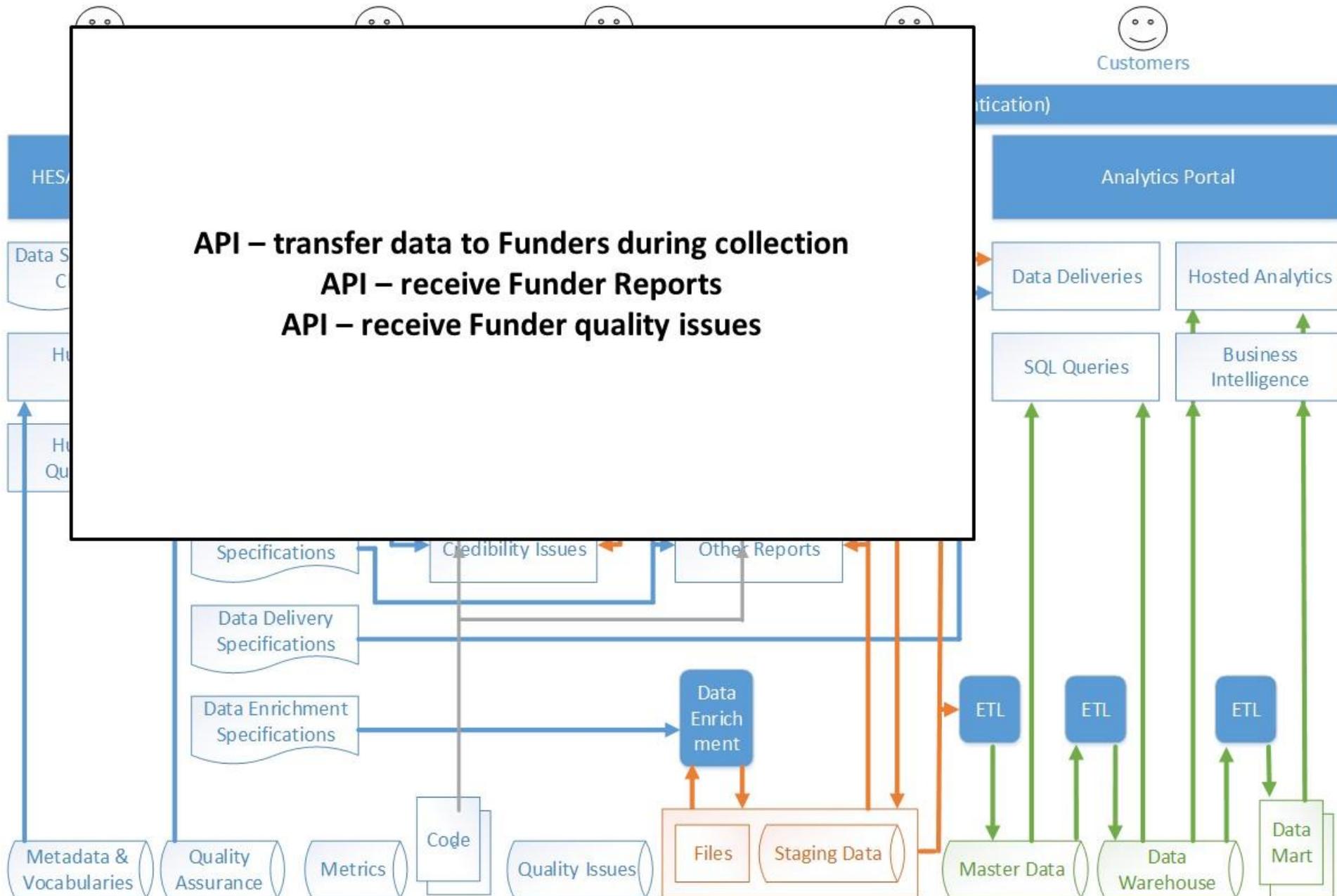
```

<Label code="C30" text="OUG" />
<Label code="C41" text="OUG" />
<Label code="C42" text="OUG" />
<Label code="C43" text="OUG" />
<Label code="C77" text="OUG" />
<Label code="C78" text="OUG" />
<Label code="C80" text="OUG" />
<Label code="C90" text="OUG" />
<Label code="C99" text="OUG" />
</Field>
</DataSource>
<CalculatedValue calculation="Difference">
  <Parameters>
    <Parameter>V1</Parameter>
    <Parameter>V2</Parameter>
  </Parameters>
  <ValueFormat>{0:#,##0;-#,##0;0}</ValueFormat>
</CalculatedValue>
<CalculatedValue calculation="PercentageChange">
  <Parameters>
    <Parameter>V1</Parameter>
    <Parameter>V2</Parameter>
  </Parameters>
  <ValueFormat>{0:#,##0%;-#,##0%;0}</ValueFormat>
</CalculatedValue>
<CalculatedValue calculation="PopulationCriteria">
  <Parameters>
    <Parameter>10000</Parameter>
    <Parameter>50</Parameter>
  </Parameters>
  <ValueFormat>{0:#,##0;-#,##0;0}</ValueFormat>
</CalculatedValue>
<ShadingRules>
  <RuleDefinition area="BodyColumns">
    <Or>
      <And>
        <Condition>
          <LeftPredicate type="Expression">V1</LeftPredicate>
          <Operator>Equal</Operator>
          <RightPredicate type="Expression">0</RightPredicate>
        </Condition>
        <Condition>
          <LeftPredicate type="Expression">ABS(V2)</LeftPredicate>
          <Operator>GreaterThan</Operator>
          <RightPredicate type="Expression">0</RightPredicate>
        </Condition>
      </And>
    </Or>
  </RuleDefinition>
</ShadingRules>

```



Funder Reports and Issues

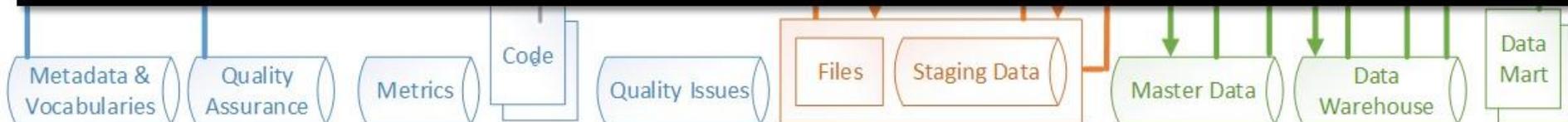


Other Reports



Latest Reports and Downloads

| Report Name | Description | Quality Issues | Warnings & Switched Errors | Issues Remaining | Downloads |
|---------------------------------------|--|----------------|----------------------------|------------------|--|
| Quality rules | Issues with your data | 2901 | 2901 | 0 | XLSX CSV |
| Quality rule compare | Select two Quality rule reports for comparison | - | - | - | - |
| Continuity | Year on year checks | 1058 | 942 | 116 | - |
| Entry Profile | Missing Entry Profile data | 0 | 0 | 0 | - |
| Credibility report | Management information tables for QA | - | - | - | - |
| Frequency Counts | Occurrences of valid entries | - | - | - | - |
| Check documentation | Management information for QA | - | - | - | XLSX |
| Checkdoc compare | Comparison with prior submission | - | - | - | XLSX |
| IRIS (HEFCE) | Funding Council output | - | - | - | ZIP |
| Data Supply | Raw data and derived fields | - | - | - | - |
| DLHE Population | DLHE Survey Population | - | - | - | - |
| NSS | National Student Survey population | - | - | - | - |
| Continuity population | Expected instance population for next year | - | - | - | - |
| UNISTATS | Data feeding into UNISTATS | - | - | - | - |



Analytics Portal

