Data governance for online learning

ODI Fellow research report

Open Data Institute

Contents

About

This report was produced as part of the ODI research fellow scheme. Its author was Janis Wong, a PhD Researcher at the University of St Andrews, with contributions from Mahlet (Milly) Zimeta, ODI.

If you want to share feedback by email or would like to get in touch, contact Janis at jccw@st-andrews.ac.uk.

If you would like to learn more about the ODI research fellow scheme, please visit the information and application page, or contact us at fellowships@theodi.org.

Executive summary

Data governance considerations and evidence for higher education institutions working with online learning data

The Covid-19 pandemic has increased the adoption of technology in education by higher education institutions in the UK.[1] Although students are expected to return to in-person classes, online learning and the digitisation of the academic experience are here to stay. This includes the increased gathering, use and processing of digital data.[2]

These changes mean a rethink of traditional data governance strategies for education is needed, to adapt them for the online learning environment. Although higher education institutions have long considered how digital technologies can be used to help shape the future of learning, little consideration has been given to data governance of online learning and how other forms of educational data impact the academic experience of students and staff in the long term.

Innovative and forward-thinking data governance strategies should take into account the collection and processing of student and staff administrative data, as well as teaching data; explore how digital tools are used to benefit learning; and look at what the data archival practices are at an institutional level. Such governance strategies create new opportunities to improve the value of data within higher education institutions.

Data governance in online learning

In order to create an online learning strategy that is fit for the future of digital education, university management needs to consider how online learning data should be governed. This includes:

- governing research data as well as student and staff personal data related to teaching, administration and the data processed by external platforms
- addressing the benefits to, and concerns of, students and staff as the largest groups of stakeholders in delivering secure and safe academic experiences
- deciding what education technology platforms should be used to deliver, record and store online learning content, by comparing the merits of improving user experience against potential risks to vast data collection by third parties
- balancing education innovation with the protection of student and staff personal data through data governance, management and infrastructure strategies.

[1] Office for National Statistics (2020), 'Coronavirus and the impact on students in higher education in England: September to December 2020'.
[2] The Observer, Fleming N (2021), 'After Covid, will digital learning be the new normal?'

Online learning data governance needs to be considered holistically, with an understanding of how different stakeholders interact with each other's data to create innovative, digital means of learning.

Introduction

Context

In our data-driven society, digital technologies are increasingly becoming a part of our everyday lives. This includes education. The framing of educational data has been described by Borgman as 'grey data',[3] where the lines of research, teaching, learning, services and administration data blur in the context of universities, raising questions around privacy and data governance. Academic data collection has become more complex due to the open access requirements of institutions, as well as the accumulation of such grey data in daily academic life. This is further complicated where private and corporate stakeholders use this data for third party purposes, such as assessing university rankings, exam proctoring or graduate surveys.

Before the Covid-19 pandemic, the challenges around data in education – specifically online learning and ethical frameworks for managing data – were already being identified. However, online learning has exponentially accelerated during the Covid-19 pandemic, with students, staff and academic institutions having to rely on technologies and educational platforms to deliver lessons, record classes and introduce new interactive methods of learning. From reports of student and staff surveillance,[4] to concerns around online learning harassment and 'Zoom-bombing' (where a person joins a Zoom meeting without invitation and aims to disrupt the session), the digitisation of online learning not only affects the academic experience, but also the relationship between educational institutions and how they collect, use and process personal data related to students and staff.

Other data-related issues involve how research data is collected and stored, how digital materials are archived, and how student and staff data is secured for purposes other than online teaching. While these data management processes are often considered separately, they all fall within the wider remit of how data, and more specifically online learning data, is governed and managed at higher education institutions.

Focus of this report

This report examines the relationship between higher education institutions and the online learning data they collect, process and share; outlining how online learning data can be better governed and managed. This includes bridging the gap between research into online learning data pre-pandemic, and the research and policy recommendations resulting from the Covid-19 pandemic. This report outlines current data governance practices and recommendations, how educational institutions can balance the need for innovative learning with privacy of the community, and new

[3] Borgman CL (2018), 'Open Data, Grey Data, and Stewardship: Universities at the Privacy Frontier'.
[4] Edwards L, Martin L, Henderson T (2018), 'Employee Surveillance: The Road to Surveillance is Paved with Good Intentions'.

opportunities for how academic institutions should govern data.

Who is this report for?

This report is for individuals and organisations such as EdTech platforms and university management that have to process, manage or govern data, as well as policymakers who have to provide guidance on data governance or develop data infrastructure strategies related to online learning. This includes:

- university and higher education management
- education policymakers
- technology platforms that create digital services for higher education institutions.

This report can be used to inform product development, data management and new data governance strategies.

How should you use this report?

This report is designed to help you and your organisation understand what online learning data is, why online learning data should be considered more widely as part of data governance, and how you can go about incorporating online learning data considerations into a data governance strategy.

Each section and subsection has summaries, action points, key points and concluding remarks, to help you identify the most important information and guide your next steps.

Links and resources have been included throughout this report, for more in-depth information on each topic.

Definitions

To clarify the scope and application of this report, the following are definitions and how we use specific terms:

Academic data

Academic data refers to data collected, processed and used by higher education institutions that relates to the academic element of education. This includes examination and coursework data, and research data.

Data governance

Data governance refers to the technical, legal, policy and management considerations that support data within specific contexts and environments. (See also the *Data governance and online learning data* section)

Data protection

Data protection in this report refers to the protection of personal data regulated under the UK Data Protection Act (DPA) 2018 and the EU General Data Protection Regulation (GDPR) 2018.

Digital technologies and platforms

Digital technologies and digital platforms are online tools, applications and services that enable the collection, distribution and sharing of information.

Education technologies (EdTech)

Adopted from the UK Department for Education,[5] education technology (EdTech) refers to the practice of using technology to support teaching and the effective day-to-day management of education institutions. It includes hardware (such as tablets, laptops or other digital devices), and digital resources, software and services that aid teaching, meet specific needs, and help with the daily running of education institutions (such as management information systems, information sharing platforms and communication tools).

Higher education institutions

Higher education (HE) institutions include universities or an institution conducted by a higher education corporation in the UK. For the purposes of this report, the terms universities and higher education institutions are used interchangeably.

Online learning

Online learning is learning and education that is conducted in a digital environment and facilitated using digital platforms.

Online learning data

Online learning data covers all aspects of data that is collected, stored and processed within higher education institutions. This includes teaching material, research outputs, lecture recordings, student administrative information and examination results. (See also the *What is online learning data?* section)

It is important to note that online learning data can have been in a non-digital format in the past, and so although the same type of data may have previously been gathered, the digital processing of such data has allowed it to be more easily accessed, managed and retrieved.

Personal data

Following the UK DPA and EU GDPR, personal data refers to information relating to natural persons who can be identified or are identifiable both directly and indirectly from different information sources. (See also the *What is online learning data?* section)

Surveillance

Surveillance refers to the observation or supervision of a group, with or without the group knowing that they are being observed, usually with the suspicion of malpractice or with the intent to prevent malpractice. Digital surveillance uses technologies to conduct this observation.

[5] UK Department for Education (2019), 'Realising the potential of technology in education'.

Data governance and online learning data

In order to examine how online learning data should be governed by higher education institutions, we need to understand what types of online learning data there are and how such data is created or used, and explore broader questions surrounding data governance for online learning.

Summary

- In higher education institutions, different stakeholders interact with and create online learning data for different purposes.
- There are many types of online learning data, including teaching data and related materials, administrative data, and research data.
- Online learning data is managed independently at different stages, but it is important to identify how online learning data follows through the education process and changes over time across different stakeholders.
- When considering how online learning data should be governed, strategies should be targeted towards specific stakeholders and data flows.

Stakeholders

Before identifying the relationship between higher education institutions, online learning data and data governance, it is important to first identify the stakeholders involved. As well as students, other internal and external stakeholders playing key roles are also involved. Internal stakeholders include teaching staff, research staff, university management, and librarians and archivers. External stakeholders include technology providers and higher education advocacy institutions.

In the table below, we list each relevant stakeholder within the online learning sector, summarising their main motivation and actions in relation to online learning.

STAKEHOLDER	MOTIVATION	ACTION
HE students	Study and make the most of resources that are available in the online learning environment	Learning to use and access online resources through traditional and new online learning platforms such as emails, moodle pages and video platforms
HE teaching staff	Provide the best possible learning	Learning to use traditional and new

	experience to students using the resources and equipment available to them in the online learning environment	online learning platforms as allowed by their institution to deliver teaching and materials
HE research staff	Conduct the best possible research and present their research	Using the most suitable platform for publishing and sharing their outputs and research data as well as archiving their work with their institution
Librarians and archivers	Enable staff and students to conduct their best work through library resources and data management techniques	Providing support for staff and students in order to find the best resources to use in their work as well as the best ways to store and access their data for dissemination and retention
Data protection officers	Protect student and staff personal data	Keeping up to date with data protection regulations, implementing those regulations in university policies, and responding to data protection requests where necessary
University management	Support their university's strategy and promote their institution to external students, staff and organisations	Create data management and data governance strategies for their institution
Educational technology platforms	Provide technology capabilities to facilitate online learning	Develop applications to be used within existing university infrastructures for online learning
Higher education advocacy groups	Advocate for the position of their members within education	Establish policies that reflect the benefits and challenges in online learning and education for their members
Education policymakers	Create supportive, competitive, long-term strategies for online learning and education for higher education in both local and international contexts	Engage a wide range of stakeholders to provide input on developing an education strategy that best reflects innovative and forward-thinking considerations for online learning

Key points

- There are a number of stakeholders in online learning, many of whom do not interact with each other during their journey through education.
- All stakeholders play an active role in their area for advancing their experience in education or the experience of their target audience.
- Understanding their motivations and actions can help explain their relationship with you or other stakeholders, as well as their data.

What is online learning data?

Online learning data can be split into three main areas: teaching data and associated materials, administrative data, and research data. Teaching data and its associated materials refers to the data that is used for the purposes of teaching to fulfil academic purposes. Administrative data refers to data that is used to keep a record of student and staff roles at the university and to track their progress. Research data refers to data that is collected for the purposes of research. Although research data itself may not play a lead role in online learning, the processes and policies by which research data is managed and stored can affect other types of online learning data. It is important to note that these types of data and how they are created or used by stakeholders is not hierarchical, and they all have important functions as part of the online learning environment.

Identifying what constitutes online learning data is important because it helps in understanding the complexity of governing data in digital education. A successful data governance strategy for online learning must consider all forms of data and how they are used by various stakeholders.

Key points

- Online learning data is varied and constantly changing.
- Understanding what types of online learning data there are can help better govern these forms of data, opening up new opportunities for both innovation and management.

Online learning data flows

To illustrate how different types of online learning data are created and passed through the online learning ecosystem, we have mapped out the data flows to demonstrate how different stakeholders interact with the same type of online learning data.

The online learning data flows can be mapped using the ODI Data Ecosystem Mapping tool and Data Landscape Playbook. The tool identifies the data, the stakeholders related to the data, and their relationships. The types of data and stakeholders are identified, and it is recognised that stakeholders and their roles may overlap or cross-over.

Online Learning Data Flows

Creating teaching data:

- Teaching staff: deliver
- Students: contribute
- University management: facilitate

Recording teaching data:

- Teaching staff: record

- Students: access
- University technology: record
- University library/archiver: preserve

Storing research data:
- Research staff and students: manage
- Administrative staff: manage
- University technology: manage
- University library: manage
- Data protection officer: protect

Publicising teaching data and research publications:
- Research staff: share
- Students: share
- University technology: share
- University library/archiver: share

Student administrative data:
- Students: provide
- University technology: collect and delete
- University management: reuse and generate
- University library/archiver: delete

Student examinations and academic data:
- Students: provide
- University technology: monitor

All online learning data:
- Data protection: govern
- University management: govern and deliver

Following the Data Ecosystem Mapping tool, these types of data and their stakeholders can be sorted into more generalised categories:

- **Data** representing teaching data, academic data, research data and administrative data.
- **Goods** representing academic data resources.
- **Money** representing direct transfer of finances in exchange for another form of data or service.
- **Services** representing the platform or tool for which data can be utilised.

These data flows can be mapped into a diagram based on the Data Ecosystem Mapping tool, demonstrating the relationship not only between stakeholders and data, but also how this could affect relationships between the stakeholders themselves.

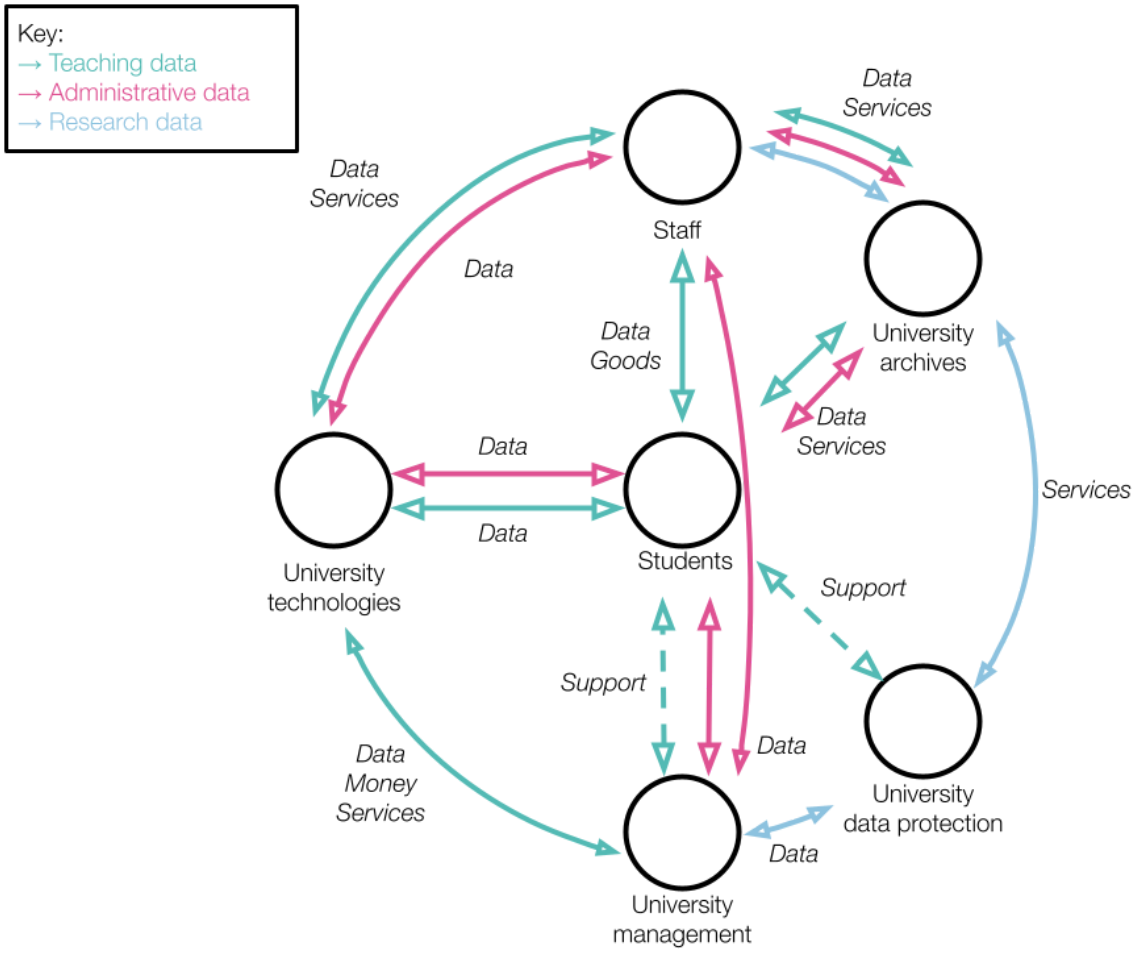

Figure 1: The primary types of information (data, goods, money and services) flow between different stakeholders. All forms of data are seen as two-way flows, with the dotted lines representing indirect relationships.

Key points

- In order to better govern online learning data, it is important to understand what the data ecosystem is like for the online learning and higher education ecosystem.
- Mapping the online learning data flows can help clarify how data travels and transforms throughout the institution, and highlight which areas may need support.

Data governance of online learning data

Data governance of online learning data includes technical, legal, policy and management considerations that help support the online learning data flows within higher education institutions.

Good online learning data governance should outline how different types of online learning data affect the institution and how the data should be protected; illustrate the benefits to the inclusion of certain technologies and data management methods; respond to stakeholders' concerns; and suggest practical solutions for how the

institution can balance innovative implementation of digital strategies with limiting the risks to additional data collection and processing.

Considerations for online learning data governance

Summary and action points

- The shift to online learning has significantly impacted how institutions interact with student and staff personal data, with the increase of data gathering, service provision, and transfer to digital resources by higher education institutions.
- This includes changes in stakeholder relationships with their personal data and the increasing importance of data protection and privacy in the digital classroom.
- When creating data governance strategies for online learning, these issues must be taken into consideration in order to establish trust between stakeholders and trust in how institutions use their data.
- Although higher education institutions and their staff members have been making the most of technological tools to limit the negative impacts of the pandemic during the early introduction of online learning, more long-term data governance strategies for online learning are needed to better support institutional data practices.
- When considering how online learning data should be governed, higher education institutions need to better understand how data flows have changed as a result of online learning, and determine to what extent online learning will continue in order to better support the university community.
- Policymakers should address how data infrastructures will be strengthened and how education data should be governed in light of the rapid acceleration of online learning.
- Current technology providers and new EdTech companies need to understand how institutions use their platforms and how data flows through institutions to deliver better services to students, staff and universities.

The impact of moving from physical classrooms to online hybrid learning

Prior to the Covid-19 pandemic, higher education institutions had been incorporating technology into their teaching, learning and research practices. From the introduction of email, to the use of laptops by students, technology has allowed for more flexible ways of working and fostered more methods for collaboration and sharing of information. However, it is important to identify how the digitisation of education and increase of online learning data has also impacted stakeholder relationships with education and their institutions.[6] As blended learning is set to continue, in order to

[6] Dhawan S (2020), 'Online Learning: A Panacea in the Time of COVID-19 Crisis. Journal of Educational Technology Systems'.

accommodate the impact of the pandemic, students are beginning to express dissatisfaction in their online learning experience, with university leaders often not consulting them on future plans. In this report, although we identify some of the new challenges that have emerged as a result of adapting to online learning through Covid-19, we recognise the efforts that students, staff and institutions have all made in transitioning to the digital classroom at short notice. This section outlines the impact of moving from physical classrooms to online learning, both pre-pandemic and during the pandemic.

While the full effects of moving to online learning have yet to be seen, some research has been done to assess the impact on students and staff at higher education institutions. As education is transferred online, it has been shown that over-reliance on lecture recordings may negatively impact student attainment.[7] During the pandemic, 74% of students report that exams and assessments are now being carried out online, with 27% reporting that some marks are being based on previous assessments.[8] For staff, chilling effects of employee surveillance in the classroom have also been identified.[9] Staff may also be less willing to teach certain topics, especially if a class is recorded, and may limit the time dedicated to general discussion. Research has recently been conducted to map how higher education institutions have implemented cloud infrastructures, comparing institutions' level of migration to public cloud services.[10]

As a result, when considering privacy-preserving solutions in education, in order to ensure that student privacy frameworks align with students' digital practices, stakeholders should incorporate robust ways for students to participate in the discussion about the best ways to tackle outstanding and future student data privacy challenges.[11]

To consider how online learning has changed the relationship with online learning data and to illustrate the impact of this, we adapt each type of data below:

Creating teaching data:

- More online teaching data is being created
- All classes are happening online
- New technologies introduced in a short timeframe to enable online teaching

Recording teaching data:

- Most if not all teaching data is recorded
- Potentially stored in different ways, even if teaching was previously recorded

Storing research data:

- More research data being stored online

Publicising teaching data and research publications:

---

[7] Edwards M, Clinton M (2019), 'A study exploring the impact of lecture capture availability and lecture capture usage on student attendance and attainment'.
[8] Edge Foundation (2020), 'The impact of Covid-19 on education: A summary of evidence on the early impacts of lockdown'.
[9] Edwards L, Martin L, Henderson T (2018), 'Employee Surveillance: The Road to Surveillance is Paved with Good Intentions'.
[10] Fiebig T, Gürses S, Gañán CH, Kotkamp E, Kuipers F, Lindorfer M, Prisse M, Sari T (2021), 'Heads in the Clouds: Measuring the Implications of Universities Migrating to Public Clouds'.
[11] Plunkett L, Gasser U, Cortesi S (2019). 'Student Privacy and the Law in the Internet Age'.

- Greater emphasis on digital sharing due to lack of physical events
- New medium for sharing explored due to international and geographical differences and allowing for as much sharing as possible
- Greater shift towards openness, both in terms of open access publication and in the dissemination of restricted information

Student administrative data:

- More stored as part of online teaching; data derived from online teaching but not part of it, for example taking attendance
- ID checks being conducted online and through software and cameras, rather than in person
- Student feedback all processed online
- Student services such as physical and mental health; any pastoral support outside of academic support

Student examinations and academic data:

- All exams are held online
- More academic data gathered online, including coursework

The impact of online learning can also be mapped onto the data ecosystem, extending Figure 1, as shown in Figure 2. This demonstrates how data flow may change as a result of the digitisation of learning and what future opportunities there are to strengthen the online learning process. Data, goods, money and services are included as direct data flows, and support as an indirect data flow. Detailed definitions of these terms can be found on the Data Ecosystem Mapping tool, but are broadly defined in the list above.

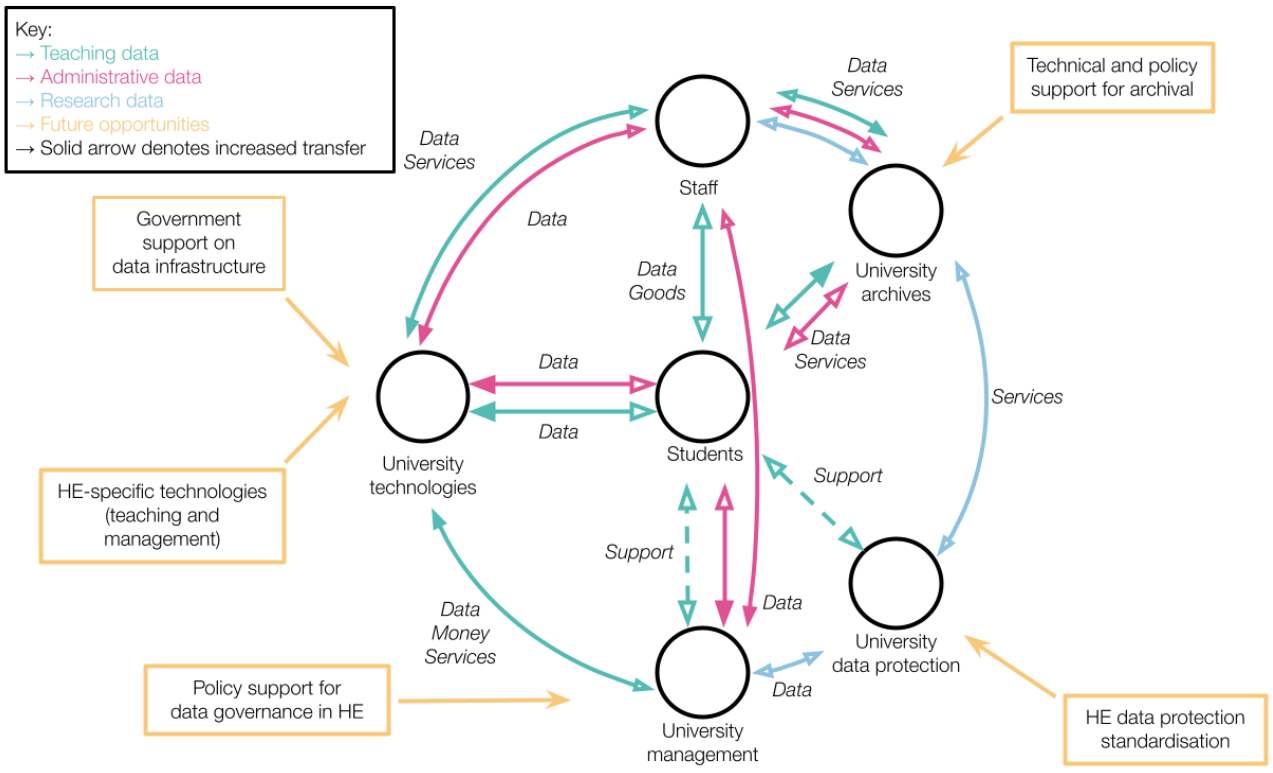

Figure 2: The primary types of information (data, goods, money and services) flow between different stakeholders, where solid arrowheads demonstrate the greater

transfer of information and data to specific stakeholders. This suggests that who, what and how data travels through the education data ecosystem has changed as a result of online learning. The boxes outside the map demonstrate the future opportunities for developing data governance strategies for online learning.

Some higher education institutions have successfully modelled their education environment to be solely digital. The main example is The Open University (OU). As the largest university in the UK, the OU has long had online distance learning courses. In the context of the Covid-19 pandemic, the OU was much better equipped to close their campus and continue online learning. This also means that the OU is in a better position to protect and manage data for education.

The OU is transparent on its website about how student and staff data is used. There is a specific page for student data and how different types of students have their data managed differently. There is also a Privacy at the OU page that outlines where data is held, what rights data subjects have, staff data protection, alumni data protection, and other data related to external organisations and corporations. Both full length documents and summary documents are provided. The OU also provides links to its strategy and policies, which have information on how data fits into its overall workings as an institution.

While the OU's policies may not be suitable for hybrid learning institutions, it demonstrates how online learning can be implemented successfully without detriment to innovative educational practices.

Key points

- The rapid transition from physical classrooms to online learning has exponentially increased dependence on digital services for delivering education, thus increasing the amount of online learning data that institutions have to govern.
- Existing research, frameworks and guidance can help support online learning infrastructures within and from outside institutions
- While this may disrupt the distribution of resources within the institution, it also highlights new opportunities for innovation, governance and policy.

Considerations: digital monitoring

Summary and action points

- When considering how online learning data should be governed, university management and institutions should identify their aims in implementing certain technologies and evaluate whether they help achieve those aims.
- Developers of platforms and technologies that deliver online learning should understand that students, staff and other university stakeholders may consider certain technological functionalities to be unnecessary, or even detrimental to academic freedom and discussion.
- Policymakers should provide guidance for data governance in relation to electronic means of assessments and examinations, taking into account the impact technology has on academic performance and progression within higher education.

Data protection and data governance for online learning

Although online learning and its socio-technical and policy impacts have been explored, there are currently no sector-wide policies for online learning, lecture recording or use of teaching platforms. The UK's Information Commissioner's Office has published The Employment Practices Code[12] on topics such as employment records and monitoring at work, to provide employers with the best data protection practices that comply with existing regulations. However, this is only suggested as a guide and is not education-specific. As a result, educational institutions have had to quickly establish working practices to accommodate new ways of learning.

As lecture capture, tutorial recording and educational monitoring start being implemented and, in many cases, become mandatory, university experience for both staff and students is affected. As classes are being recorded and monitored, students and staff may less willingly participate in them, creating a chilling effect on education. Apart from academic content, the shift to online learning also affects how institutions process data subjects' (those about whom data are collected, used and processed; in this case students and staff) data. The GDPR offers solutions such as applying fairness and transparency principles and exercising data subject rights. However, this may be insufficient where data subjects have no choice but to accept their institutions' terms or be locked out of academia. There is also a lack of consideration of group rights under the GDPR, which could affect specific classes and demographics of students and staff, further highlighting the power imbalances

[12] ICO (2018), 'The Employment Practices Code'.

between them and their institutions.

Theoretical and empirical work has found that the use of e-proctoring could further students' desire to cheat and rebel against surveillance systems. In a recent example, a student exercised their data protection rights under the GDPR to find out how their e-proctoring software works.

There is also friction between online learning and the GDPR within higher education institutions when it comes to student and staff personal data.[13] For example, it may be unclear what the roles and responsibilities of protecting personal data are; there may be a lack of transparency around the processing and possibility to effectively exercise data subjects' rights; there may be uncertainty around data transfers outside the EU; and there may be challenges to the legality of e-proctoring systems.

Online learning and technology

Digital technologies have long been implemented for online learning at universities. The Open University offers flexible distance learning for undergraduate and postgraduate degrees. Massively Open Online Courses (MOOC) offer courses on platforms such as Future Learn, EdX and Coursera by partnering with academic institutions. However, in the space of two weeks, the Covid-19 pandemic propelled universities that didn't have the technological infrastructure to run online classes for thousands of students, not only across the country, but across the globe. Online learning platforms have also experienced a spike in their reach during lockdown in the UK. This raises the question of the extent to which institutions were equipped to handle an exponential amount of data. Additionally, it also raises concerns about how data may be collected, used and processed during this period of online learning.

Although measures have been taken to ensure that the data gathered through online platforms is protected, some of these technologies have been used to monitor students and staff behind the scenes. For example, the Microsoft Office Productivity Score, running in the background of Microsoft applications such as Microsoft Teams, tracks the time and activity of its users, producing data on the extent to which individuals are working on its platform. Initially, this data could be accessed by institutions and linked to specific usernames. Only after privacy concerns were raised online did Microsoft remove usernames and change how the data gathered is presented. Even if higher education institutions do not access or use this data, it could still be collected by digital platforms and may be shared and sold to third parties. Particularly where universities are public institutions, these data collection and processing practices should be made transparent to those who use these technologies for online learning.

Home-working has also introduced new aspects of informal staff monitoring, blurring the line between home and work environments. For example, if a worker would normally log in to meetings by turning on their video, but one day they are in a car or a new location, the employer might think they are not committed or focused enough. While traditional forms of register taking or swipe card access data were collected in

[13] Angiolini C, Ducato R, Giannopoulou A, Schneider G (2020). 'Remote Teaching During the Emergency and Beyond: Four Open Privacy and Data Protection Issues of 'Platformised' Education'.

physical classrooms, the shift to online learning introduces new questions around the ethics of care related to online and remote work. New organisational conditions need to be addressed to support institutions' duty of care for the health and safety of their community, without necessarily needing to deploy new technologies or conduct long-term monitoring.

Lecture recording and remote proctoring

To accommodate for students and staff working across the globe due to the pandemic, lecture recording has been more widely adopted to ensure the access of content across time zones and provide academic resources to accommodate for potential absences. While lecture capture may be beneficial, the recording of lectures and tutorials may also have an unsettling effect on staff and student participation in classes, affecting their content and contributions. More generally, with online learning being accessed in the home environment, the pandemic has resulted in a shift from workplace monitoring to taking surveillance technologies home, for both staff and students. It may also encourage more informal monitoring, such as assumptions about a person's commitment or attention to work based on whether they have their camera on. There may also be extraterritorial harms where politically sensitive topics may be taught. Despite these challenges, it is important to note that staff and institutions use lecture capture and recording tools as a means to continue their responsibilities for providing students with the means to study during the pandemic. As we transition away from using these technologies, institutions should reassess the purpose and usefulness of these tools and whether they enhance learning and teaching.

In addition to lecture recording, higher education institutions have also introduced online proctoring and remote proctoring services as part of online learning, to monitor students during tutorials and exams as a means to track engagement and prevent potential malpractice. While the aim of proctoring technologies may be to ensure that students make the most out of their online learning experience and level the playing field when it comes to assessments, the use of these technologies could result in academic and data-related risks. Remote proctoring services can negatively impact students where it requires access to technology that not all students are guaranteed to have; can constitute an invasion of privacy for students; and can discriminate against students of colour and disabled students.[14] In order to balance these risks, institutions could consider scrapping proctoring systems in favour of a people-centred approach, through investing in instructional design staff and faculty development programming to help transition to authentic assessments.

[14] Silverman S, Caines A, Casey C, de Hurtado BG, Riviere J, Sintjago A, Vecchiola C (2021), 'What Happens When You Close the Door on Remote Proctoring? Moving Toward Authentic Assessments with a People-Centered Approach'.

Considerations: data protection and privacy

Summary and action points

- When considering how online learning data should be governed, university management and institutions should consider data protection and privacy as part of the data governance process, particularly as stakeholders spend more of their university experience online.
- Developers of platforms and technologies should work with institutions to establish clear data protection practices that follow university guidance and codes of conduct.
- Policymakers should provide guidance and standards for data protection and privacy to ensure more consistency in how institutions across the country can better protect their stakeholders.

Technology

Developers of digital platforms have responded to privacy and data protection concerns as their technologies have scaled and become staples for online learning. Following incidents of Zoom-bombing, all Zoom meetings now require a password to join unless attendees are joining by directly clicking on a meeting link with an embedded password. Zoom has also committed to implementing end-to-end encryption for all users; however this was challenged in a class action privacy lawsuit. Microsoft Teams now allows users to create private channels on their platform and IT administration has been integrated into the platform. Microsoft has also committed to supporting end-to-end encryption for one-to-one Teams calls.

There have also been crowdsourced, public efforts for improving the ethical data and technology landscape during online learning. For example, online documents such as A Comprehensive Guide to Tech Ethics and Zoom Class and the Coronavirus Tech Handbook were created to support students, staff and researchers in making ethical and data protection-friendly decisions when it comes to online learning.

Platforms

As more privately owned technologies are used for online learning, the digitisation of higher education may result in further commercialisation of public education, with limited benefit for students and staff.[15] Many of the online learning platforms adopted by higher education institutions were not created with education as their primary

[15] Williamson B, Hogan A (2021), 'Pandemic Privatisation in Higher Education: Edtech & University Reform'.

purpose, but for business-oriented cloud services. These technologies were implemented by higher education institutions due to existing paid licensing, limited alternatives, and relative ease of adoption due to already being part of the digital ecosystem. Institutional barriers remain higher for adoption of new technologies for online learning due to the need for scale and accessibility from global cohorts of students. The cost of adopting and using these platforms is also shifted to students and faculty, rather than being internalised by universities and management.

Another consideration with the shift to online learning is how intellectual property and information rights are managed for digital materials. Although the digitisation and increased openness of data and research dissemination through libraries and public means are encouraged, where copyright should not be used as reason to paywall a publication, caution must be shown when repurposing materials that are used for teaching and learning.

Intellectual property and ownership of teaching materials remains a challenge with the shift from physical to online learning. In an extreme example, during Covid-19, a student tweeted that when attempting to send their professor an email, they discovered that the professor delivering their recorded lecture was in fact dead. Where classes are recorded on EdTech platforms, it is unclear who has intellectual property rights to the content. Further, pre-Covid, academic institutions have been caught repurposing academic content for teaching without the consent or knowledge of the staff who created the teaching materials in the first place. For students, there have been long-standing questions about the ownership of intellectual property over essays submitted through plagiarism detection platforms. FAIR (Findability, Accessibility, Interoperability, and Reuse) use often covers what and how resources are to be used for academic purposes, but the shift to online learning and increased use of digital learning technologies have introduced new examples that not only demonstrate issues related to how data protection may be neglected where the collection and use of data may not have been transparent, but also highlight the potential mismanagement of educational data in the digital environment.

Institutional policies for online learning

From an institutional perspective, organisations have written guides to promote best practices for online learning and offered guidelines for adapting institutional policies for online learning. For example, JISC has written guides for recording lectures[16] and the legal considerations required under copyright and data protection legislation. The organisation has also worked alongside the National Union of Students to provide a code of practice for learning analytics.[17]
The UK Department for Education has also outlined how cloud services and educational applications[18] should be governed following the UK's DPA, identifying schools' data protection obligations to be: overarching legal requirements, data processing, data confidentiality, data integrity, service availability, considerations of transfers beyond the EEA, and concerns related to the use of advertising. While this publication focuses on schools, the approach could be applied to higher education

[16] JISC (2020), 'Recording lectures: legal considerations'.
[17] JISC (2018), 'Code of practice for learning analytics'.
[18] Department of Education (2014), 'Cloud (educational apps) software services and the Data Protection Act: Departmental advice for local authorities, school leaders, school staff and governing bodies'.

institutions due to the similar nature of content delivery and provision.

Governing online learning data

Summary

- Covid-19 has significantly sped up the digitisation of learning and education. It is now necessary to think beyond the pandemic and strengthen data governance for online learning in the long term.
- E-proctoring and similar monitoring technologies need to be properly governed through institutional data protection practices and not have the responsibility passed on to technology providers.
- Data protection must be considered as part of the data governance strategies in online learning as it plays a key role in increasing transparency and trust between staff, students and institutions. This goes hand-in-hand with government investment in digital learning and teaching infrastructure, ensuring that new online learning pedagogies include robust data protection and data security considerations.

The impact of Covid-19

With the accelerated shift towards online learning and the knowledge gained through the implementation of online learning in 2020, organisations and institutions have conducted empirical research to identify how good online learning practices can be established for the future. This section provides an overview of some of the research and potential guidance on strengthening data governance for online learning in the long-term, identifying the opportunities for advancing the adoption of digital technologies for higher education, while considering stakeholders' potential data-related concerns.

In considering the future of online learning post-pandemic, JISC has provided four key recommendations for the 2021/22 academic year for the digital content and delivery of education:

1. Individual institutions need to strategically plan the digital transformation of learning and teaching.
2. The HE sector should develop new economic models for developing digital teaching materials.
3. Universities should move to blended learning, with students co-designing curricula.
4. And lastly, the government, universities and funders should combine to tackle digital poverty. While teaching and learning through the digital environment has helped protect staff and students from health risks, it has also resulted in education that can take place at any time and anywhere, which can make communication itself more difficult in the online space.

Regarding the deployment of technologies to facilitate online learning for the future, higher education institutions should consider expanding the deployment of digital infrastructures more broadly, creating online opportunities that enhance the academic experience as a whole. This not only encourages creative and innovative thinking when it comes to adopting technology for online learning, but facilitates wider discussion about how data can be better controlled, supported and protected at an institutional level. As students and staff become more aware of how their data is being used by their institutions, university management also has to consider how data and data protection policies are communicated to them in order to facilitate trust in online learning and the digital university experience.

Digital monitoring

When considering whether to deliver online learning and, if so, what platforms should be used, higher education institutions need to balance the need for widespread accessibility and scale with other concerns such as privacy and ethics under existing guidance and regulations. Rather than being governed by technology, higher education institutions should actively create standards and ethical guidelines for online learning with evidence of staff and student impact.[19] If e-proctoring and monitoring technologies are found to be ineffective, detrimental to wellbeing, or not fit for purpose, educational institutions should stop implementing them as soon as possible, rather than waiting for licences to elapse or hoping that the technology will improve in the future. If these technologies are used, institutions have a responsibility to explain how and why they are being implemented, what the platform's purposes are, and what the platform will not be used for should there be additional functionalities.

Assessing the impact and ethical challenges of online learning, The Berkman Klein Center for Internet & Society examined the ethical, human rights and societal aspects of digital transformation.[20] The considerations include assessing how offline and online forms of surveillance are different, the boundaries of agreements between institutions and technology companies, and the impact of surveillance on the social contract between students and parents, and teachers and schools.

Data protection and privacy

In addressing data protection and privacy in relation to online learning, much work has been done examining the impact of the digitisation of primary and secondary schools. The ODI has examined the data about children's lives in a pandemic.[21] For children, their families and teachers, data needs to be inclusive and represent the needs of those with disadvantaged socio-economic status. The data needs to be shared in a safe way to enable analyses to inform organisations, policymakers and schools in their work with children, families and teachers. The project highlighted the need to build a better data infrastructure within the education sector, and has demonstrated how real-time, open, shared data can inform the response of schools

[19] Williamson B, Hogan A (2021), 'Pandemic Privitisation in Higher Education: Edtech & University Reform'.
[20] Berkman Klein Center (2020), 'Digital Ethics in Times of Crisis: COVID-19 and Access to Education and Learning Spaces'.
[21] ODI (2020), 'Data about children's lives in the pandemic'.

and other organisations to the various challenges posed by the pandemic. Similarly, an ODI report on data about teachers' lives in a pandemic demonstrated that reliable, systematic data on the wellbeing and professional views of teachers can support better learning outcomes for teachers and students.[22]

In The State of Data 2020 report[23] Defend Digital Me notes that the digitisation of education can result in children losing their ability to control their online footprint once they start school. Data protection alone may not be adequate and so the government and policymakers should work to build better data infrastructures at national, regional and local levels, and give it independent oversight through a statutory ombudsman to champion children's rights for national data in education.

Focusing on the future of digital learning and teaching, JISC produced a report on the response to Covid-19 based on over 1,000 stakeholder experiences.[24] Looking at the projected short-term and long-term inclusion of online learning, JISC have provided recommendations for universities, sector agencies and government that include strategic investment in digital learning and teaching, prioritising blended learning by rethinking the scale and scope of learning and teaching activities with students, and including digital training, peer support mechanisms and reward and recognition incentives to encourage upskilling in professional development plans.

Models for the governance of online learning data

Digital technologies and platforms will remain in educational spaces even as physical classrooms return. Therefore, educational institutions must plan for the blended learning environment in the long term. This requires rethinking online education pedagogies and data governance for online learning.

While it may seem logical to focus on the data for better data governance, educational institutions should instead focus data governance strategies on the impact on students and staff, improving teaching and learning experiences, and sustaining data infrastructures to scale in the long run. To ensure that the academic community's privacy is respected, Borgman[25] suggests that:

- Universities should focus on their core missions of teaching, research and services to address priorities for data collection and stewardship through accountability.
- Whenever data is collected, the consequences on the individual and the community should be considered. Only data that matters, not data that is easy to gather, should be collected.
- Joint data governance between faculty and administrators should be considered.
- Universities should promote awareness and transparency to maintain high levels of public trust.
- Patience and broad consultation of the use of data at institutions should be prioritised.

[22] ODI (2021), 'Data about teachers' lives during the pandemic'.
[23] Defend Digital Me (2020), 'The State of Data 2020'.
[24] JISC (2021), 'Learning and teaching reimagined: a new dawn for higher education?'.
[25] Borgman CL (2018), 'Open Data, Grey Data, and Stewardship: Universities at the Privacy Frontier'.

The Office for Students' Gravity Assist report,[26] which collected empirical evidence during the pandemic, identified six components to successful digital teaching and learning:

- Digital teaching must start with appropriately designed pedagogy, curriculum and assessment.
- Students must have access to the right digital infrastructure.
- Good access enables staff and students to build the digital skills necessary to engage.
- Technology can then be harnessed strategically, rather than in a piecemeal or reactive way, to drive educational experience and outcomes.
- Inclusion for different student groups must be embedded from the outset.
- All the elements need to be underpinned by a consistent strategy.

Ofsted has produced a report on remote education research in schools.[27] The UK Department for Education has also produced policy work on how technology can be implemented to overcome barriers within education.[28] This could be extended in light of the pandemic and with greater consideration of data governance and data infrastructure innovation. The EU has also considered the relationship between emerging technologies and the teaching profession.[29] That report emphasises that emerging technologies and the datafication of education offer an opportunity to rethink and re-frame success in education and training, encouraging the creation of new practices to support effective learning in the future. The French data protection authority CNIL has gone one step further, calling for French academic institutions to reassess and consider alternatives to US-based collaborative tools on privacy grounds, especially related to sensitive data such as research data or data relating to minors. In response to Covid-19, the European Commission has established the Digital Education Action Plan 2021-2027 to address the impact of the pandemic on education and its strategy for the long-term digital transformation of education and training systems. The report identifies key enabling factors for effective digital education and training:

- connectivity and suitable digital equipment for learners and educators
- teachers and trainers that are confident and skilled in using digital technology to support their teaching and adapted pedagogy
- leadership
- collaboration and
- the sharing of good practice and innovative teaching methods.

Applying existing research on online learning to potential policy solutions, data management models can be applied to digital education to reveal new ways of work that are privacy-friendly, secure and innovative. For example, the ODI Data Ethics Canvas can help identify ethical challenges for collecting certain types of data and implementing new technologies. The Data Ecosystem Mapping tool used in this report can also be applied to specific institutions to help map out data governance strategies based on institution-specific goals. Identifying the value of data from online

[26] Office for Students (2021), 'Gravity assist: propelling higher education towards a brighter future'.
[27] Ofstead (2021), 'Remote education research'
[28] UK Department for Education (2019), 'Realising the potential of technology in education'
[29] Vuorikari R, Punie Y, Cabrera Giraldez M (2020), 'Emerging technologies and the teaching profession: Ethical and pedagogical considerations based on near-future scenarios'.

learning, balanced with ethical decision-making, can support better cost-benefit analysis when introducing new technologies, gathering new forms of data, and implementing new ways of learning.

Beyond institutional decision-making, institutions should be aware of wider strategies on data and data governance as applied to online learning. The UK government has focused its effort on rebuilding the country's education strategy, some of which is powered by data. With the Skills and Post-16 Education Bill, the government aims to ensure that higher education is up to date with the latest technological knowledge and skills. The Online Safety Bill may place more responsibility on institutions to govern data and content used for education. The UK may also take inspiration from initiatives within the EU such as the GARR Consortium in Italy, a national computer network for universities and research with the aim to design and manage high-performance network infrastructure that delivers advanced services to the academic and scientific community. The Scottish government has also published 'A changing nation: how Scotland will thrive in a digital world', a new digital strategy that aims to establish Scotland as an ethical digital nation and increase digital skills through education. JISC, in its response to the National Data Strategy Consultation, suggests that a new national data infrastructure for research is needed for the benefit of education and research, related to the organisation's Education 4.0 targets for 2025. The UK Department for Education is also recruiting for a No. 10 Innovation Fellow to work specifically on harnessing the potential of digital solutions in education. As the government implements digital governance within public policy, institutions can look to these strategies to support their short-term and long-term data goals, mapping out the types of data their institutions are planning to collect, what stakeholders need and what can help meet those needs, and whether new technologies are necessary.

Gathering empirical data

Acknowledging that online learning and data governance strategies will look different for different institutions, individual educational institutions may look to gather stakeholder perspectives to identify the benefits and risks of online learning from their own community.

In this section, based on existing research and proposed solutions, we suggest a set of questions that provide a survey template that can be used by university management to better understand how online learning has impacted the academic experience of students and staff outside traditional feedback documents. Institutions are encouraged to adapt the template to their needs and provide as much information as possible on their data-related policies to better inform stakeholders of existing online learning considerations.

Firstly, the beginning of the feedback form should include changes that the institution or organisation has made when it comes to data governance and data protection practices for online learning. This may vary across departments or schools within an institution. Institutions may also wish to share their future plans for feedback. The form should also include an introduction to explain key terms and help the respondent understand what constitutes data in online learning. For example:

Introduction

Digital technologies are increasingly becoming a part of our everyday lives. This includes the area of higher and further education, where students, staff and academic institutions can use technologies and educational platforms to deliver lessons, record classes and introduce new interactive methods of learning.

This form aims to gather information about how different stakeholders have experienced data being collected, used, managed and protected in online learning. Data around online learning refers to quantitative or qualitative information collected in a digital educational environment or technology platform. This data can be about any stakeholders that engage with online learning, such as students, teachers, researchers and parents. Examples of data related to online learning include student information collected so they can access digital educational tools, learning analytics and tutorial recordings. Data that has been collected may also be used for purposes other than delivering online learning, such as increasing the quality of online learning platforms, repurposing recorded tutorials for academic and non-academic use, and monitoring student and staff engagement during sessions.

Next, the feedback should contextualise the role of the respondent to support better understanding of their perspective. For example:

Stakeholder category

a. Student
b. Teaching staff
c. Research staff
d. School/university management
e. Technology company or platform
f. Education provider
g. Regulation and policy
h. Public or non-profit organisation
i. Other [free text]

2. How prepared were you or your organisation for online learning with regards to data management and governance from March 2020 (when the first Covid-19 lockdown started) onwards?
   a. Very prepared
   b. Slightly prepared
   c. Slightly unprepared
   d. Very unprepared
3. If you are interested in our research findings, please leave your email address and we will notify you of our final publication. This information will not be used for other purposes:
   a. [free text]

The main section of the feedback should focus on the specific impact of data

governance and data-related changes faced by the respondent. This could includes questions such as:

Data governance questions

1. How prepared do you think your organisation was for online learning with regards to data management and governance going into online learning from March 2020 onwards?
   a. Very prepared
   b. Slightly prepared
   c. Slightly unprepared
   d. Very unprepared
2. Are you confident in your organisation's ability to implement hybrid or blended learning, including support for online learning data, for the 2021/22 academic year?
   a. Very confident
   b. Slightly confident
   c. Slightly doubtful
   d. Very doubtful
3. How satisfied are you about the changes (or lack thereof) that have been implemented by your organisation in response to data governance for online learning since March 2020?
   a. Very satisfied
   b. Slightly satisfied
   c. Slightly unsatisfied
   d. Very unsatisfied

Finally, the last part of the feedback form should be future-focused and encourage free form reflections both on how changes have impacted the education experience for the respondent, as well as what they hope to see in the long term:

The future of learning questions

1. In hindsight, what do you wish to have known regarding online learning and how online learning data is used within your organisation?
2. Is there anything that you wish was implemented from the start of online learning?
3. What online learning practices would you like to continue to implement in a blended learning environment?
4. What online learning practices would you like to continue to implement when fully in-person learning returns?

By gathering direct feedback, institutions can better understand how the community has responded to previous changes in policy and shape the future direction of data governance and online learning within the organisation.

Discussion

Summary

After one year of online learning, and as the UK slowly transitions into blended learning models for the upcoming 2021/22 academic year, it is clear that some of these digital technologies and ways of learning will remain post-Covid. Technology and blended learning have become the new normal. As a result, educational institutions should rethink how they can better manage, protect and govern online learning data and personal data to improve the online academic experience.

Limitations and caveats

The aim of this report was to provide insight and evidence on the data protection and data-related concerns that students and staff may have over the course of online learning during the pandemic. We acknowledge that during this time, students and staff, as well as their institutions, have given their best efforts to continuing their academic activities and responsibilities despite these challenging times. This report aimed to illuminate the areas in which higher education institutions should consider when developing data governance strategies for the future, not only in preparation for future emergencies, but also to balance digital innovation in education with maintaining student and staff autonomy over their personal data.

It should also be noted that as the Covid-19 pandemic continues in the UK and many parts of the world during this report's publication, more research needs to be done and more evidence needs to be collected to better understand how the pandemic has affected students, staff and academic institutions in the long term. This report has highlighted some of the work that has been conducted on online learning prior to the pandemic, as well as during some of these transitions.

Although the focus of this report is on online learning data and data governance, we recognise other related challenges of online learning such as access to education and online learning, internet connectivity, and socio-economic challenges that could make online learning undesirable.

Further, this report focuses on higher education and may not be suitable for schools and the governance of children's online learning data. However, this report could inform data considerations required in schools and colleges, where more caution should be taken for children and young people.

As the authors of this document are based in the UK, the data protections regulated within the UK may be different in other countries. While not all rights and privileges may be granted outside of the UK and the EU, we encourage the careful consideration of data governance for online learning globally to ensure that student and staff personal data are better protected.

Concluding remarks

In order to create an online learning strategy that is fit for the future of digital education, university management should consider how online learning data ought to be governed as part of their long-term strategy. This not only includes research data but also student and staff personal data related to teaching, admin and data processed by external platforms.

From an institutional perspective, when the implementation of innovative technologies is supported by better data governance, a more dynamic and engaging academic environment can be created for staff and students. This involves careful consideration of what is and what isn't necessary when it comes to online learning data. When creating new data governance strategies, educational innovation and data protection best practices go hand-in-hand, where good data management and forward-thinking data infrastructure planning can make academic institutions more attractive and create a better foundation for success.

Higher education institutions and university management need to better understand how online learning has affected the university community, not only with regards to teaching and learning, but also with the university experience. Universities should explore different options when it comes to introducing new technologies, establishing regular evaluations to ensure that the technologies that are implemented are fit for purpose. With this in mind, policymakers should support institutions in making these decisions, providing guidance on how online learning and education data should be governed, as well as outlining the data infrastructure and data governance strategies for the education sector as we continue to rely on hybrid education delivery for the 2021/22 academic year.

Looking ahead

Thinking beyond the pandemic and looking to the future of online learning in higher education institutions, universities should reflect on the pandemic online learning experience and update their long-term strategies on how they want to deliver academic content and what they want the university experience of the future to be. This also requires support from education technology companies and policymakers in considering funding data infrastructure across the education agenda. In order to attract a more diverse and international body of staff and students, institutions should focus on improving the academic experience beyond implementing technologies. Having experienced a year of online learning from March 2020, students and staff look to create meaningful relationships with the university community outside of the classroom. This requires re-imagining the purpose of higher education and considering how an institution's values and aims can be carried forward both online and offline.

Resources list

Useful resources linked in the report for more in-depth research and guidance.

Open Data Institute resources

Open Data Institute (2021), 'Data Landscape Playbook'

Open Data Institute (2021), 'Data Ethics Canvas 2021'

Open Data Institute (2019), 'Data Ecosystem Mapping Tool'

COVID-19 online learning resources

Office for Students (2021), 'Gravity assist: propelling higher education towards a brighter future.'

JISC (2021), 'Learning and teaching reimagined: a new dawn for higher education?'.

Edge Foundation (2020), 'The impact of Covid-19 on education: A summary of evidence on the early impacts of lockdown'.

Berkman Klein Center (2020), 'Digital Ethics in Times of Crisis: COVID-19 and Access to Education and Learning Spaces'.

Vuorikari R, Punie Y, Cabrera Giraldez M (2020), 'Emerging technologies and the teaching profession: Ethical and pedagogical considerations based on near-future scenarios'.

Other resources

Fiebig T, Gürses S, Gañán CH, Kotkamp E, Kuipers F, Lindorfer M, Prisse M, Sari T (2021), 'Heads in the Clouds: Measuring the Implications of Universities Migrating to Public Clouds'.

Silverman S, Caines A, Casey C, de Hurtado BG, Riviere J, Sintjago A, Vecchiola C (2021), 'What Happens When You Close the Door on Remote Proctoring? Moving Toward Authentic Assessments with a People-Centered Approach'.

Angiolini C, Ducato R, Giannopoulou A, Schneider G (2020), 'Remote Teaching During the Emergency and Beyond: Four Open Privacy and Data Protection Issues of 'Platformised' Education'.

Borgman CL (2018), 'Open Data, Grey Data, and Stewardship: Universities at

the Privacy Frontier'.