

## Data futures

### Purpose

1. This paper provides an update on the HESA data futures programme following the discussion at the board meeting on 10 September 2020. The paper and seeks approval from the board to fund the Data Futures programme until 31 March 2022 and to set aside budget of £1.8 million in future years to support delivery of HESA Data Futures.

### Recommendations/Decisions required by the board

2. The board is invited to:
  - a. Note the progress of the programme since September at Annex A.
  - b. Agree funding in line with the board discussion on 10 September 2020 of £6,236,772 for the period until 31 March 2022 including
    - i. Core funding of £4,765,244
    - ii. Rolled forward underspend of up to £600,000
    - iii. Contingency of up to £871,528
  - c. Agree in principle a budget of £1,590,949 (including contingency of £250,000) to fund HESA and Jisc from April 2022 to March 2024.

### Issue

3. The board received an update on the HESA Data Futures programme at its meeting on 10 September 2020. That update included a detailed proposal covering the full data futures programme. HESA and Jisc have continued to work to deliver against the programme plan as set out in the proposal, copies of the September proposal can be made available to board members.
4. The board agreed funding of up to £2,582,117 to cover the period until the end of March 2021 and a conditional budget of £7,227,720 for the remainder of the programme. The board requested a further check in on programme delivery before committing this funding. This paper provides an update on the programme to date, with more detail in annex A, and requests confirmation of funding up to 31 March 2022.
5. The proposed funding relates to the investment in the technology uplift, core data collection capabilities, and the move to a new data model the frequency of collection which we plan to begin consulting on later this month, as set out in paragraph 10 below does not directly impact

these costs other than to the extent that where requirements diverge between nations this will increase the overall cost.

## Discussion and options

### Progress since September

6. Since September, HESA and Jisc have continued to deliver the data futures programme against the agreed project plan and specification. Like all organisations HESA and Jisc have faced issues due to the pandemic. In particular, recruitment has been more difficult than anticipated as experienced staff have tended not to enter the jobs market. The uncertainty over long term funding has also led to some retention issues and increased reliance on contractors. Slow recruitment and some lost productivity due to reduced resource through the pandemic means that the programme is behind the original plan although grant funding is largely tracking the delivery. The underspend of grant funding to March is projected to be around £540,000 representing around 21 per cent of total grant.
7. The delay in delivery does not create an immediate issue for the programme which uses an agile delivery approach and was designed to include slack although the ability to absorb further delays is now limited. HESA and Jisc are considering options to bring the programme back on track to ensure some contingency is retained in the programme. Options include standing up further development teams by utilising the current underspend and planned contingency. We are also discussing options with HESA to phase or delay implementation. While a delay to implementation is not desirable it can be mitigated by collection of aggregate data in the short term. In addition, a delay may be necessary due to the timing of the proposed consultation on in-year collection. We are also aware that provider readiness has been impacted by the pandemic meaning that providers may struggle to make systems changes to adapt to the new data model for 2022-23 irrespective of the frequency of collection.
8. The governance structures for the programme are now fully established and working well, the programme board meet monthly and are providing good challenge to the HESA and Jisc teams. OfS staff continue to engage in project governance at all levels ensuring that there is good visibility of decision making and objective reporting through to the programme board. The quarterly review group met for the first time in December 2020 and will meet again on 5 March 2021. A fuller update on progress of the programme is given in Annex A.
9. We have agreed with DfE that we will review the burden universal in-year collections would place on providers, this commitment is reflected in the DfE guidance letter of 14 September 2020. We plan to undertake this review in the autumn alongside consultation on the new strategy and data strategy. Conducting the review at this time adds uncertainty to the programme. However, it allows the review to be conducted in the context of the new OfS strategy and with the outcomes of our consultations on quality and standards, TEF, the strategic priority grant and NSS known. All of these activities are expected to use individualised data collected by HESA to some extent and will therefore have an impact on the burdens and benefits of the collection of in-year individualised student data.
10. While we cannot yet conduct the full burden review we believe that there is value in starting work with stakeholders to explore the options to reduce burden and highlight that we expect to need some in-year student data for some or all providers. To this end we intend to publish a working paper later this month that sets out the background to data futures, how in-year data is

likely to support our work, and options for meeting the needs for some in-year student data for some or all providers. We also intend to explore how we can reduce the burden of any individualised student data that are collected. We will also use this paper to remind providers about the changes that will happen to the HESA systems and the structure of the data collected to allow them to plan their work. We will use feedback from our working paper to inform a full consultation in the autumn alongside consultation on the new data strategy.

11. The burden review is focussed on England. Our current understanding is that the devolved administrations intend to require in-year individualised student data and plan to proceed with data futures as currently specified. If the outcome of the burden review is that in-year individualised student data is not required in England then there will be a need to replan work to account for the divergence and this is likely to significantly increase costs as HESA will need to replicate parts of their systems to deal with the different frequencies of collection for example, some quality rules rely on the period data covers. This increased complexity will not only impact data futures but will lead to ongoing increased costs at HESA. HESA are aware of this possibility and are actively seeking to minimise additional costs by ensuring that flexibility is built into the design and that the elements most impacted by any change are delivered later in the programme when there will be greater certainty.
12. Despite the uncertainty in the need for in-year individualised student data the investment in the technology uplift, core data collection capabilities, and the move to a new data model will continue to deliver value for providers through an improved submission process and a data model that better reflects the diversity of providers. Providing certainty to HESA on the funding of this core programme for a twelve-month period will enable them to recruit and retain staff many of whom are currently on short term contracts reflecting the funding uncertainty. In his guidance letter of 19 January 2021 the secretary of state advised that he was content for funding for the HESA Data Futures programme to continue although this is a decision for the OfS rather than government. The board is therefore invited to approve funding for the period to 31 March 2022 and agree in principle funding to March 2024 by which time the programme will enter business as usual.

Recommendation: The board is invited to:

- a. Note the progress of the programme since September at Annex A.
- b. Agree funding in line with the board discussion on 10 September 2020 of £6,236,772 for the period until 31 March 2022 including:
  - i. Core funding of £4,765,244
  - ii. Rolled forward underspend of up to £600,000
  - iii. Contingency of up to £871,528
- c. Agree in principle a budget of £1,590,949 (including contingency of £250,000) to fund HESA and Jisc from April 2022 to March 2024.

## **Paper publication date**

13. This paper can be published immediately.

**Richard Puttock, Director of data, foresight and analysis, [richard.puttock@officeforstudents.org.uk](mailto:richard.puttock@officeforstudents.org.uk), 9 March 2021**

# Annex A – Project delivery since September 2020

## Programme highlights for quarterly review group

1. The Data Futures programme has been impacted by recruitment and resourcing challenges arising from the pandemic, and the uncertainty associated with the burden review, with the pandemic and burden review also negatively impacting provider engagement.
2. Throughout the October 2020 to March 2021 delivery period, the Data Futures programme has re-ordered or staged work to minimise future rework associated with the burden review, and prioritised items delivering most value for Alpha<sup>1</sup> in light of capacity constraints and associated underspend (grant underspend for HESA and Jisc of £540,000 to £600,000 by the end of March). Some items have therefore been deferred to the next delivery increment.
3. Data Futures is on track to commence Alpha (with minor deferrals) in the timescale planned with provider training and preparations in April 2021. Careful management and resourcing will be required to deliver for the 2022-23 timeline.

## Delivery summary

4. Progress against the delivery roadmap for the October 2020 to February 2021 period has been reported on a monthly basis to the Data Futures Programme Board, and to the Quarterly Review Group in December 2020 and March 2021.

	Key Deliverables / Activities	Sep-20	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Apr-21	End Date	
<b>External</b>	Impacting Programme Direction	Burden Review announced						DF funding decision Burden review start			
<b>Programme</b>	Proposals & Governance	Proposal to OFS Board	Quarterly Review TOR		1st Quarterly Review			Confirm funding / ORG			
	Respond to Burden Review						Commence new mandate and options analysis				
	Design Decisions	XSD Generation Spec Publication (1)									
	Programme level deliverables	Grant letters to Sept 20	DPIA - first draft Ways of working with OFS	Grant letters to March 21 Move hosting to HESA	Framework Agreement			Alpha pen test DPIA Alpha Agreements			
	Recruitment & Resourcing		Team E Compliance role v1.5.0	PM Tech Lead		Programme Manager PM Transition Support					
<b>Data Workstream</b>	Data Model		v1.5.0	v1.6.0	No release	No release	v1.7.0	No release until Team E work completed			
	Throughout Programme	Guidance and minor model updates, derived fields / ECF and support QA, specification of 'lift and shift' rules, thematic review of Data Model, specification of credibility rules and reports, other QA reports and output specifications - throughout programme**								Throughout	
	Data deliverables	Review C19051/4 model changes and rules						Beta RP1 coding manual			
		Core Table & Output Spec			Alpha coding manual						
		Data Migration - Forward Mapping			2019/20 Data Migration				Data Migration Testing		May-21
		Team C - target 260 Rules Coded by end March	47	53	81	100	150		Continued Quality Rules		
<b>Technology Workstream</b>	Team A							Interface with HESA IMS System Notifications		Apr-21	
								Audit trail, logging and monitoring		Apr-21	
	Team B							Credibility Report **		Apr-21	
								Entry Profile Lookup **		Apr-21	
								Fix Pen Test Issues		Apr-21	
	Team D									Jan-22	
										Apr-21	
<b>Business Change</b>	Sector Engagement							Pre-Alpha Communications			
	Transition Planning										
	Preparation for Alpha									Apr-21	
<b>Note:</b> roadmap shows sequencing and concentrated efforts - some tasks may ramp up / ramp down and sprints may not align exactly to months; features will be prioritised by Product Owners in line with agile practice											
<b>Key:</b>		Due to complete post March	Impacted by burden review	Complete to plan (or late)	On Track	Moderate Risk	Significant Risk	Issue	Deferred		

5. The majority of the planned roadmap is on track to be delivered by the end of March.

<sup>1</sup> Alpha is the first testing phase of the system with users and will have 'skeleton' feature functionality to test the end to end solution including the ingestion of data, quality assurance and the production of output reports

6. A number of items originally planned for delivery prior to alpha have been delayed until after alpha or will be delivered during Alpha. In all cases these features are not essential to delivering the key objectives of alpha and will not adversely affect the programme.
  - System notifications
  - Audit trail, logging and monitoring
  - Data delivery to HESA's statutory customers
  - Issue Management
    - Approval of changes to quality tolerances
    - Filtering by rule across all issues.
    - A number of items relating to email notifications.
    - 'summary view' and the issues download report.
  - Look up data APIs (12 of 19)
  - Coding manual publication
  - Some planned quality rules and derived fields
7. The risk of delivering these alongside the rest of the roadmap is mitigated by:
  - a. Reprofile of underspend (up to £600,000 grant underspend for HESA and Jisc)
  - b. Potential use of contingency for contractor resource if permanent recruitment remains a challenge, and/or to increase Transition Team resource post April 2022
8. No requirement for any increase in funding from the figures set out in the September 2020 proposal has been identified currently, however some use of the contingency funds identified in the September 2020 proposal is likely to be required post March for:
  - Specification publication and guidance management (Sept 2020 proposal included £225,000 contingent costs for "Contingent Team E post Feb 21" to protect programme timescales, to be considered at key points during the programme)
  - Potentially additional contractor resource given recent turnover and the challenges recruiting at permanent rates in the current market
9. The September proposal identified several scenarios that would require additional funding. These scenarios are still valid and any of the following potential changes would need to be impact assessed and would require additional funding
  - Change to transition year
  - Change to collection approach (discrete to cumulative)
  - Any requirement to support annual and in-year collection concurrently

### **Delivered October 2020 to February 2021**

10. Delivery in this period includes:
  - Recruitment to address resource gaps and turnover, with over 30 positions recruited and over 10 positions requiring recruitment for the next phase
  - Programme governance fully established including the quarterly review group and monthly programme board, with a healthy level of input and challenge from attendees
  - Design governance continuing with OfS and UCISA invited to design authority meetings where key design decisions are made, the design principles approved, and the end-to-end solution demonstrated and reviewed in February with a focus on Alpha readiness

- Grant letters signed and Framework Agreement agreed
- Data workstream delivery includes Coding Manual 1.7.0 for Alpha (ready for release to the sector in early March), 1000+ rules with initial specification published, 200+ derived fields specified, 100 enhanced coding frames specified, 150 rules developed and 110 derived fields developed, output specification drafted, and data migration backward mapping complete
- Technology workstream delivery including continued development of the HESA Data Platform, the continued development of the Reference Data Store, the implementation of the XSD generation solution, and the ability to publish a sample guidance page as well as part of the data dictionary using the new approach for Coding Manual Publication (previously named Specification Publication and Guidance Management)
- Business Change workstream delivery includes continued sector engagement through Provider Forum and Software Suppliers workshops and updates to operational contacts, the development of a Data Futures E-Learning module which has been accessed by 100 users from 62 providers (21 per cent of providers), the definition of the approach to Alpha (adapted in response to provider pressures and feedback), and communications drafted to support the canvas for Alpha participants
- An end-to-end demonstration of the solution to the programme board in October 2020 with the relevant videos published on the HESA website (this will be repeated following the completion of this increment)

11. In addition, key members of the Data Futures Team have worked with OfS and the Devolved Administrations to explore options and implications in light of the burden review, and to provide information to support decision making for data collection in 2022-23.

### **Expected delivery by March 2021**

- Continued development of the data privacy impact assessments
- A number of pre-Alpha deliverables including environments, penetration test and Alpha Agreements
- Continued feature development for the HESA Data Platform, Reference Data Store, Issue Management and Coding Manual Publication
- Data Migration forward mapping and 2019/20 Data Migration

### **For alpha**

12. During Alpha providers and statutory customers will be able to:

- Work on the end-to-end solution.
- Engage with the data model, guidance and rules including automated checks, managing tolerances and raising issues.
- Give usability input for consideration in the development process, and enable bugs and errors to be identified.
- Test the model and provide user input on the rules and guidance to inform future work on rules and guidance.
- Support work on the approach to data migration that will support transition year quality assurance.

13. The features that will be delivered for the start of the Alpha pilot are:

- Submitting and support of multiple lifecycles
- Provider onboarding
- Approval and sign-off process
- Offline data entry tool

- Credibility reporting – 4 reports, 2 simple and 2 complex
- Quality Rules report
- Student continuity and on entry tables
- Regulation populations
- Data Delivery specified and delivered
- Monitoring of providers submissions
- Managing of tolerances and approvers
- System monitoring via process logging, health-checks, and log aggregation (excluding user interface)
- GDPR tasks can be performed
- Managing multiple specifications
- IDS and roles working as expected
- Integration with Issue Management
- Variety of quality rules and derived fields
- Transition Year QA using Migrated Data
- Entry Profile Look-up
- Snapshot views of historic data

## Underspend

14. The tables below provide the financial position as reported at the Programme Board in February. Data futures is being funded by a mix of grant funding from the OfS and the other UK funding bodies and HESA's own resources generated through their statutory subscriptions

October 2020 - March 2021							
	Resource - Grant (Incl OH)	Non-resource-Grant (Incl OH)	Total Grant Funded (Incl OH)	Resource-Subs (Incl OH)	Non-resource-Subs (Incl OH)	Total Subs Funded (Incl OH)	Total Budget (Incl OH)
HESA	1,479,369	0	1,479,369	402,043	371,861	773,904	2,253,273
Jisc	1,054,748	48,000	1,102,748	0	0	0	1,102,748
<b>Total</b>	<b>2,534,117</b>	<b>48,000</b>	<b>2,582,117</b>	<b>402,043</b>	<b>371,861</b>	<b>773,904</b>	<b>3,356,020</b>

Total Programme - Actuals plus Feb-Mar forecast								
	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Under Spend	Oct-Mar Total
<b>HESA Resource including Overhead</b>	214,892	226,472	234,328	263,909	263,595	259,077	392,957	1,855,230
<b>HESA Contingency</b>	-	-	-	-	-	-	26,182	26,182
<b>Jisc Resource including Overhead</b>	109,119	129,807	115,877	142,003	153,019	173,756	231,168	1,054,748
<b>HESA non-resource cost</b>	20,000	14,860	20,073	2,060	31,616	31,616	251,636	371,860
<b>Jisc non-resource cost</b>	3,978	3,748	3,787	6,018	3,986	3,986	22,496	48,000
<b>Total</b>	<b>347,989</b>	<b>374,887</b>	<b>374,066</b>	<b>413,989</b>	<b>452,216</b>	<b>468,435</b>	<b>924,438</b>	<b>3,356,020</b>
<b>Actuals</b>								
<b>Forecast</b>								

Dedicated Only - Actuals plus Feb-Mar forecast								
	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Under Spend	Oct-Mar Total
<b>HESA Resource including Overhead</b>	162,793	171,740	203,749	210,291	201,996	240,935	287,864	1,479,369
<b>HESA Contingency</b>	-	-	-	-	-	-	-	-
<b>Jisc Resource including Overhead</b>	109,119	129,807	115,877	142,003	153,019	173,756	231,167	1,054,747
<b>HESA non-resource cost</b>	-	-	-	-	-	-	-	-
<b>Jisc non-resource cost</b>	3,978	3,748	3,787	6,018	3,986	3,986	22,496	48,000
<b>Total</b>	<b>275,890</b>	<b>305,296</b>	<b>323,414</b>	<b>358,311</b>	<b>359,001</b>	<b>418,677</b>	<b>541,527</b>	<b>2,582,116</b>
<b>Actuals</b>								
<b>Forecast</b>								

## Programme RAG status

15. Overall, the programme is currently amber, the programme board's risk assessment is given below. The assessment of the programme board and HESA is a fair reflection of the status of the programme and aligns with the view of the OfS staff involved in monitoring the programme. The main risks facing the programme are uncertainty caused by the burden review and provider readiness.
16. The proposed review of universal in-year data collection as part of the burden review has had two significant impacts on the programme. The first is that this has led to the need to replan and re-prioritise work to ensure that the programme is agile and can respond to changing requirements. The second is that the uncertainty over the collection frequency and pressure in resources caused by the pandemic means that providers have either ceased or significantly reduced their planning for the introduction of data futures. We will use our working paper to provide clarity over those items that we intend to move forward with irrespective of the burden review which should allow providers more certainty to plan around. We, and HESA, are monitoring provider readiness and engagement with the programme including through the alpha piloting stage but recognise that changes to delivery dates may need to be made because providers (rather than HESA) are not able to respond.
17. Data Futures represents a significant technology investment and is not without risk as indicated by the previous phases of the project. However, the revised approach significantly reduces the risk and progress to date has been good with many of the most significant challenges already having been overcome and the core end-to-end system having been stood up. As the programme moves forward the risk will reduce as the work remaining will become increasingly well understood.



Key Aspects	RAG	Status
Overall	A	Funding agreed for October 2020 to March 2021 - constraint hinders permanent re... Decision required on data model and approach for 22/23 Good progress on delivery against roadmap despite lockdown and resourcing cha... Work being prioritised to deliver for Alpha and minimise potential future rework Replanning required post March 2021 following 22/23 decision
Scope	G	Roadmap to September 2020 delivered; Increment 1 signed off by from Programm... Future roadmap set out in September proposal - next increment completes end of... funding completes end of March Scope may be impacted by the outcome of the burden review (now expected late...
Costs	A	Funding agreed for October 2020 to March 2021 - further funding decision in Marc... Up to £600k underspend in the Oct 2020 to March 2021 period needs to be carri... Grant letter signature process underway
Schedule	A	Schedule may be impacted by the outcome of the burden review (now expected la...
Resources	R	Funding decision to March impacts recruitment and increases resource risk Increasing pressures due to staff turnover, vacancies and impact of Covid-19 lock... No significant contention with BAU currently
Stakeholders	G	Approach and communications for Alpha drafted for review by Programme Board... Continued engagement with Provider Forum and software suppliers Ongoing provider concerns related to bandwidth and burden review impact
Benefits	G	Updated in joint proposal for September; summary of benefits for Feb 21 Program...
Risks	R	Short term funding decision increases resourcing risk Outcome of burden review may increase risk depending on the level of change an... divergence of SC requirements Lack of provider engagement with DF whilst dealing with Covid and in light of the u... associated with the burden review Impact of Covid-19 on resource availability remains a risk, exacerbated by recent