

## Data Futures

### Issue

1. This paper seeks approval from the board to fund the Data Futures programme until 31 March 2021 and to set aside budget in future years to support delivery of HESA Data Futures.

### Recommendations

2. The board is invited to:
  - a. Note the progress of the programme since March.
  - b. Note the work on reviewing the burden on providers from in-year collection as part of Data Futures.
  - c. Note the importance of the data model and technology refresh for HESA.
  - d. Agree funding of £2,582,117 until 31 March 2021.
  - e. Agree in principle a budget of £7,227,720 (including contingency of £1,121,528) to fund HESA and Jisc from April 2021 to March 2024.

### Further information

3. Available from Richard Puttock ([richard.puttock@officeforstudents.org.uk](mailto:richard.puttock@officeforstudents.org.uk)), Director of Data, Foresight and Analysis.

## Background

4. The board received an update on the HESA Data Futures programme at its meeting on 28 January 2020 and a further update at its meeting on the 16 March 2020. The board noted the regulatory importance of in-year data relating to individual students in the longer term. The board agreed to release up to £1 million to Jisc to fund the core data collection capability, and up to £1.3 million to HESA for development of the data model, quality rules and other items OfS agrees are on the critical path during the period March – September 2020. This agreement was subject to prior approval for staged payments being delegated to a group comprising the OfS chair, a member of the Risk and Audit committee, the chief executive, the Director of Resources and Finance, and the Director of Data, Foresight and Analysis. Based on proposals submitted by HESA and Jisc, this group agreed funding of £725,389 for HESA and £874,000 to Jisc, with an additional £831,349 funded via HESA's statutory subscriptions.

## Discussion

### Progress since March

5. Since March, HESA and Jisc have issued 3 proposals for approval all focussed on the period to September 2020. These proposals represented a significant improvement in clarity over the December proposal and have better evidence underpinning the costs.
6. HESA have established new project governance structures with OfS staff represented at all levels. The programme is being overseen by a programme board including senior staff from HESA, Jisc, the programme funders and sector representatives. The new governance structure builds on the recommendations made by BDO, with clear accountability, more regular programme board meetings and significant improvements in the reporting to the board.
7. Recognising the importance of OfS oversight of the programme KPMG have conducted an internal audit of the governance arrangements which has led to a finding of significant assurance with minor improvement opportunities and three recommendations in total. A table of recommendations, proposed actions and completion dates from the internal audit is given in annex B. These were discussed with the Risk and Audit Committee at their meeting on 8 September.
8. Since the March board meeting HESA and Jisc have delivered the majority of items included within the three proposals despite the challenges presented by the pandemic in particular the difficulty in recruiting and on boarding staff. HESA and Jisc are working well together with integrated delivery teams and programme artefacts.
9. HESA and Jisc have delivered working software that allows data collections to be created, providers to submit data and the data to be delivered to statutory customers. Although significant work is still required, this approach has ensured that the software delivery has been significantly de-risked. In addition, HESA have delivered significant amounts of guidance, validation rules and other information published on their website to allow providers to confidently develop their systems. In the original project, these objectives were not met. Further details of project delivery since March are included at annex A.
10. While there have not been significant issues, HESA and Jisc have been transparent where issues have occurred, with significantly improved reporting to monthly programme boards. OfS

staff engaged in assurance activity continue to engage with HESA and Jisc staff at all levels and have not identified any issues that were not being escalated appropriately.

11. A few items included in the proposals have not been delivered. The delivery issues are driven by a failure to recruit staff or staff being diverted to HESA business as usual work due to pressures created by the pandemic, rather than issues with pace of the teams in post or quality of delivery. If left unchecked these delivery issues could create a long-term delivery risk for the programme. HESA and Jisc are aware of these issues and are mitigating them by creating a new team to ensure development stays on track and, where possible, backfilling specialist roles from within Jisc. Currently the project is underspent by around £389,000 and this is reflected in the grant payments being made to September. The cost of the additional team required to bring the project back on track is £225,000 which is included in the grant being now being sought to cover the period September 2020 to January 2023. Importantly, this should result in HESA and Jisc achieving everything as set out in the March to September proposals, excluding a small amount of security related work (roughly 2 weeks' work for a development team) with a combined underspend of around £164,000. In agreement with OfS, HESA have delivered some additional work beyond the agreed proposals where specialist resource had capacity. The table below shows proposed vs actual costs for the period January-September 2020:

Costing	Grant funding	HESA Subscription	Total cost	Notes
Proposals for March – September	£1,599,389	£831,349	£2,430,738	The board approved of £2,300,000 grant funding in March 2020 based on 6 months of 2020-21 financial year expenditure
Likely actuals for March – September	£1,100,448	£941,426	£2,041,874	This figure does not include cost of additional HESA team (which is approximately £225,000) – that cost is included in September proposal, or the cost of deferred security work

12. HESA senior management have started to demonstrate management challenge in their behaviours, actively looking to de-scope work where the business case for an area is weak. This, coupled with the decrease in expenditure gives OfS staff confidence that HESA senior management are not allowing the work to expand to fill the time available, are delivery focussed and are actively controlling cost.
13. Despite the good progress to date during this stage of the programme given the significant investment and previous issues it remains prudent for the OfS to continue with a high level of assurance activity.

### **Burden reduction**

14. We have agreed with DfE that we will review the burden universal in-year collections would place on providers if we were to continue with the current data futures model where all providers make termly returns, this commitment is reflected in the DfE guidance letter of 14 September 2020. We will want to test the extent to which this additional regulatory burden is

necessary, particularly for providers that do not represent significant regulatory risk and the review would be designed to ensure that data collection is proportionate. This might lead us to seek more frequent data from those providers that present significant regulatory risk, for example, where a provider has breached a condition relating to student outcomes we may wish to collect continuation data more frequently to identify evidence of necessary improvement. The review will also look at how data can be collected more quickly as the systems are improved.

15. We have agreed with DfE that we will complete the review of in-year student data collection by the end of October 2020. We therefore plan to approach the review through a targeted consultation exercise with senior managers and professional groups in the sector including utilising HESA's networks of practitioners. This approach will allow us to understand the nature of the burden placed on providers. The review will look at the three drivers for data burden through three lens':
  - a. Data required
  - b. Quality requirements, and
  - c. Timing and frequency of the data requirements
16. The current HESA student data collection requirement consists of up to 114 fields for each student. In 2018-19 this meant English providers returned 151 Million items of data for the OfS via the HESA Student Record. Of these, 35 million were module level data items and 10 million related to students' qualifications on entry. We will consult on our ability to remove these items from the model. For example, qualifications on entry information could be obtained through linking to other data sources, although it would require DfE support to obtain this data from the devolved administrations.
17. We recognise that it is not practical for providers to ensure that every data item for every student is accurate, although there is often a perception that this is our expectation. We will consult with providers to identify the most burdensome items to quality assure, and either identify alternative approaches to collection or to express our quality expectations more clearly so that providers can understand what is fit for purpose. It should be noted that two of the known most problematic areas for student data are qualifications on entry, which is mentioned above, and funding completion which we plan to review as part of our review of funding.
18. The board confirmed at its meeting on 16 March 2020 the importance of in-year individualised student data noting that a patchwork of solutions could increase burden and decrease utility. However, given the concerns expressed by providers over the last 6 months, it is right that we now explicitly test with providers whether the alternatives considered by the board in March represent a less burdensome approach. In considering the burden of comprehensive in-year individualised student data collection it will be important that this is considered in the light of possible changes to content and quality expectations. We will use the outcomes of this consultation to help inform a new data strategy which will be consulted on in early 2021.

### **Impact of the review**

19. Irrespective of decisions around the frequency of the collection of individualised student data, we believe that Data Futures is required to achieve two other requirements:

- a. A technology uplift for HESA.
- b. An improved data model reflecting the updated data requirements for the OfS.

Both of these reasons are covered in annex C and have been accepted by DfE as important work. The need for investment in the data collection infrastructure is separate to the outcomes of the burden reduction review.

20. The Data Futures proposal assumes that the OfS will require data three times a year and that this requirement is common across the UK. Any change to this requirement in England will not result in any decrease in the cost of this proposal as the costs in this proposal are largely an investment in capability; the additional costs to HESA of collecting data termly are the recurrent costs of running multiple data collections per year which will be met from providers' annual subscriptions. The possible divergence of requirement between OfS and devolved administrations, and the need for some rework will result in cost increase which are likely to fall to OfS or subscriptions for OfS registered providers as these costs will no longer be split across the UK. Any increase in cost is hard to quantify without knowing our exact requirements.
21. HESA and Jisc are aware of the potential changes to requirements in England and are considering this in their detailed planning to ensure that the costs of changing approach are minimised by scheduling work that reflects the frequency of collection later in the programme. Even if the OfS decides that it only requires data once per year it is unlikely that the devolved administrations will take the same view which would lead to overall increased costs as some project artefacts, such as validation, would be duplicated. There would be further increased costs if OfS required termly data from some providers but not others as HESA would need to build new systems to manage the changing frequency. Divergence on timeliness requirements also increases the risk of divergence on the collection of specific data items, increasing burden on HESA, and on the OfS for UK wide analysis such as TEF. Annex Q of the proposal deals with how HESA and Jisc will adapt to changing regulatory requirements and possible divergence between the requirements of their statutory customers.
22. It would theoretically be possible to pause the programme while further work was undertaken to establish the OfS requirements and confirm the requirements for the devolved administrations. However, such a pause would mean that project teams at HESA and Jisc would need to be disbanded. Experience in establishing teams since March has shown that there is considerable cost to forming teams and decreased productivity as they familiarise themselves with the project and each other. Therefore, any pause, even a short one, is not only likely to lead to more significant delays in delivery but will increase costs as the savings from pausing activity will be more than offset by resumption costs.
23. We are confident that, irrespective of the outcomes of the work on regulatory burden, the revised HESA platform and data model can be adapted to meet our needs. The data model has been designed to work irrespective of the length of the collection period, so limited change will be needed. The data platform has been designed so that it can ultimately support all of HESA's data collections so will be able to flex to any foreseeable new requirements for periodic data. The improved functionality and user experience from the new data platform and a model that reflects modern higher education should reduce the long-term burden on providers of making HESA returns, and decrease HESA running costs.

## Proposal

24. HESA and Jisc have estimated the total cost of this project as £15.6 million, this includes costs from January to September 2020. The full proposal, including a separate annex on how HESA and Jisc intend to manage possible changes to scope, is available here. Comparisons to previous costings are below, in comparing these costs the January proposals should be compared to the total expenditure line and it should be noted that once the proposal was not approved in January options 1 and 2 were no longer possible.

Costing	Grant funding	HESA Subscription	Contingency (subscription)	Contingency (grant)	Total budget
Option 1 and 2 (in year data 2021-22) January proposal	£10,621,441	£2,955,056	£0	£1,846,832	£15,423,329
Option 3, (in year data 2022-23) January proposal	£10,689,514	£3,666,164	£259,291	£1,913,082	£16,528,051
September proposal	£8,688,309	£3,646,748	£156,499	£1,121,528	£13,613,034
January to September 2020	£1,100,448	£941,426	£0	£0	£2,041,874
<b>Total proposed expenditure</b>	<b>£9,788,757</b>	<b>£4,588,174</b>	<b>£156,499</b>	<b>£1,121,528</b>	<b>£15,654,958</b>

25. Total proposed expenditure is £15.65 million, which is comparable to the equivalent option in the January proposal (£16.53 million) once the reduced contingency reflecting increased delivery confidence is taken in to account. However, these figures are not directly comparable. The December proposal moved several items to post go live as they were not essential for a minimum viable product and their inclusion would have significantly increased delivery risk. It was always intended that these features would be required to ensure that the system was operable. OfS staff agree these items will significantly improve long-term operability of the system and value to providers. The new proposal includes these features as it is now possible to deliver them within the programme timescales. The items now included are:

- a. Monitoring of validation rule usage.
- b. Integration with HESA's CRM system to facilitate provider sign off.
- c. Excel readable individual level outputs of a provider's submitted data with derived fields included.

26. OfS staff have engaged significantly with HESA to ensure the costings are evidenced and reasonable. This process has involved deep dives on a number of specific programme items where either the rationale for the item was unclear or the effort associated with it was out of line with OfS expectations. As a result of this engagement some of the resource requirements scaled back. This has allowed OfS staff to gain confidence that the proposal is delivery

focussed and represents good value for money. The detailed estimation and OfS access to underlying programme tracking tools means that it will be straightforward to measure progress.

27. HESA have engaged with and received feedback from the programme board on the proposal. This has resulted in a significantly more focused proposal with a clear understanding of deliverables and timescales.
28. HESA intend to use the same governance structures they have been utilising since March, but enhanced to include quarterly Tripartite group meetings, which an OfS board member will attend. OfS staff involved with the earlier programme iteration have noted that the clarity of materials provided to the programme board has improved significantly and the programme board is much more actively engaged in the programme.
29. Provision for contingency at 10% of costs is prudent and aligns with the PWC Value for Money Review points on contingency. In particular, the contingency allows for staff turnover and use of short-term contractor resource to cover recruitment and on boarding timeframes. The proposal includes a top-level tri-partite governance group comprising HESA, Jisc and OfS board members.
30. In addition to this, and included in the table above as contingency, a small amount of resource has been costed separately as contingent resource. This resource covers areas where there is less certainty over the requirements and has been separately identified to ensure that the programme is not over-resourced. The contingency does not allow for additional costs resulting from divergence in requirements between the UK nations as a result of the review. Once the implications of the review are known the board will be asked to provide additional grant to cover any additional costs.
31. The release of funding, including contingency, will be tied to the monthly reporting to the programme board and contingent on delivery against milestones with advice sought from the tri-partite group where necessary.
32. Despite the good progress to date there remain risks with the proposal which HESA and Jisc have recognised. These risks are being appropriately managed by HESA and Jisc and escalated to the programme board. There are a number of factors that mean the risk of failure is as low as is reasonably practicable including: good working relationship between HESA and Jisc, high levels of motivation and expertise demonstrated by both organisations, and decrease in delivery being linked to a larger than expected decrease in cost. There is also a high degree of transparency which gives good visibility of emerging issues.
33. HESA's functions as designated data body are funded through subscriptions from registered providers using powers under section 67 of HERA and enforced by condition G3. It would be possible to fully fund this activity rather than through grant payment from the OfS. Any grant payments made by the OfS will come from programme funds, that is funding that would otherwise be distributed to registered providers to support the teaching of students. The current proposal assumes a mix of grant and subscription funding that avoids a significant rise in the HESA subscription. The overall costs of the programme are invariant to the source of funding. An approach that includes significant grant funding is recommended as this allows OfS a greater role in monitoring progress with the programme and to cease, or scale back, funded activity if the programme encounters difficulties.

34. The grant funding for each financial year is given below. Given the uncertainty over OfS programme budgets over the coming years the board is asked to approve funding of £2,582,117 up until 31 March 2021 and agree in principle to funding of £7,227,720 up to March 2024. Funding in future years will be confirmed as part of the annual budgeting process once we get confirmation of our programme funding from DfE. Contingency has been assumed to fall proportionately in to the 2021-22 and 2022-23 financial years.

Period	Funding (contingency)
October 2020-31 March 2021	£2,582,117
1 April 2021-31 March 2022	£4,765,244 (£880,777)
1 April 2022-31 March 2023	£1,302,526 (£240,751)
1 April 2023 – 31 March 2024	£38,422

### Recommendation:

35. The board is invited to:

- a. Agree funding of £2,582,117 until 31 March 2021.
- b. Agree in principle a budget of £7,227,720 (including contingency of £1,121,528) to fund HESA and Jisc for the April 2021 to March 2024.

### Risk implications

36. The proposed review of universal in-year data collection as part of the review is likely to cause providers to cease planning for the introduction of data futures. While this risk will be manageable in the short term any prolonged period of uncertainty could significantly increase the risk that providers are not ready to provide data to the new data model in 2022-23 or in-year data from 2023-24.

37. Data Futures represents a significant technology investment and is not without risk as indicated by the previous phases of the project. The revised approach significantly reduces the risk as an end-to-end system will exist by the end of September 2020. While the system is still some way from being a minimum viable product it does address the core functionality required of the system. The risk is being further managed by much closer project monitoring by OfS staff including participation in all of the main programme meetings and full access to the programme backlog.

### Communications and engagement

38. Significant engagement will be required with providers during the review of data futures outlined in paragraphs 14 to 18.

## **Paper publication date**

39. This paper can be published immediately.

## Annex A: Project delivery until September

1. Progress against the roadmap has been shared with the Programme Board on a monthly basis, and most deliverables due by the end of September 2020 have been completed with the exception of the deferred items outlined below.
2. Delivery headlines include:
  - a. Programme setup, governance, reporting, ways of working, recruitment and key design decisions.
  - b. Key platform functionality including data storage framework and data delivery to statutory customers, reference data store and associated interfaces, schema generation and identity and access management updates.
  - c. Data model published and being updated monthly including updates to guidance, along with backward mapping.
  - d. Output specification files created and coded.
  - e. Significant proportion of quality rules specified and published, with a small number of rules coded in the data platform.
  - f. Derived fields specified and published with a small number coded to support end to end demonstration.
  - g. Knowledge share webinars, communications, training and transition plans.
3. A user of the HESA data platform, that is currently functioning with Data Futures data model v 1.2.0, would be able to complete operations including:
  - a. As a HESA user, create a specification version consisting of a schema version and ruleset version, view the configured rules, derived fields and metadata, view and configure a collection by assigning a specification version and managing collection state, access each provider's submission for a given collection via collection monitoring screen, and generate/regenerate a final delivery to a statutory customer.
  - b. As a Provider, submit to an open and assigned collection, see your data securely segregated and checked using the matching schema for the associated data model, note additional state checks performed on previously submitted data, see your data enriched with the appropriate primary and specialist regulation populations and derived fields, and with quality rules applied and results visualised in online report, download an enriched version of the submitted data in XML, and utilise a simple sign-off step to support other key features and end-to-end flow.
  - c. As a Statutory Customer, access an overview of statutory customer deliveries and associated metadata, and download a final delivery package containing a subset of the full delivery files aggregated in accordance with an initial delivery specification.
4. Additional work outside the planned roadmap has also been delivered including:
  - a. work on forward mapping which was planned for post September but commenced earlier than planned due to work on the backwards mapping and derived fields completing earlier than expected.

- b. work has been completed on a recommendation for the transition year including preparation for consultation with the sector, with timing to be agreed with Statutory Customers.
  - c. the Data Futures Data Quality Strategy has been defined.
  - d. Team A work on regulation populations (previously cohort) and tolerance management, to manage dependency.
  - e. Data platform support for required tolerance types (count, percentage, and range).
  - f. Implemented quality rules executed in accordance with 'applicable to' in the rule config (covers both primary and specialist).
  - g. Different default tolerances supported across both primary and specialist regulators per rule.
5. Deferred as agreed by the Programme Board:
- a. E-learning content which was agreed by the Programme Board as deferred to post September due to turnover in the Training team – this has been mitigated by publishing the Statutory Customer knowledge share webinars on the HESA website, and is anticipated for completion by the end of 2020.
  - b. Deliverables arising from the resource profile being impacted by delayed recruitment and the impact of Covid-19 including:
    - i. Team A (120 points out of 400 originally allocated, with 28 points of work pulled forward)
      - 1. Guidance management including revision history (60 points)
      - 2. Specification publication (60 points)
    - ii. Team D (35 Story Points out of approximately 230 Story Points)
      - 1. API for Spec Publication
      - 2. Security policies integrate with IDS
      - 3. Security event logging, audit logs, Cache for APIs
      - 4. Reliability performance testing of APIs
      - 5. Build production environment
      - 6. Support DPIA

## Annex B: Internal audit action plan

#	Priority	Recommendation	Response
1	Medium	<p><b>Lack of detail in the documented agreement on the nature and extent of OfS involvement in the programme.</b></p> <p>a) The principles guiding the nature of OfS assurance and oversight over the programme were agreed between OfS and HESA in the Ways of Working document. The document includes a commitment from HESA to ensure OfS “involvement in key design decisions”, as well as to tracking progress and benefits in line with “OfS requirements”.</p> <p>However, the definitions of which design decisions are deemed key, the exact level of involvement in them, or what OfS exact requirements for progress tracking are, have not been specified or agreed in writing.</p> <p>b) As part of the assurance arrangement the OfS Head of Data Infrastructure attends programme governance bodies (such as the Programme Board, the Data Futures Delivery Group, and the Design Authority) allowing them to stay appraised of the progress and raise concerns if necessary. The OfS Data Assurance Analyst attends weekly calls about various design decisions, as well as feeding into them via email or a Teams channel as needed.</p> <p>Both OfS representatives have been noted to consult with SMEs within OfS to ensure that they can sufficiently challenge assumptions or decisions made by the programme team. The OfS representatives are also able to escalate issues to the Director of Data, Foresight and Analysis as needed.</p> <p>However, there is no documented mapping of who from OfS should be consulted or involved, beyond the individuals mentioned above, in making decisions or dealing with programme issues. The criteria to consult or escalate are not defined, and are based on the professional judgement of those individuals.</p> <p><b>Risk</b></p> <p>a) There is a risk that that the views on the exact level of OfS involvement in the programme are or would become misaligned between HESA and OfS. This could lead to disagreements and</p>	<p><b>Recommendation 1a –Reject recommendation / accept the risk, but with review</b></p> <p>OfS visibility of all programme decisions is good due to current involvement at all levels of the programme. All decisions, beyond the trivial, are reported to the programme board on which OfS has a clear challenge role with escalation to the quarterly group (see recommendation 2). The number and nature of decisions is large on a project of this scale so cannot be perfectly specified other than in crude terms around expected impact on time or budget. Therefore, adding further specificity to the decision making is likely to constrain rather than enhance OfS ability to challenge and escalate decisions. If OfS engagement in the project is reduced the risk level will increase and further specificity is likely to be necessary.</p> <p><b>Responsibility:</b></p> <p>Richard Puttock</p> <p><b>Timescale:</b></p> <p>Review if OfS engagement reduces</p> <p><b>Recommendation 1b –Partially accept recommendation</b></p> <p>The involvement of staff beyond the immediate assurance team is necessary to ensure that gaps in the assurance team’s knowledge do not expose the programme to risks. OfS is not the main risk owner for the programme, this lies with HESA. Therefore, the risk is that HESA do not identify risk and that OfS does not identify this failing. There is a related risk that the assurance team are not able to secure input from staff with expertise.</p>

		<p>deterioration of the relationship, ultimately impacting the successful delivery of the programme.</p> <p>b) Additionally, there is a risk that relevant OfS stakeholders would not be appropriately consulted, which could lead to the end product delivered by the programme not meeting the OfS requirements in full.</p> <p><b>Recommendation</b></p> <p>a) The exact nature of OfS involvement in the programme should be elaborated on. This should include agreeing the specific requirements that HESA should meet for tracking progress and analysing benefits. It should also be specified in which design decisions OfS will be involved in and in what way. This should be documented and approved by HESA and OfS.</p> <p>b) OfS should pro-actively identify areas where and from whom additional input might be needed, on top of the day to day activities of individuals directly involved in the oversight process.</p>	<p>Mitigations:</p> <ul style="list-style-type: none"> <li>-The assurance team should be risk averse when identifying the need for additional internal input that is to say if they have concerns about gaps in their expertise input should be sought</li> <li>-Commitments should be secured from internal teams to providing resource to support the assurance team.</li> </ul> <p><b>Responsibility:</b></p> <p>Richard Puttock</p> <p><b>Timescale:</b></p> <p>31 October 2020</p>
2	Medium	<p><b>Lack of formal Board level oversight beyond September 2020.</b></p> <p>Since April 2020, and leading up to September 2020, additional assurance (on top of day to day operations outlined in recommendation 3) was obtained through the scrutiny applied to the HESA delivery proposals submitted to the OfS board for grant funding approval. The grant funding requested in the final proposal (due in September) is to take the programme up to the technical programme go-live in 2022 and through finalising the transition into using the new data platform by 2024. Hence, this assurance will no longer be provided beyond September 2020.</p> <p>It is important that the assurance model continues to involve a senior management/board level challenge which is not directly involved in the day-to-day details of programme delivery. We noted, that the programme governance structure proposed in April 2020 included a “Quarterly Progress &amp; Relationship Review and Ad-hoc Escalation Route” meeting which could serve this role beyond September 2020. However, at the time of this audit, this part of the governance model had not been setup yet.</p> <p><b>Risk</b></p>	<p><b>Recommendation 2 –Accept recommendation</b></p> <p>The quarterly review group should be established and should meet within 3 months of the next phase of the project commencing</p> <p><b>Responsibility:</b></p> <p>Richard Puttock</p> <p><b>Timescale:</b></p> <p>22 December 2020</p>

		<p>Individuals involved in the day to day oversight of the programme will naturally become very close to the details of programme delivery and might not be able to focus on assurance over the “bigger picture” (i.e. whether the programme is progressing in line with senior management/board expectations). There is a risk that without an established regular senior management/board level challenge, any potential high level key delivery issues would not be identified. The risk could potentially increase over time, as the longer the assurers are involved in the programme, the more likely their views are to become aligned with those from the programme team.</p> <p><b>Recommendation</b></p> <p>We recommend that the “Quarterly Progress &amp; Relationship Review and Ad-hoc Escalation Route” is implemented promptly, with a formal Terms of Reference specifying its role agreed and documented. The meeting should be attended by Board Level executives from the relevant parties, and have the power to stop and/or request a re-baseline of the programme if needed.</p> <p>For the meeting to be effective, the attendees should receive periodic targeted report updates that would allow them to understand if sufficient progress is being made.</p>	
3	Low	<p><b>Individuals involved in the programme oversight becoming “single points of success”.</b></p> <p>The day to day OfS oversight over the programme is carried out by the Head of Data Infrastructure and the Data Assurance Analyst. They both share programme oversight responsibilities with their regular Business-As-Usual (BAU) roles, and we were advised that their programme oversight were roles designed to take 0.4 and 0.5 FTE time respectively.</p> <p>However, we noted that their current programme workloads are more time consuming than originally intended. This does not appear to have caused any major challenge so far, both because some of the increased workload is temporary (as it relates to working with HESA on the delivery proposal due in for September) and because their BAU workload is currently lighter than usual. The BAU workload appears to be reduced due to Covid-19, and this reduction is not likely to be permanent.</p>	<p><b>Recommendation 3 –Reject recommendation / accept the risk</b></p> <p>Data Futures is a complex project covering a wide range of activities in order to have good oversight and provide effective assurance very high levels of engagement are needed comparable to the levels of engagement from the existing team. Both members of the assurance team have 3 month notice periods which helps mitigate this risk. The reduction in risk from inclusion of additional staff is not proportionate to the cost.</p> <p><b>Responsibility:</b> N/A</p> <p><b>Timescale:</b> N/A</p>

	<p>Both staff members are well placed to carry out their programme oversight responsibilities due to their background, expertise and having varied networks within OfS, allowing them identify relevant SMEs to consult with.</p> <p>However, they are becoming increasingly crucial to the effective operation of the programme oversight, to the point where they are close to becoming the “single points of success” –i.e. where they would be difficult to replace. It is recognised that having two individuals involved in the day to day assurance, already provides potential cover for unavailable staff.</p> <p><b>Risk</b></p> <p>There is a risk that, should one of the staff members involved leave OfS or become unavailable long term, the quality of assurance provided over the programme would diminish. The risk of individuals becoming unavailable on long term sick leave is heightened due to the Covid-19 pandemic.</p> <p>Furthermore, there is a risk that if the staff members BAU or programme workloads increase, they would be unable to fulfil their roles effectively.</p> <p><b>Recommendation</b></p> <p>We recommend that consideration is given to introducing another staff member to the programme oversight model so that there is potential cover for unavailable staff involved in the process and greater flexibility for the individuals to cover each other’s responsibilities. It is recognised, that the benefits of introducing a new staff member would mostly be realised only if one of the staff members became unavailable.</p> <p>Addressing the recommendations 1a and 1b, outlined on the previous pages, would also reduce the risk outlined above, as increasing the level of documentation would facilitate any potential staff replacement.</p>	
--	--	--

## Annex C: Need for technology uplift and new data model

### Technology uplift

1. HESA have not invested in their current data collection systems significantly since 2015 reflecting the intention for a major refresh as part of Data Futures. The current data platform is built on outdated technology using primarily on-premise computing. This outdated design means that the system is hard to maintain, has high running costs and delivers a poor user experience as providers need to queue to have their submissions processed.
2. The current data collection systems has evolved over a number of years meaning that many of the operations required were not designed in to it, leading to large amounts of manual effort and high costs transitioning from one collection to the next. The levels of manual intervention are significant even with annual data returns and introduce a high likelihood of human error.
3. The current system includes a large range of both automatic and manual checks on providers' data to quality assure them. The manual checks would be hard to automate on the current platform. The use of manual checks not only creates burden on HESA but adds delay for providers as even once submissions have been processed they need to wait in a further queue to be manually quality assured, this can, during peak times, mean providers waiting up to a day for complete feedback on their submissions.

### Data model

4. HESA currently collects individualised student data in one of two different forms. This is historical and relates to a time when data required from providers that were then considered alternative providers (APs) was not required to be directly comparable to other providers. In addition, we currently collect individualised student data from FECs using the ESFA's individualised learner record (ILR). This means that there are three data models. Consolidating the 2 HESA returns would decrease the burden on HESA and improve consistency in outputs across providers. We are not proposing at this point to further consolidate the ILR return into the HESA return as this would significantly increase burden on FECs. However, this creates a risk that the data is less comparable and a significant risk as requirements evolve. The OfS has less control over definitions due to the need to work within the constraints of the ILR. For example protected characteristics are not fully captured within the ILR and ESFA are currently not inclined to add them.
5. In addition, the current student data model struggles to adequately reflect flexible provision. It was originally developed in the 1990's, when the HE sector, and the data requirements, were very different. Many of the structures assume students are studying structured courses on a traditional September to June pattern which is no longer true for large groups of students.
6. The data model developed for AP's deals with flexible provision better. However, we have learned some significant lessons since its introduction. In particular, the flexibility afforded in the model means that it is not always possible to ensure providers' data can be compared. In addition, it was only designed for DfE regulation of APs, rather than the more diverse regulatory requirements of the OfS.
7. The revised Data Futures model is designed to be a hybrid evolution of the two existing data models taking the best elements of each model to provide something that is fit for OfS regulation across a wide range of providers.
8. Data Futures has been designed with extensibility in mind. Providers currently need to make data returns to a large number of other organisations such as professional, statutory

and regulatory bodies (PSRBs). The primary reasons why PSRBs do not use current HESA returns are the timing (the data is too lagged) and, while the record collects much of what they require, it does not collect everything. The Data Futures data model is designed so that it will be straightforward to extend the model to collect additional information required by PSRBs. This will allow PSRBs to use Data Futures data and support the OfS' long term aim of decreasing the overall data burden on the sector by working with other bodies that collect data from providers.

9. It should be noted the aim of decreasing burden through wider use of HESA data is a long term one. It was considered central to the original Data Futures approach, leading to significant increases in the complexity of the model to ensure that all users' requirements could be captured. [Exempt from publication]. The intention now is to build a system which supports the core regulatory uses of the data and then extend it once success has been demonstrated. This has led to a Data Futures data model that is more focussed and closely resembles what has gone before meaning that while there will be work for providers in moving to the new data model this is significantly lower than the initial approach.