



RESPONSE TO DCMS
CONSULTATION
“DATA: A
NEW DIRECTION”

HESA

A horizontal bar composed of several colored segments: blue, green, grey, and dark blue.

Dan Cook
Deputy Director, Data & Innovation
HESA
95 Promenade
Cheltenham
GL50 1HZ
E dan.cook@hesa.ac.uk
W www.hesa.ac.uk

Higher Education Statistics Agency Limited is a company limited by guarantee, registered in England at 95 Promenade, Cheltenham, GL50 1HZ. Registered No. 02766993. Registered Charity No. 1039709. Certified to ISO 27001. The members are Universities UK and GuildHE.

DATA REFORM CONSULTATION 2021

CONTENTS

Foreword	4
Introduction	7
Chapter 1 – reducing barriers to responsible innovation.....	9
Research Purposes	9
Further processing	12
AI and Machine Learning.....	15
Data minimisation and anonymisation	16
Data intermediaries	16
Chapter 2 – Reducing burdens on businesses and delivering better outcomes for people	20
Privacy management programmes	20
Data protection officers.....	20
Data protection impact assessments	21
Prior consultation.....	21
Record keeping	21
Reporting thresholds for breaches.....	22
Voluntary undertakings.....	22
Further questions	22
Alternative scenario proposals.....	22
Subject access requests.....	23
Privacy and electronic communications	23
Chapter 3 - Boosting trade and reducing barriers to data flows	26
Chapter 4 - Delivering better public services	27
Digital economy act	27
Reliance on lawful grounds of public bodies	28
Transparency mechanisms for algorithms	28
Public safety and national security.....	30
Chapter 5 - Reform of the Information Commissioners’ Office.....	30

FOREWORD

The Government's proposals are wide-ranging and provide a useful opportunity for public debate and proactive consideration of the UK's approach to data protection and data governance legal and regulatory frameworks.

The UK's higher education (HE) sector is the focus of numerous and invaluable national resources – in research capabilities across a very wide spectrum of expertise, in education and training of a highly skilled workforce, as a highly successful export industry that also acts as a focus for the UK's global reputation and presence, and as the principal anchor for much knowledge- and innovation-led local and regional development. The HE sector is an especially intensive producer and user of data across the full breadth of its operations, and it maintains a sophisticated UK-wide data, digital, and technology infrastructure on which much of this activity depends. This infrastructure includes the HE sector's own shared data centres and high-speed network, connecting a collaborative network that stretches around the globe. Shared digital services support discovery, productivity, innovation, and efficiency. This infrastructure includes HESA, which operates the range of data collection, management, and dissemination activities that have been established as lawful, fair, and transparent methods for serving the information needs of all those with a stake in HE.

Most endeavours in the HE sector are undertaken in the public interest, principally by charitable bodies. Working within universities and colleges are more than 2.6 million students and 250,000 academics striving for understanding, insight, and progress. Higher education is its people, and they are collectively seeking answers to some of humanity's most intractable problems: disease, poverty, conflict, and the climate crisis, to name a few. The good reputation and achievements of our HE sector rest ultimately on the scientific, professional, and ethical commitments of these students, researchers, teachers, and other professional staff. This sounds rather grand, but we recognise that:

*'reputations rest – every day – on decisions made by individuals: the examiner marking the 'finals' paper; the admissions officer taking that difficult telephone call; the researcher deciding if a necessary adjustment to an investigative protocol has ethical implications; and so on.'*¹

Commitments to establishing the truth, being both relentless and careful in telling the truth, avoiding harm, keeping promises, and acting with respect for persons are all part of the 'psychological contract' that sustains HE. Questions of fairness and legitimacy, transparency and integrity necessarily dominate the HE sector's approach to collecting, managing, and using data, across the full range of its activity.

HESA's goal is to maintain as far as possible the trust of students and staff in our custodianship of their data, to deliver valuable insights to them and to the wider stakeholders in HE. HESA therefore welcomes the Government's commitment to putting the protection of personal data at the heart of future plans, and to maintain and strengthen public trust in responsible data use. We believe that these commitments are also closely aligned with a commitment to scientific integrity, and in this response, we explore some of the benefits and risks that the Government's proposals may present to the HE sector and wider scientific community.

¹ Watson, D. (2009). The question of morale | managing happiness and unhappiness in university life. Open University Press. <http://site.ebrary.com/id/10394831> p.122

HESA is a producer of official statistics, and in addition to data protection legal frameworks, we are also required by law to operate in accordance with the Code of Practice for Statistics. The three pillars of the code of practice are Trust, Quality, and Value. As statisticians we recognise the mutually reinforcing effects that integrity, professionalism, and a focus on the interests of both data subjects and other end users have on the data economy. We are also aware of the risks to data quality and usefulness that can be introduced by non-responses and other forms of bias that occur in environments with lower trust. We recognise that the Government's proposals, while extensive, are still in development. We welcome the wide-ranging consideration being given to data governance issues, and we are pleased to have an opportunity to input at this formative stage.

The Open Data Institute (ODI) theory of change² is a useful and respected tool for understanding how we can create impact from data. The ODI identifies a 'farmland scenario' as the optimised conditions where there is high public trust in the data economy, and hence widespread willingness to share data. The ODI compare this on the one hand to a 'data oil field scenario' where data is hoarded by organisations and shared insufficiently, requiring greater innovation, capability and infrastructure to create value; and on the other hand, to a 'data wasteland scenario' where low levels of trust require a focus on ethics, equity, and engagement to promote confidence in data collection and use. The ODI's work is not cited in the consultation, but the Government's proposals are consistent with a diagnosis that we are currently in 'the oil field' scenario, where data is hoarded, and innovation, capability, and infrastructure are needed to establish 'a new direction' and 'a bold new data regime'. However, the consultation mainly focusses on 'build[ing] on the current regime – aspects of which remain unnecessarily complex and vague'. The substance of the consultation appears mainly to address these issues by adapting the existing approach: this could be understood as meaning that the Government believes we are currently closer to the ODI's optimal 'farmland scenario'. The dissonance in language leaves the reader somewhat unsure as to intentions – is what is planned a bold change, or an enhancement of the current rules? Some additional reflection on how rhetoric relates to substance would be advantageous in helping understand intentions.

Clarity over the current situation and future direction is important in maintaining public trust and confidence. We therefore recommend the Government develops and publishes a clearer and more systematic evidence base (including on public perceptions of trust in data collectors and users) to inform diagnosis of the current situation and potential options. Given that there is still a perception of a deficit of public trust in data collectors despite the introduction of GDPR³, there is a possibility that we are in fact closer to the ODI's 'data wasteland scenario' where the public are more fearful of the misuse of data, and hence are less likely to share information.

Various data scandals have been internationally prominent during the past few years and may resonate in the public's consciousness. A 2020 report by the Government's Centre for Data Ethics and Innovation (CDEI)⁴ identified that the 'sharing of personal data must be conducted in a way that is trustworthy, aligned with society's values and people's expectations. Public consent is crucial to the long term sustainability of data sharing activity.' CDEI go on to note that public trust is 'tenuous' and that 'there has been relatively limited effort by the Government and wider public sector to address public trust explicitly'. Subsequent polling and other work has been modest and

² <https://theodi.org/about-the-odi/our-vision-and-manifesto/our-theory-of-change>

³ Bauer, P. C., Gerdon, F., Keusch, F., Kreuter, F., & Vannette, D. (2021). Did the GDPR increase trust in data collectors? Evidence from observational and experimental data. *Information, Communication & Society*, 0(0), 1–21. <https://doi.org/10.1080/1369118X.2021.1927138>

⁴ <https://www.gov.uk/government/publications/cdei-publishes-its-first-report-on-public-sector-data-sharing/addressing-trust-in-public-sector-data-use>

fairly operational in scope. Work on public attitudes to data collection and use more broadly would be valuable to underpin any more substantial change of direction. This is important, because research by Frontier Economics shows that ‘there is robust quantified evidence that greater trust consistently leads to increased data sharing.’⁵ We believe that there are several areas where the current arrangements could be adapted to provide greater benefit and clearer understanding, but that the case for a bolder or more sweeping set of changes has not yet been established. We also reflect on some of the nascent possibilities that ‘data intermediaries’⁶ or ‘data institutions’⁷ can play in supporting public confidence as well as responsible innovation in the data economy, and thereby support many of the benefits the Government aims to achieve.

In most respects we follow the reasoning of the Information Commissioner’s Officer (ICO) in their response to this consultation⁸, and HESA’s corporate position is closely aligned with the Information Commissioner’s thoughtful and careful response. Our experience is that strong regulation and enforcement acts as a platform for responsible innovation and development. Extensive effort has been put into implementing and explaining the GDPR, and people are now familiar with it. The UK’s GDPR has been a success story, and in the three years since its inception it has led to increased focus on data protection and greater clarity of thinking around the ethics of data management. In our sector, ‘thinking with GDPR’ has improved both the effectiveness and safety of large-scale data processing systems.’⁹

The current legal and regulatory framework broadly works well for HESA and the HE sector. It has had the effect of promoting high standards and responsible behaviour, and alignment with regulatory standards in place elsewhere supports research collaboration across borders. Working through GDPR’s implications has helped clarify and simplify HESA’s role as the hub for information about HE, and the central role that HESA’s data sharing agreements (underpinned by productive stakeholder relationships and professional independence and integrity) play in maintaining a valued UK-wide data architecture that has evolved to serve the differing needs and policy goals that devolved legislative competency over HE brings. As the Government develops a more concrete package of proposals and a clearer direction of travel, it will be important to consider the impact of all the changes taken together, both to establish the distribution of costs and benefits of change following a period of intensive implementation of the current UK GDPR, and to avoid undesirable unintended consequences occurring as a result of implementing a further set of complex interrelated changes.

⁵ Fitzpatrick, N. (2021). Economic impact of trust in data ecosystems. Frontier Economics & Open Data Institute. http://theodi.org/wp-content/uploads/2021/03/RPT_Trust-in-data-ecosystems-23.02.21-STC-final-report.pdf

⁶ <https://www.gov.uk/government/publications/unlocking-the-value-of-data-exploring-the-role-of-data-intermediaries>

⁷ <https://theodi.org/article/what-are-data-institutions-and-why-are-they-important/>

⁸ <https://ico.org.uk/about-the-ico/consultations/departments-for-digital-culture-media-sport-consultation-data-a-new-direction/>

⁹ Cormack, A. (2021). Thinking with GDPR: A guide to better system design. Information Services & Use, 1–9. <https://doi.org/10.3233/ISU-210107>

INTRODUCTION

HESA, the Higher Education Statistics Agency Limited, is a not-for-profit private limited company owned by its members. The members are Universities UK and GuildHE. HESA is funded by subscriptions from the higher education (HE) providers from whom we collect data.

Our mission is to support the advancement of UK HE and to do this HESA operates as an independent, expert data collection, assurance, and dissemination agency, delivering reliable data services for our customers and stakeholders. We do this to enhance the competitive strength of UK HE through providing a trusted source of information, supporting better decision making, and promoting public trust in HE.

We collect, assure and disseminate data about HE in the UK. This includes a wide range of functions to support HE providers and data users in submitting, processing, and understanding the data.

We collect data from universities, HE colleges and other specialist providers of HE. Collectively, we refer to these organisations as HE providers. Much of the data we collect is published as open data.

HE providers have a statutory requirement to report data to HE funding and regulatory bodies. We work closely with providers to support this data collection process and to analyse and quality check the data they submit.

In April 2018 HESA became the Designated Data Body (DDB) for England as defined by the Higher Education and Research Act (HERA) with a statutory relationship to the Office for Students. We perform a similar official role in supporting the statutory funders and regulators of HE in Wales, Scotland, and Northern Ireland.

As part of the UK's 'statistics family', HESA also produces official statistics. These outputs are regulated by the UK Statistics Authority, which has a direct line of accountability to Parliament, independent of any Government department. As a trusted statistics agency with a statutory function, we are committed to the highest standards of conduct.

The Statistics and Registration Service Act 2007 introduced the concept of official statistics. This encompassed the previous category of National Statistics, which continues to exist as a sub-category of official statistics. The Official Statistics Order 2008 first designated HESA as a producer of official statistics, having previously published National Statistics under the former statistics framework.

HESA publishes four types of official statistics as defined by the Code of Practice for Statistics:

- **Official statistics:** statistics produced by crown bodies, those acting on behalf of crown bodies, or those specified in statutory orders, as defined in section 6 of the Statistics and Registration Service Act 2007.
- **Experimental official statistics:** a subset of newly developed or innovative official statistics undergoing evaluation. Experimental statistics are published in order to involve users and stakeholders in the assessment of their suitability and quality at an early stage.
- **Ad-hoc official statistics:** statistical analyses produced and released where there is a pressing need for official statistics in the public interest.

- **National Statistics:** official statistics assessed as fully compliant with the Code are given National Statistics status by the Office for Statistics Regulation, in line with the Statistics and Registration Service Act 2007.

HESA releases official statistics in three types of publication:

- **Statistical Bulletins:** provide high level statistics at a national level.
- **Open data releases:** larger suites of interactive open data tables providing detailed views of the data.
- **Ad-hoc Statistical Bulletins:** statistical analyses published where there is a pressing need for official statistics in the public interest

HESA's regular official statistics releases are listed below.

- Higher Education Student Statistics (Statistical Bulletin) (National Statistics)
- Higher Education Staff Statistics (Statistical Bulletin)
- Higher Education Student Data
- Higher Education Staff Data
- UK Performance Indicators for Higher Education
- Higher Education Provider Data: Finance
- Higher Education Provider Data: Business and Community Interaction
- Higher Education Provider Data: Estates Management
- Higher Education Graduate Outcomes Statistics (Statistical Bulletin)
- Higher Education Graduate Outcomes Data
- Unistats Dataset

HESA's user engagement practices are detailed in our user engagement strategy for statistics¹⁰, and we commit to continue operating and demonstrating continuous improvement, including strengthening our research function, ingesting additional data sources to improve our quality assurance and outputs, and working with our stakeholders to improve coherence in HE statistics.

In writing this response to the consultation we refer to the relevant section titles of the Government's consultation document¹¹ throughout. We number our paragraphs, making reference to the numbered paragraphs in the consultation where appropriate. Elsewhere, we condense our responses where we have concise views on an area where the consultation asks multiple related questions.

¹⁰ <https://www.hesa.ac.uk/about/regulation/official-statistics/engagement-strategy>

¹¹ Version at:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1022315/Data_Reform_Consultation_Document_Accessible_.pdf

CHAPTER 1 – REDUCING BARRIERS TO RESPONSIBLE INNOVATION

RESEARCH PURPOSES

1. Proposal to consolidate and bring together research-specific provisions (paragraph 40)

While HESA has not experienced problems in this area, we are aware of difficulties faced by others in the scientific community, and we are generally supportive of efforts to bring together provisions in one place where they can be more easily navigated.

2. Proposal to incorporate a clearer definition of ‘scientific research’ into legislation (paragraphs 41 & 42)

Establishing the concept of a statutory definition of scientific research within the UK GDPR could be helpful in providing greater certainty to HE sector researchers. It would need to be narrow enough to support public trust and confidence, and also to retain the trust of the expert community it seeks to support in delivering public benefit. We are concerned that Recital 159 may be overly inclusive and permissive as to what activity might be understood as ‘scientific’ research, as opposed to the more general use of the word ‘research’ in everyday speech. A statutory definition based on Recital 159 would have the practical effect of including uses of data that would go beyond what many members of the public might understand as scientific research. It would appear to include various sorts of ‘finding things out using data’ for purposes that might stretch public credulity over their scientific worth and be undertaken by parties who are not generally understood to have scientific aims or purposes. We therefore wish to alert the Government to a potential risk that requires careful consideration.

3. There is already a quasi-statutory definition of research in operation, which is familiar to practitioners, and trusted widely by governments and citizens across the world, as well as being embedded in practice and policy within the UK. The Frascati definition of research and experimental development from the OECD¹² is very longstanding and used widely by statistical bodies, research funders, and scientific organisations as a helpfully inclusive definition of what is meant by scientific research. It has the benefit of covering scientific research in the social sciences, humanities, arts, and professions, which the other UK quasi-statutory definition (that used by the HMRC for awarding R&D tax credits) does not generally achieve. Adding a further, separate statutory definition of scientific research based on Recital 159 could have the effect of introducing some additional complexity in the administration of science and result in the scientific research status of individual projects being different depending on whether the answer is sought on a legal, scientific, or financial basis. This would be unhelpful. Our suggestion is that the Government consider the merits of utilising an established standard such as Frascati for a single statutory definition of scientific research.
4. We recommend that the Government undertake further investigation and consultation with the scientific community, especially the HE sector, and most especially with individuals working both practically and theoretically on scientific integrity, data ethics, research ethics, grant administration, and scientific practice.

¹² <https://www.oecd.org/sti/inno/frascati-manual.htm> (pp.44-45)

5. Proposals on access to data for research purposes (paragraphs 43-45)

Our experience is that navigating lawful purposes for scientific research can be difficult. We have encountered examples of projects where, for instance, the researcher has initially and historically relied on consent, but post-GDPR realised that public task would have been a more appropriate ground, and that the original reliance on consent was better understood as an ethical consent. Given that the proposed activity involving HESA was nothing more than a data quality assurance task that was clearly in the public interest and would not have infringed the data subjects' rights in other than a very technical manner, the legal confines of the present arrangement felt, in this case, suboptimal for a good outcome. While clarifying guidance is somewhat helpful, such examples demonstrate that public good scientific inquiry is sometimes hampered by the present arrangements.

6. The establishment of scientific research as a lawful ground in its own right could be helpful in supporting the scientific community in pursuing such projects. However, we do also see some potential risks, such as a potential risk to scientific integrity if such projects are not subject to a form of ethical approval that all parties agree is sufficient, or if the definition of scientific research is too broad and does not command public confidence.
7. There would be a number of ways in which these risks could be managed, such as configuring scientific research (suitably defined) as a public task in the public interest, either by duplicating some of the language describing the public task, or by establishing the new lawful ground as a separately identified 'sub-category' of the existing public task lawful purpose. Other approaches to risk management could be achieved by limiting the bodies that can conduct such work to those with appropriate attributes of scientific credibility and public trust, or by the establishment of research ethics approval as a necessary requirement in law within a suitably trustworthy legal and institutional framework.
8. We again recommend that the Government undertake further investigation and consultation with the scientific community, especially the HE sector, the statistics community, and most especially with individuals working both professionally and academically on scientific integrity, data ethics, research ethics, and scientific practice.

9. Further processing for scientific research purposes (paragraphs 45-50)

We welcome the Government's attention to the matter of reusing data in scientific research. The nature of research in extending human understanding can mean that the value of data collected now may not be fully realised for a considerable period of time. This could be years or decades depending on how our understanding evolves, and the urgency of the problems we have yet to encounter. We are familiar with scientific data about the natural world retaining and increasing in long-term value. Centuries-old data on temperatures and solar activity remain useful today and are used for public good in ways that could not have been anticipated by the original collectors. It is arguable that our societal capacity for collecting and storing data about the social world should also be considered in this light. However, given that this is data about people, a range of ethical safeguards are required to ensure that public perceptions do not put long-term uses of data for scientific purposes at risk. This has two logical corollaries: the first is that the lawful basis for scientific research should be made resistant to subversion or 'scope-creep' for other purposes; the second is that the scientific establishment must also maintain the highest standards of integrity over the long-term.

10. Recent Government research¹³ showed that ‘people have become more positive in their attitudes towards science and scientists’, and that trust in scientists making a valuable contribution to society is very high (89%), but that the public feel there is little option but to trust them – 47% feel that rules are insufficient to stop scientists doing what they want, and 74% feel that scientific independence is compromised by the interests of funders. This perhaps goes some way to explaining public scepticism about some aspects of data sharing, and the value of institutions in stewarding public trust. For example, when considering sharing personal health information, 90% trust the NHS, with only 35% trusting private companies. Research organisations score 73% and Government 61%. We therefore contend that the evidence is consistent with suboptimal trust in the use of data for both scientific research and public interest purposes. This indicates that the problems that it is most important to overcome with legislation relate to bolstering the ethical considerations relating to data collection, management, and use; ensuring that there is equity around who accesses, uses, and benefits from data; and engagement of the public in science¹⁴. Measures to support these would reduce public fears around misuse and increase trust.
11. However, the evidence from the Government’s research into public perceptions of science¹⁵ also demonstrates that there are reasonably high and rising levels of ‘scientific capital’ among the public and that scientific capital is correlated with greater trust in science and scientists. Under these circumstances it seems sensible to permit data subjects to give broader consent to the use of their data for scientific research. It is standard for research participants to understand how their involvement can help scientific aims, with engagement work typically seen as central to the conduct of good science. These processes as well as ethical approval and oversight will tend to increase the scientific capital of participants. Public attitudes to science could be enhanced through researchers being able to demonstrate that by participating in a single specific research project, a data subject is also potentially able to contribute to the scientific endeavour more generally. Recognising this willingness on the part of a data subject to make their personal data available for a wider range of scientific endeavours in law, through a lawful basis for scientific research and an associated form of broad consent, appears a very welcome idea.
12. Research participants can sometimes be difficult to maintain contact with, so some limited disapplication of the absolute requirement to provide information about further processing could be helpful. However, this does raise potential ethical issues, and it may be that an appropriate ethical practice could be evolved by the scientific community to make appropriate use of disapplication in ways that ensure scientific and ethical integrity is maintained. Examples such as the publication of processing notice information at a stable URL that is maintained and updated over the long term might give data subjects confidence, and a point of contact should they wish to exercise their rights or express their opinions on onward use.
13. It is not clear to us how the ‘broader areas of scientific research’ would be constituted, since interdisciplinarity may bring researchers from various perspectives together in a manner that defies easy categorisation. Categorisation of academic disciplines becomes harder the

¹³ <https://www.gov.uk/government/publications/public-attitudes-to-science-2019>

¹⁴ This would also be consistent with a diagnosis that in many ways the UK is closer to the ODI’s ‘data wasteland scenario’ and requires measures to increase trust to derive better value from data. See: <https://theodi.org/about-the-odi/our-vision-and-manifesto/our-theory-of-change>

¹⁵ <https://www.gov.uk/government/publications/public-attitudes-to-science-2019>

closer one gets to the frontiers of knowledge, and insights from one discipline can transform another profoundly. Moreover, science is a single endeavour incorporating and synthesising diverse bodies of knowledge and investigatory techniques in the empirical pursuit of truth. Ascribing arbitrary limitation on the 'type of science' that can be done would likely pose practical difficulties on research organisations and is unlikely to resonate with members of the public with sufficient scientific capital to wish to contribute their own data.

14. The proposed clarification of the lawfulness of further use of data for scientific research always being compatible with the original purpose, appears helpful, though where the original purpose was consent for a purpose other than scientific research (whether construed narrowly or broadly) this potentially raises ethical concerns. Also, applying this provision to other 'research' that does not conform to a suitable definition of 'scientific research' could risk public confidence. Further clarity from Government would be welcome on achieving balance between ensuring public interest scientific re-use and maintaining public confidence that such activity falls within data subjects' reasonable expectations. We therefore recommend that the Government undertake further investigation and consultation with the HE sector in particular.

FURTHER PROCESSING

15. Proposals for allowing further processing subject to compatibility tests (paragraphs 51 – 54)

The Government raises three key areas of uncertainty and recognises that there will be challenges to ensure re-use remains fair and within reasonable expectations. These proposals on further processing raise fundamental issues about fairness, consent, and transparency. As our response to the use-case around further processing for research makes plain, this is a complex area that raises important issues of principle, such as legitimacy, fairness, and transparency. We think that these issues need to be handled carefully and thoroughly. Data is not a fungible asset and individuals can think differently about different personal data. Context is key, and exactly what data is requested, and by whom, and for what purpose, and under what conditions, will all influence data subject behaviour. There is a delicate balance to be struck between compatible purposes and individual rights. The proposals are unclear on how this would work in practice. Further clarification and examples would be required to come to a more informed view.

16. Greater clarity over the compatibility test would be welcome. Although we perceive benefits in HESA gaining greater flexibility in how it ingests, processes, and shares data in order to fulfil its public task and charitable objectives, we perceive a greater number and magnitude of potential risks.
17. The first risk we can see relates to the idea that further processing could be lawful if incompatible but based on a law that safeguards an important public interest. This is unclear; all laws are attempts to safeguard public interests, and have been deemed important enough to legislate on, so this brings all public law within scope. This potentially very wide range of lawful use cases would put at risk public confidence in public interest data collection and use. As statisticians we are acutely aware of risks to data quality if public trust in data is undermined, and these risks to data quality could impact the value of data and statistics to Government. If for example data subjects choose to withhold or frustrate data collection efforts, the insights that can be drawn can be damaged through the

introduction of (largely unmeasurable) bias. Boycotts of statutory data collections for ethical or political reasons are not unknown in the HE sector.¹⁶ We therefore consider the maintenance of public trust among relevant data subjects to be of high importance in securing reliable data to serve public tasks in the wider public interest.

18. The Government's proposals on limitations and safeguards are not clear enough for us to give a considered response but ensuring that purposes are ethically compatible in the view of the overwhelming majority of data subjects is vital. Later questions on the role of data intermediaries / institutions may offer a potential mechanism for identifying and regulating further processing in the public interest. Charitable organisations, which must be apolitical as well as serving a public interest, may offer a suitable regulated route to supporting trustworthy stewardship of public data free from influences that could introduce bias.
19. We would welcome further clarity over when processing can be undertaken by a data controller different from the original controller. In practice we believe HESA will continue to require some form of additional application for any additional processing by a data controller other than that specified in the license we have issued, whether the additional processing is deemed further or new. Further clarification is needed on how the Government anticipates this provision working in practice, with examples.
20. The proposed changes to further processing would potentially give HESA greater flexibility in how it processes, ingests, and shares data in order to fulfil its public task and charitable objectives. However, this benefit is likely to be relatively modest. We also perceive a range of concomitant risks.
21. Risks associated with the general approach proposed include the growth of undesirable practice that damages the data economy more generally. Sufficient safeguards would need to be in place to prevent this. Under the outline proposals, it is not clear how HESA could be sure that its data will be handled appropriately, by a suitable person, in a sufficiently secure environment, and for appropriate projects. There is therefore a risk to public trust and confidence in the data industry which could result in the behaviour of (even inadvertent) bad actors impacting on the viability of responsible custodians. Poor practice can be avoided with sufficient forethought and scrutiny, but if allowed to grow, could resonate in the public consciousness rapidly. The resulting damage to public discourse would prove difficult and costly to address and may not be wholly rectifiable. In particular, we perceive a potential risk to statistical data quality from various forms of measurement error arising due to changes in data subject behaviour.
22. If the legal 'gateway' created by the proposals of further processing is too wide or permissive, this opens the potential for HESA to effectively lose control of the microdata it issues under license. This would risk the trust of our customers and data subjects in the services we offer, and would effectively void the controls we have established for permitting and limiting various broad categories of lawful onward use of HE providers' data under GDPR.¹⁷
23. The proposals also present new risks to supporting data users to raise data literacy, both among the general public as well as the users of our data specifically. Not all data, even from reputable sources, is always properly catalogued. Even excellent catalogues and user

¹⁶ For example, the 2017 boycott of the National Student Survey: <https://www.ucu.org.uk/boycott-the-nss>

¹⁷ <https://www.hesa.ac.uk/support/provider-info/subscription/onward-use>

guides cannot cater for every possible use case for the data they seek to explain. As data would propagate more easily and extensively under the proposals, there would be a greater reliance on traceability and documentation. Even capable and trustworthy analysts may misunderstand or misuse data when they do not have access to information, advice, and guidance from knowledgeable producers of data. Responsible custodians seek to support users in understanding the explanatory power of data, so that they can obtain the value they wish to achieve. If the proposals mean responsible custodians may not know who exactly is using their data and how, their ability to support users is limited, and the risks of misinterpretation or misuse are raised. Consequently, where data is of considerable public interest, this may mean responsible controllers have to devote additional resources to monitoring and intervention strategies than at present. Monitoring and intervention is a reactive and suboptimal approach to handling public domain data quality issues that are more easily, cheaply, and effectively resolved at source. Current licensing arrangements enable us to know who our users are and to understand and support their needs.

24. Limitations and safeguards should be set to introduce very specific and limited legal easement for particular use-cases where a well-evidenced and widely agreed limitation on the pursuit of public interest justifies a reduced administrative burden.

25. Proposals for further processing on the grounds of legitimate interests (paragraphs 55-62)

Since HESA mostly works with data collected under other lawful bases than consent, we limit our response to the general observation that there must be sufficient safeguards in place to protect data subjects, to protect public confidence in data, and to avoid the growth of non-response and measurement error, where consent is the lawful basis for processing. Further processing for the purpose of scientific research would be desirable, subject to appropriate ethical safeguards. Further clarification of the proposals, with examples, would be helpful.

26. We agree that the proposal to create a limited, exhaustive list of legitimate interests for which organisations can use personal data without applying the balancing test would simplify the process of applying the balancing test in many cases. This would be a helpful piece of legislative streamlining that would benefit data controllers, which could be made to avoid impacting on data subject rights or public confidence.
27. The list of activities offered in the proposal for which the legitimate interests balancing test would not be required appears helpful and correct to us. In most cases the risks arise not from the legitimate interest itself, but from the apparent license it affords to data controllers in stewarding their data.
28. There is a likely reaction on the part of responsible data stewards to exercise additional caution in releasing data. Risk-aversion could result in a move towards 'gold-plating' privacy regimes, which could be counterproductive to the aim of supporting innovation in the data economy.
29. Professional standards and accreditations are important in assuring the quality of organisations, practices, and skills in using data. Deregulation will likely promote growth in these industries as many data stewards will perceive the value in third-party assurance of quality, in order to reassure subjects that their data will be handled with respect. We

encourage the Government to give thought to mandating or supporting an accreditation regime that promotes public confidence in the industry, fosters efficiency and good practice, and supports the social benefits of ethical data sharing.

30. HESA agrees that the legitimate interests balancing test should be maintained for children's data, irrespective of whether the data is being processed for one of the listed activities. Much data on or about children is collected without their consent or knowledge. We have a duty of care to children in particular: to engage with protecting their interests on their behalf, and in demonstrating this to them and their parents/guardians. The proposal is a proportionate response to the limitations of applying a more adult form of engagement, in managing this ethically important issue.

AI AND MACHINE LEARNING

31. Proposals for the responsible development of Artificial Intelligence and Machine Learning (paragraphs 63 – 113)

HESA agrees with the Government that all artificial intelligence (AI) and machine learning (ML) approaches are built on data. As statisticians we recognise the potential for bias most especially in data that is not drawn from the population of interest at random. Like all other kinds of analysis, AI and ML suffer from the limitations introduced by data missing not at random. That the explanatory power of AI and ML approaches is limited by the characteristics of the training data is well-known. More data is not always helpful in addressing issues of bias.

32. AI and ML are important groups of advanced techniques for deriving meaning from data. We welcome the Government's focus on fairness in the use of AI and ML. We do not consider that the regulatory framework for AI and ML approaches should handle them in a fundamentally different manner than other processing techniques that derive meaning from data. Data protection is the right legislative framework for handling questions of lawfulness, fairness, transparency and proportionality in the use of data in AI and ML applications. It would therefore be appropriate and efficient to ensure the ICO retains full oversight of this area. Other approaches would risk creating a confusing and potentially overlapping regulatory regime.
33. HESA's central goal could be characterised as sense-making, rather than decision making. Our use of advanced techniques is focused on pursuit of our charitable mission, to support the advancement of UK HE by collecting, analysing, and disseminating accurate and comprehensive statistical information in response to the needs of all those with an interest in its characteristics and a stake in its future. However, we are aware of uses of data in learning analytics applications, where decisions about interventions can be made on the basis of data, and some automation of aspects of the teaching and learner support can be automated by using AI and ML techniques. In these scenarios the opinion of our sector has been that the same principles of equity, ethics, and engagement must be observed as with other uses of personal data. A good example of this is the Code of practice for learning analytics¹⁸ produced by our data processing partners Jisc, to support the development of responsible, appropriate, and effective learning analytics practices.

¹⁸ <https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics>

DATA MINIMISATION AND ANONYMISATION

34. Proposals for data minimisation and anonymisation (paragraphs 114-125)

HESA agrees with the proposal to clarify the test for when data is anonymous by giving effect to the test in legislation. Clarity over the tests would promote confidence and aid alignment of practices. HESA agrees with the ICO position that anonymity is largely a matter of resources of the data controller. We favour anonymisation approaches that take into account objective measures of the factors bearing on re-identification of an individual, such as time, cost, and technology. We also recognise the value of controls over improper motivations using researcher accreditation processes. The proposals could allow HESA the opportunity to streamline our due diligence activities in areas where we are confident, and have applied a test, that data is anonymised. This could be particularly relevant to the area of tailored datasets and other microdata that do not contain direct identifiers, supplied for research purposes.

35. Recital 26 of the UK GDPR would provide a basis for formulating the text in legislation, as it is familiar to practitioners, and simple and consistent both with existing legislation and practice.
36. HESA considers that the re-identification test under the general anonymisation test is a relative one. We recognise a variety of different capabilities among our users. HESA's widely used data capability framework¹⁹ categorises different levels of maturity, and explicitly recognises capabilities can change over time, especially as technology develops and achieves higher levels of market penetration, and as skills prevalence around data and technology grows. Judgements have to be made in the light of imperfect information on third party capabilities, so have to be based on reasonable likelihood or the balance of probabilities. Ongoing requirements for due diligence persist as the relationship between a stewardship organisation and a licensed user of data develops over time.
37. We look forward to the Information Commissioner's forthcoming guidance on privacy-enhancing technology (PET). It is important to recognise that PET is still in its infancy, and there is not yet a mature market for translation of theoretical conclusions into operational solutions for making data available more widely. Our continuing exploratory work in this area has so far mainly emphasised the limitations of such technologies. Active promotion of such technologies is therefore premature, but there may be value in revisiting this in a few years' time, once PET products and standards are more widely used.

DATA INTERMEDIARIES

38. Proposals on innovative data sharing solutions (paragraphs 126-137)

HESA welcomes the Government's focus on the role of data intermediaries in supporting responsible data sharing. HESA is an established data intermediary, a company limited by guarantee and a registered charity, working principally with HE providers and government bodies to deliver data services within a range of statutory frameworks. Longstanding leadership from HESA has helped to foster a HE data landscape characterised by honesty, impartiality, and rigour, in order to pursue the following aims for our stakeholders:

- Promote better outcomes.

¹⁹ <https://www.hesa.ac.uk/support/tools/data-capability>

- Increase operational efficiency.
- Minimise burden.
- Rationalise, harmonise, and inculcate best practice.
- Demonstrate adherence to fair, transparent practice.

Our data collection Codes of practice²⁰ and associated work on the HE data landscape have been central to pursuing these aims. However, we face several persistent barriers in pursuing a data economy that liberates the power of HE data in a responsible and ethical manner. We therefore welcome the Government's interest in the role that bodies like HESA can play in supporting better outcomes through sharing data.

39. We have successfully navigated many of the barriers we have faced over nearly 30 years of operation. Our experience could be helpful as a case study in supporting the Government to play various roles to stimulate and support an improved data economy around the HE sector, and many of these lessons may be transferrable to other sectors. HESA operates within a complex regulatory landscape. We operate at an intersection of civic interests that sees us affected by legislation on charities, on data and information, on official statistics and on the devolved competency of HE. Voluntary compliance with various established standards for information management and good governance are deemed essential, and we manage risks carefully. However, as an ambitious organisation in a knowledge-rich sector we have long pursued goals of advancing HE through influencing the development of a richer and more efficient data landscape in the interests of HE providers, their students, and their staff.
40. This consultation from DCMS brings into plain sight a gap that has been evident to HESA for many years. Working principally under statutory frameworks designed for HE, we have demonstrated that the regulatory bases for data collections can act as a platform for establishment of a wider data sharing economy, which has wider economic and social benefits sometimes well beyond the regulatory or funding roles of our statutory customers. For example, HESA's subscribers in the HE sector have education and research at their heart, but they are also major employers, important civic actors, business incubators, hospitals, schools, and construction sites. These roles are salient to the data economy around the sector but are arguably not the intended focus of legislation that is primarily about student interests and academic standards.
41. The interests of our statutory data customers regulating and funding the HE sector will naturally vary over time, and when they change this can impact on the availability of data for other legitimate uses. This means that the proper and legitimate decisions of statutory bodies can have ramifications on data sharing that are felt beyond their sphere of responsibility, and for which interests they are not necessarily permitted or able to do data governance. Regulators and funders of HE are important users of HE data, and we pursue their needs with vigour, just as we pursue the collective interests of the HE sector in obtaining robust and consistent data to support planning and public information. In our experience, regulatory functions benefit from good stewardship and governance of a data ecosystem that supports an efficient and high-performing competitive HE sector, and this is tacitly acknowledged by various public sector bodies, but rarely a focus of their activities as it is the focus of ours. While our experiences of working with statutory data customers on these matters has broadly been positive, the statutory data customers are ultimately bound by their own lawful purposes.

²⁰ <https://www.hesa.ac.uk/innovation/data-landscape/Codes-of-practice>

42. If the role HESA plays as a sector data steward were recognised under a general public law framework for data and information outside our specific responsibilities under the various statutory frameworks for HE, this could support the wider economic and social aims the Government has expressed through the consultation, and we could shape our own data governance activity around broader Government aims focused on regional and national economy and society, while retaining focus on the specific needs of our statutory customers. The establishment of a regime where the value of trusted data institutions²¹ is recognised in law could offer HESA a modest claim for consideration by Government for the role we play as a key data steward in a high-value sector dominated by public spending and low-margin charitable endeavour. We would be happy to produce detailed case study materials on request showing how various styles of Government action could help with this.
43. In our devolved sector decisions about onward use are influenced heavily by the different attitudes of regulators of HE in each part of the UK. HESA must expend considerable efforts to explain and lobby for the consideration of public bodies in sustaining a little-understood set of mutually beneficial data sharing agreements. As the HE policy contexts across the UK diverge, the value of consistent data standards to permit comparisons becomes increasingly evident. This is a heavy burden to bear for a relatively small charity. Being able to demonstrate to our customers that we are playing a stewardship role that is valued by being codified or protected in legislation would strengthen our convening power substantially. UK-wide legislation could also support UK-wide coherence and comparability of data, and support HESA in meeting the needs of users in Scotland, Wales, England, and Northern Ireland within a clearer framework, as well as offering support to the needs of city regions. By configuring the role of data institutions as part of a governed and statutory space outside our sectoral interests in HE regulation, HESA could play a central role in supporting the use of HE sector data in the domains of innovation, economic performance, and generation of public value.
44. Turning to the lawful grounds other than consent that might be applicable to data intermediary activities, as well as the conferring of data processing rights and responsibilities to those data intermediaries, HESA does not generally utilise consent as a lawful basis for its data stewardship activities. Most of our activity to date utilises the lawful grounds of public task, or legitimate interests. HESA normally seeks to establish its own public task ground where it is collecting data that one of its statutory customers requires to fulfil a legal obligation.
45. Similarly, long-established market information mechanisms in HE such as the newspapers' university league tables might arguably be considered within scope of HESA's legitimate interests as a charity charged with the advancement of HE, rather than the performance of a public task in the public interest (which is arguably already served by mechanisms such as the Office for Students' Register and TEF in England). The balancing test has been performed for the supply of pseudonymised data to league table compilers. There is an argument that where data has been collected for a legal obligation, and where sufficiently strict disclosure control, ethical consideration, and license terms are applied that data collected under a legal obligation could be supplied usefully under our legitimate interests. HESA operates under a charity law framework. This means that our legitimate interests are themselves regulated by charity law, which requires that we produce public benefit.

²¹ We are influenced by the work of the Open Data Institute in this regard: <https://theodi.org/article/what-are-data-institutions-and-why-are-they-important/>

Anything outside our charitable purpose could not be considered a legitimate interest. Moreover, it is in our corporate interests to continue providing this benefit, rather than undermining the business case for our ongoing operation. While these conditions do not necessarily prevail for all data institutions, we believe charity law can offer an important part of a framework that the Government could choose to utilise in codifying the role of responsible data intermediaries in legislation.

46. Government could help bolster HESA's ability to increase the value it creates with data and retain a position of trust in the following ways. First, by stewarding the scaling and mainstreaming of data institutions by producing a policy statement and sponsoring and funding a programme of activity to identify the levels of data capability and coverage of data institutions across all industry sectors. Second, through leadership in setting the agenda for showing how the legitimate interests of charities can support the growth of a public interest data economy. Third, by undertaking further legislative early-stage intervention to float the concept of a specific function for data institutions in primary legislation in a green paper that recognises and regulates the conduct of such organisations in the public interest. We suggest building on the provisions in part 5 of the Digital Economy Act 2017 to create an appropriate governance regime and legal basis for establishing data sharing agreements in the public interest. Additional Government accreditation beyond mandating charitable status would be helpful – as this is a relatively new area of public life and not necessarily well understood by data subjects, establishing a strong digital Government regulatory oversight would help promote public confidence, support compliance, and foster the formation of appropriate standards, such as assessing the alignment of charitable purpose with a statement of legitimate interests. We think that a funded role to undertake a quasi-regulatory data governance activity would be better in the first instance than an accreditation scheme, but in time a pathway to a more market-based approach might be credible. We would welcome the cooperative involvement of the ICO as a strong independent economic regulator in overseeing the growth of data intermediaries. However, we also recognise the potential complexity of regulatory structures around data intermediaries, and we would be keen to be involved in work to ensure that the proposals are not unduly burdensome on organisations that are already highly regulated.
47. Turning to the impacts of the proposals on people who identify with the protected characteristics under the Equality Act 2010, the Government has already recognised the special care that must be taken with data about children, and this is welcome. The reform already seeks to recognise the fine balance between the privacy rights of individuals and the public good in a number of areas. The proposals are not yet fully formed enough to know what the impacts might be. Since the legislation deals with fundamental issues concerning asymmetries of power around data and information, it would be appropriate to note that digital exclusion is a consequence of poverty and disadvantage, and that poverty and disadvantage are unevenly distributed by protected characteristics. We would therefore suggest that Government's equality impact assessment takes into account the prevalence and impact of digital exclusion among individuals identifying with the various protected characteristics. This approach would assist in leveraging maximum opportunity from this policy development. The value of an individual's data is not correlated with their income or status, but with the explanatory power it affords the researcher in obtaining scientific insights that lead to public goods.

CHAPTER 2 – REDUCING BURDENS ON BUSINESSES AND DELIVERING BETTER OUTCOMES FOR PEOPLE

PRIVACY MANAGEMENT PROGRAMMES

48. Proposals on privacy management programmes (paragraphs 138-183)

HESA welcomes the focus on accountability and sees this as essential to public trust in the data economy. HESA believes that the responsibility of a controller to demonstrate accountability rises with the risks associated with the data they control. While the arguments around reducing requirement, increasing flexibility, and making accountability more risk-based appear generally sound, they will make little difference in the HE sector. Higher education is a complex human-centred professional services industry, dominated by large organisations controlling and stewarding very large quantities of sensitive data. Cybersecurity and ethical responsibilities are, in our experience, taken seriously, and relatively conservative attitudes to data handling predominate. As a responsible controller of the personal data of students and staff in HE, HESA would continue to take a relatively risk-averse attitude following any legislative change.

49. HESA does not anticipate any particular benefits occurring as a result of being required to develop and implement a risk-based privacy management programme. While the apparent flexibility afforded by the proposal is superficially attractive, HESA cannot determine any specific benefits created by it, and has identified some potential new risks. As a controller of sensitive personal data to fulfil legal obligations and in the public interest, HESA will remain accountable for demonstrating a very high standard in its privacy regime. We already operate a regime that seeks to identify and manage risks in an appropriate manner. Our assessment is that under the proposals, we would likely continue to operate much as we do at the moment, undertaking Data Protection Impact Assessments (DPIAs) as a matter of course as part of our privacy regime. Our strategy recognises security of data as a core aim²². Under the liberalisation implied by the Government's proposals we may in future face pressures to differentiate our conservative data stewardship from that of organisations that take a more permissive attitude to sharing personal data. This could result in looking to 'gold-plate' our approach through certifications, or to engage in other forms of 'privacy theatre' that are designed more to promote the image of HESA as a responsible controller, than to add materially to our approach to the rights of data subjects. This could add costs without achieving additional benefit. Neither do we believe that data subjects of the HESA data would benefit from the proposals, since we would be unlikely to wish to adjust our approach significantly, unless and until public attitudes change. We believe that accountability must lead to enforcement, in order to maintain high standards and public trust in the privacy regime.

DATA PROTECTION OFFICERS

50. Turning to the role of Data Protection Officer (DPO) on implementation of GDPR we did not have to appoint additional staff as we already had the relevant role in place. However, this is an area where there has been a growth in provision, and we have invested in this area over the past few years to enhance our approach. We would not benefit from removal of the requirement to appoint a DPO. From HESA's perspective the designation of a DPO is part of a long-term strategic capability. Given the nature of the data handled within the HE

²² <https://www.hesa.ac.uk/about/strategy#Security>

sector, we think it is likely that most HE providers will also perceive a similar continuing need to designate a DPO, and we would encourage this as a good practice approach for operating their own privacy regimes. However, we recognise the increased burden on smaller HE providers. Nevertheless, in the HE sector we believe continued identification of a DPO role will remain standard, even if smaller HE providers choose to combine it with other functions.

DATA PROTECTION IMPACT ASSESSMENTS

51. HESA finds that Data Protection Impact Assessments (DPIAs) are helpful in the identification and minimisation of data protection risks to our activity, and HESA would not benefit from the proposal to remove the requirement for organisations to undertake DPIAs. Data Protection Impact Assessments form part of the information security management system we have invested in. They support HESA in taking a proportionate and risk-based approach to its work, and in supporting lawful use by our data processing partners and customers in Government. They have proved to be a useful tool for describing how we use data, identifying the risks to data subjects and providing a means for recording, monitoring, and mitigating these risks. Our DPIAs form a part of our general approach to risk management. We have decided to retain DPIAs or something very like them in any future privacy regime we operate.

PRIOR CONSULTATION

52. The consultation asks us to comment on why few organisations approach the ICO for 'prior consultation' under Article 36 (1)-(3). We imagine that this relates to a perceived risk of inadvertently notifying them of something on which they may then take enforcement action - as indicated in the consultation. There has been and is an ongoing need for authoritative advice and guidance on challenges faced by data controllers in specific circumstances. We are open to the value of encouraging consultation with the ICO by identifying this as a mitigating factor during future investigation or enforcement. The main point in approaching the ICO is to better manage information security and privacy in specific instances where their specialist guidance and cross-sectoral experience would be helpful, and disincentives to approach them could lead to worse outcomes for data subjects. We therefore welcome proposals that help to foster open and honest discussions with the ICO. We are not sure whether removing the requirement for prior consultation will necessarily achieve this. We understand the ICO's point about needing to retain the capacity to use the full range of enforcement tools, and all actors in the data economy benefit from this. However, the data economy and data subjects also benefit from the application of mercy to temper justice in deserving cases. This is a difficult area to resolve through legislation, but we can see the potential for greater reassurance to be offered to parties considering approaching the ICO through the ICO publishing notices or position statements that help set expectations.

RECORD KEEPING

53. HESA does not see any value in the proposal to reduce the burden on organisations by removing the record keeping requirements under Article 30. Organisations are required to demonstrate understanding of the personal data they process, and keeping good records is central to achieving this. While there may be room for exploring more efficient ways of handling record keeping requirements, a common baseline offers a level playing field, and promotes the growth of information management solutions. Improving the competitiveness

of this field is likely to bring benefits for data subjects (through the maintenance of high record keeping standards) and for businesses (through a more competitive market for software and services).

REPORTING THRESHOLDS FOR BREACHES

54. We think that exploring and clarifying the threshold for reporting data breaches could be beneficial. It will be important to ensure a full risk assessment is produced and recorded to justify the assessment of whether a breach meets the threshold in order for organisations to have confidence that they are managing privacy risks appropriately. Internally, we devote time to analysing 'near misses' where no breach has occurred, but we have perceived a risk that potentially one might have occurred under different circumstances. The privacy management profession is still somewhat in its infancy, and in the coming years the insights from such scrutiny can be helpful to organisations seeking to improve their practice. The ICO make a comparable argument in paragraph 54 of their response, applying this thinking to the whole cyber security landscape. We echo the value of that insight from our own experience.

VOLUNTARY UNDERTAKINGS

55. HESA cannot see value being created by the proposal to introduce a voluntary undertakings process. Responsible custodians should in any case be identifying actions they can plausibly take in the event of a breach, and this should form part of their engagement with mitigating the risks realised as a result of the breach, and form part of their dialogue with the ICO. There are risks to the credibility of regulation if voluntary actions reduce the risk of bearing enforcement costs, when such actions were the responsible ones to take in the first place. This has the potential for rewarding poorer practice, without providing concomitant benefits for organisations following good practice.

FURTHER QUESTIONS

56. We have not identified any specific weaknesses in the current regime that demand attention. Our perception of the material risks will not be likely to change following the proposed reforms, and hence our approach to managing the risks will not be likely to change. If the reforms lead to a lower level of public trust in the data economy, there would be a benefit in promoting Government-backed accreditation for responsible organisations to demonstrate their standards compliance and differentiate themselves as quality providers in the marketplace. However, we hope that the reforms, when fully clarified, will increase public confidence rather than decrease it.

ALTERNATIVE SCENARIO PROPOSALS

57. Turning to the Government's alternative scenario proposals, we do not see particular burden in the slight duplication in Article 30. These requirements form part of our approach to privacy impact assessment, and we see continued value in them. The alternative breach reporting approach seems identical with the main scenario, so our opinions are the same.
58. We think that there should be comparability between private and public requirements for the appointment of a DPO. A threshold for whether to appoint or not, based on the risks inherent in the data being processed, would be the fairest and most reasonable approach.

A level playing field between public authorities and private businesses that are not public authorities, would also present the fairest and more reasonable approach.

59. Considering the two options presented at paragraph 184d(V), further development work would be required, as we can see advantages and disadvantages of both models. The level playing field of 184d(V)i. is attractive, whereas the recognition of the important role that DPOs play in driving standards in 184d(V)ii. is also attractive. Clarity on the proposals would be needed to come to a firmer opinion.

SUBJECT ACCESS REQUESTS

60. We consider the requirements around Subject Access Requests (SARs) to be clear and have not experienced 'vexatious' requests. Due to the nature of our data, we cannot normally rely on a unique and shared identifier to help us identify all records relevant to a SAR across all our datasets. While we have put in place several process improvements over the years, technology improvements only help us so much, as we also need to cater for features such as name changes, study at different times and levels with different providers, and changing roles as staff and students. The costs we currently bear in servicing SAR are therefore higher than any reasonable charge that might be implemented post-consultation. SARs can divert the same scarce resources we use to produce official statistics – another time-limited activity. However, at the current level of activity this has proved manageable. Were this activity to increase significantly, or bulk requests to become common, it could become more difficult to manage for us.
61. We have not sought to make use of the 'manifestly unfounded' threshold justification for refusing compliance or to attempt to levy a fee. We support the exploration of practical measures to ensure that supporting individuals in exercising the right remains manageable for organisations facing unprecedented demand for SARs. Our costs per SAR remain higher than any charge that we would consider compatible with supporting individuals to exercise their fundamental data protection rights. A fee set at a reasonable level is unlikely to make a sufficient contribution to the costs of service to make collecting the fee worthwhile for us. The main issue for HESA is likely to be the rationing of resource to meet conflicting demands over a short space of time. Time extensions for compliance would likely be a more helpful regulatory response if demand were to temporarily outstrip our capacity to supply SAR data.

PRIVACY AND ELECTRONIC COMMUNICATIONS

62. Email analytics, such as whether an email has been received, seen, or opened, and mobile analytics, such as whether an SMS message has been received or viewed are arguably types of data collection or other processing activities that should fall under the definition of 'analytics'.
63. Limitations covered by Regulation 6 of Privacy and Electronic Communications Regulations (PECR) prevent us from various analytics-based optimisations that we might choose to pursue in operating our national Graduate Outcomes survey. This is a population survey of all graduates 15 months after finishing their HE course and is conducted under statutory authority across the UK.
64. The limitations we face on utilising analytics in the Graduate Outcomes survey results in:

- a. Some wasted effort and expense for our subscribers (HE providers) in targeting communications through non-responsive channels.
- b. Throttling our opportunity to undertake experiments designed to improve the salience of communications for respondents, and to drive up response rates to meet the expectations of the sector's regulator.
- c. Difficulty for us to obtain statistically representative data, which may be resulting in unobserved biases occurring in the data.
- d. Possibly some unnecessary irritation on the part of potential respondents as we utilise the full range of contact details available to us, rather than homing-in on those channels most likely to garner a response.

We would be happy to see limitations on more extensive analytics confined to a limited range of purposes, such as scientific research, public tasks, legal obligations, if there were privacy concerns regarding their wider application.

65. We have identified some areas where the proposed removal of consent requirements would assist us in conducting the activities for which we have statutory backing, and which further our charitable aims, such as obtaining feedback on respondent activity with respect to our communications, in order to improve their effectiveness. However, while we would have concerns that too extensive a liberalisation of the consent requirements could be counterproductive, the evidence offered by the Government regarding the implementation of GDPR in France, along with the research commissioned by the ICO on cookies, demonstrates that there would be value in exploring a reduction in the regulatory requirements for both data subjects and organisations. We believe that browser- and non-browser-based methods for capturing and making queryable or broadcasting user preferences, and the development of an information standard for these preferences, would be a laudable goal.
66. There are categories of cookie that have no apparent justification for exemption from prior consent, as they relate to individuals' fundamental rights to refuse consent from certain forms of processing. There is potential for individuals to express their consent preferences in an automated fashion through browser and software settings, or through a data fiduciary. Government leadership on this issue could help create and sustain a technology framework for supporting subjects and businesses to operate lawfully while reducing the fatiguing aspects of compliance-focussed pop-ups.
67. Sectoral codes under Article 40 of UK GDPR could have a potential application in the HE sector to support learning analytics technology. Research by our analytics partner Jisc has focussed on this area. A sector code focussed on carrying out learning analytics responsibly, appropriately, and effectively, is already in existence. While this code is offered in support of HE providers, it could be developed into something more formally recognisable under Article 40.
68. There are benefits to data subjects in being able to avoid 'fatigue' when seeking to consistently observe their rights in the least permissive manner, by automating the transmission/broadcast of their general preferences for handling cookies. There are benefits to users in maintaining free access to information without requiring the individual to 'pay' with their data by requiring agreement with processing before granting access. Organisations would need to have mechanisms in place to receive and act on the preferences expressed by users. There would be risks to public confidence if organisations were to ignore data subjects' preferences so there would need to be an adequate

enforcement regime in place. One risk would be that an organisation that does not rely on consent for lawful processing of certain analytics data might face confusion about how it should respond to user preferences. This could present dilemmas of both an ethical and a legal nature. Either the full range of lawful possibilities should be encodable into software, or guidance produced on handling user preferences in situations where consent is not the lawful basis for processing.

69. If users were able to encode their default position on consents for cookies, etc. into browsers or other software using a consistent data standard, this could theoretically permit selection of privacy options to be automated.
70. Government consideration of legislating against 'cookie-walls' that limit access to information to users who are unwilling to be surveyed would be welcome.
71. HESA would like to see the soft opt-in extended to non-commercial organisations. Removing this limitation could improve HESA's future opportunity to undertake follow-up surveys or similar data collection activities of benefit to our users. While HESA does not currently face any specific limitations from not having access to the soft opt-in, we recognise the value in establishing a level playing field between organisations of different kinds as the initial starting point of legislation (even though there may be good reasons for deviating from this starting position in the light of scrutiny). We are not aware of the arguments for excluding charitable or public sector organisations from the soft opt-in, and, as long as there are no strong countervailing arguments of which we are unaware, we would be in favour of establishing a level playing field for its use.
72. Turning to the issues raised by the potential for updating the ICO's enforcement powers so that they can take action against organisations for the number of unsolicited direct marketing calls 'sent', we make the following observations. We do not consider that communications for our Graduate Outcomes survey (which has an annual sample of over 750,000 recent graduates) are marketing. However, HESA places a high volume of calls. We would be concerned if the scale of this lawful activity, which has been commissioned by the state regulators and funders of HE, were to become likely to result in enforcement action. We operate an opt-out process and a complaints mechanism, and we aim to operate to very high standards, consistently. There would be a benefit if the harm to data subjects is reduced, but it is not clear from the proposal that a high volume of calls in itself represents a problem. Our view is that regulatory interventions should continue to be based on the risk of harm, rather than arbitrary measures like high volumes of activity. There could be a risk to organisations if legitimate activity is curtailed for insufficient reasons. For instance, a disruption to our data collection could jeopardise the availability of official statistics. Furthermore, the definition of direct marketing would need to be clear – would this, for instance, include genuine market research activity undertaken following the Market Research Society's Code? While the general aims seem laudable, we are concerned at the possibility of unintended consequences, and would welcome further clarity on how the proposed approach would work.
73. HESA can see the benefit of introducing a 'duty to report' on communication service providers, but ensuring the reporting, monitoring, and enforcement regime remains risk-based is important to avoid over-reporting or mis-reporting. Given the provenance of many such attempts is outside the UK, it would seem prudent to extend the territorial application of PECR in line with the UK GDPR. Alignment of the PECR enforcement regime with the

UK GDPR and Data Protection Act (DPA) 2018 would improve the consistency of approach and simplify understanding and observation of the law for organisations. It would be the measure most likely to combat nuisance calls. Bringing the PECR enforcement regime in line with the UK GDPR and DPA 2018 would be a welcome regulatory simplification. However, there is little evidence that the fine acts as a significant deterrent to fraudulent and nuisance calling, so other mechanisms (such as extraterritorial applicability of PECR) would need to be considered.

74. The benefits of mandating communications providers to do more to block calls and text messages at source mainly focus on the reduction of fraudulent, scam, and nuisance calls, which not only affect some of the most vulnerable data subjects directly, but may also sour public attitudes towards legitimate activity, impacting on their efficacy. For HESA there may be benefits to the production of representative official statistics based on high response rates, if users generally believe that calls are less likely to be scams. There may, however, be risks to data collection if legitimate calls and texts are blocked through overzealous or too-narrow application of rules, with the resulting over coverage damaging legitimate activity.
75. There may be a benefit in a user having control over which calls can be received through a free of charge blocking service. This may be especially helpful for helping individuals who may be particularly vulnerable to fraudulent calls, scams, or other forms of abuse. However, there may also be risks in inadvertently cutting off individuals from legitimate sources of help and support, and in limiting legitimate activity that may be of benefit to the individual. For HESA, we can see the potential for this sort of service making it much more difficult to collect data for our Graduate Outcomes survey, and thereby damaging the quality and availability of official statistics and planning data.

CHAPTER 3 - BOOSTING TRADE AND REDUCING BARRIERS TO DATA FLOWS

76. HESA welcomes discussions about ways of improving regulation that support organisations to import and export data straightforwardly, while at the same time following the high standards of data protection for which the UK is known, and which protect its citizens in exercising their fundamental rights.
77. UK universities rely on the ability to transfer data with the EU in order to conduct research, for example in disciplines such as medical and pharmacological research, and to support education and mobility experiences. The EU's positive data adequacy decision was welcomed by the HE sector, as it allows research partnerships and other activities reliant on personal data flows to continue without further safeguards being put in place. When sharing data in other territories, universities expect to take a risk-based and pragmatic approach, utilising expertise to achieve compliance with the UK's standards. Retaining the EU's adequacy decision will therefore be important in permitting research collaborations across the continent to flourish, and for the UK to retain and extend its role as an acknowledged 'science superpower'. This would likely be the case whether UK association to the Horizon Europe programme is initiated as planned, or whether the Government follows its back-up plan of spending the monies earmarked for Horizon on domestic funding regimes.
78. The aspiration of the consultation for achieving the safe flows of data across a wider range of jurisdictions is welcome. More detail would be required to understand what the impacts

of the proposals would be in detail. These arrangements could theoretically be helpful in supporting the establishment of educational and research partnerships. We would need to understand more about how the proposals for monitoring and review of adequacy decisions would be managed, and the factors that would be taken into account in making them. Opening up new territories for easier data flows is in our view valuable when it is additive and has the effect of extending the existing realm within which data can flow safely, without jeopardising existing adequacy arrangements. We note, for example, the new Chinese Law PIPL²³ bears many similarities to the current GDPR. Therefore, research partnerships in China will likely also effectively require the maintenance of adequacy with the EU.

79. We welcome exploration of how Advanced Information Technologies for Management (AITMs) and other mechanisms such as the forthcoming international data transfer agreement (IDTA) can be used to support the ambition of simplifying and assuring the lawful transfer of data outside of the existing EU/UK GDPR adequacy regime in a manner that protects the rights and freedoms of data subjects. Formal recognition of AITMs by the Secretary of State could potentially offer welcome flexibility.
80. Given the sensitive personal information that HESA processes, we would anticipate operating a privacy management programme that conforms to the highest standards. In a scenario where accreditation becomes necessary to support the continued pursuit of established data relationships, we would anticipate valuing and seeking such an accreditation, and note that this would likely bear an increased cost of business for us.
81. We would welcome clarity on scenarios in which the repetitive use of derogations would be beneficial – this part of the consultation was not clear to us.

CHAPTER 4 - DELIVERING BETTER PUBLIC SERVICES

DIGITAL ECONOMY ACT

82. We welcome Government's commitment to improving public service delivery for businesses as well as individuals, through extending the public service delivery powers available under the Digital Economy Act 2017. By businesses, we hope and anticipate that the Government also includes the provision of services to third sector organisations such as regulated and exempt charities, which could also benefit from the provision of better public services.
83. Powers to share information under the Digital Economy Act (DEA) are subject to useful safeguards that help support public confidence. We believe that retaining these safeguards remains important. HESA offers a range of data-supported services to HE providers across the UK, both directly, and through our data processing partners. Under the proposals we perceive the possibility that some aspects of our work might potentially be brought within scope of the DEA 2017. We would be interested in understanding more about how the Government sees these proposals operating, beyond the couple of good examples provided. We perceive a potential fruitful interplay between proposals on enhancing and recognising the role of data institutions or intermediaries elsewhere in this consultation, and the proposals for enhancing the DEA. Organisations such as HESA would be well-placed to utilise such a framework to extend the efficiency and scope of data-driven public services to a wider range of organisations and offer the strong governance and accountability

²³ <https://www.wired.com/story/china-personal-data-law-pipl/>

frameworks necessary to retain the confidence of data subjects and other stakeholders. We would welcome further exploration of these ideas following the consultation.

RELIANCE ON LAWFUL GROUNDS OF PUBLIC BODIES

84. As a data controller undertaking public tasks and with legitimate interests of its own alongside responsibility for fulfilling legal obligations on behalf of various statutory bodies, HESA believes establishing our own lawful grounds for processing is an important part of our due diligence. Given the nature of our work we usually identify a public task that is complementary to the lawful basis identified by our statutory customers. In rare cases where we seek to utilise the lawful ground of a statutory body only, without identifying a public task or other lawful basis for processing that applies to ourselves, we have always regarded HESA as a data processor and not a controller. In practice, the requirement for HESA to undertake this due diligence allows us to demonstrate good data governance, and where we encounter issues, this can be an indicator to one of our statutory customers that a proposed data collection activity breaks fundamentally new ground and that the administrative burden on the HE sector may require stronger justification than they had at first perceived. We understand our stakeholders find this service helpful. Scrutiny has never prevented a necessary data collection but has generally resulted in better (and more widely supported) outcomes than would otherwise have been achieved. We therefore do not believe this proposal is necessary on practical grounds. The ICO have outlined their view on safeguards, which we defer to. However, we do not think the proposal has sufficient value to be worth introducing.

TRANSPARENCY MECHANISMS FOR ALGORITHMS

85. Our experience is that even experienced data users can find it challenging to understand the detail of how information is manufactured through data business processes, and that more could be done to support and educate the public on these matters. While AI is often the focus, there is a good deal of straightforward algorithmic processing work involved in data production that is also relevant to maintaining public trust and confidence. Making algorithmic definitions available is extremely helpful in revealing the decision-making process. We recommend they are accompanied by various metadata, as well as data definitions and explanations to help users understand how the business process leads to the decision.

86. In our experience, while most users perceive the value in curating metadata, making the business case to support its production is often quite difficult, because the benefits are not often felt by those shouldering the burden of production. Metadata manufacture is therefore often seen as an activity that is difficult to justify against other more immediate data priorities.²⁴ Recognition of the value of metadata in law would enhance the compliance aspects of business cases for metadata production. There would also be value in Government leadership on establishing harmonised standards for metadata description and production, to make the systematic cataloguing of such information efficient and sustainable.

87. Metadata production and data standards work more generally are some of the key features of the emerging class of organisations referred to in the consultation as 'data

²⁴ <https://dataarchitecture.blog.gov.uk/2019/10/07/why-metadata-is-important/>

intermediaries'. Supporting the open data publication of metadata descriptions of important datasets could be a useful contribution such bodies could make to the data economy. HESA would welcome further exploration of how placing these responsibilities on a statutory footing might engender economic and public service benefits that go beyond the accountability issues that rightly feature in the consultation proposal.

88. HESA already follows a practice of transparency in publishing the variables we engineer from our data collections, as 'Derived Fields'. Users find this information helpful and useful, however, user needs are quite varied, and while some users would like us to publish code, others prefer human-readable technical descriptions, while still others require background information on the choices and judgements involved. Some exploration of minimum mandated standards, plus a risk-based materiality test for more extensive documentation, would be welcome.
89. Considering mechanisms for supporting algorithmic transparency, the following should ideally be put in place (though undoubtedly in practice not every one of these will always be achievable):
 - Use of a machine-readable and human-readable metadata description for all important data items and algorithms
 - Use of a clear and standardised versioning system.
 - Established standards for metadata production have been observed.
 - Application of a materiality test for when reporting is required, and identification of what data items are considered important.
 - Paradata (data that describes an aspect of the operation of a system) is also described where it is material to a decision-making process.
 - A business justification for production, based on the specification of user needs, has been produced.
 - Clarity about the data governance arrangements for transparency reporting, which recognises the importance of stewardship of all those involved in the production process, as well as data users.
90. National security, and possibly law enforcement may offer grounds for democratic oversight of algorithms to take place in private. In such circumstances Government would need to consider how democratic representatives would secure independent professional advice on matters with privileged access to data. We can identify no other areas within the general processing regime established in DPA 2018 (the UK GDPR) where transparency reporting would be completely inappropriate, but pragmatically, the production of transparency information can be complicated, especially where older software, proprietary datasets, or paradata are used.
91. We recommend exploring a materiality test that:
 - a. Identifies what constitutes a decision for the purpose of the duty.
 - b. Relates the scope of the transparency information production to the public interest being served.
 - c. Identifies what standard(s) should be followed to serve different types of public interest (perhaps understood as 'user personas' representing classes of user needs that will be addressed by the duty).
 - d. Offers tests for when an algorithm or data item would be considered out of scope for the duty.

- e. Makes clear the date from which all new decisions must conform to the standard., and
- f. Makes clear the date by which all decisions within scope must conform to the standard. This second date should offer adequate time for public authorities to update or replace their systems and processes to meet the standard.

PUBLIC SAFETY AND NATIONAL SECURITY

92. Intellectual capabilities and knowledges in areas of basic and applied science as various as nuclear physics and vaccine development have been cited as essential to national security both now and in the future. Many of these capabilities are located within the UK's world-class HE sector. This has been recognised in warnings from the National Cyber Security Centre (NCSC) about state-sponsored attacks targeting researchers and universities, and this is an area where the HE sector and Government has invested to harden its critical infrastructure. HESA urges the Government to consider the potential direct and indirect impacts of legislative changes on the HE sector.

CHAPTER 5 - REFORM OF THE INFORMATION COMMISSIONERS' OFFICE

- 93. HESA values the ICO and the strong role it plays in safeguarding and promoting high standards in the UK data economy.
- 94. We believe independence from both Government and other special interests is important to the good governance of data and information generally. It is right that independence in the conduct of public functions should come at the price of enhanced scrutiny.
- 95. The considerations offered by the ICO in its response to this section of the consultation resonate with our own approach, and we are therefore supportive of their analysis and proposals.



Dan Cook
Deputy Director, Data & Innovation
HESA
95 Promenade
Cheltenham
GL50 1HZ
E dan.cook@hesa.ac.uk
W www.hesa.ac.uk

