Department of
Social Policy

Working Paper
February 2025

In the Fine Print:

Investigating EdTech Providers’ Data Privacy Commitment

K. Barud[1], T. Henne[1], V. Hillman[2], E. Radkoff[3], C. Saillant[1], E. Zdravevski[4]

Department of Social Policy
London School of Economics and
Political Science Houghton Street
London WC2A 2AE

Email: soc.pol.webteam@lse.ac.uk

Telephone: +44 (0)20 7955 6001

lse.ac.uk/social-policy

To cite this paper: Barud, K., Henne, T., Hillman, V., Radkoff, E., Saillant, C., Zdravevski, E. (2025) In the Fine Print: Investigating EdTech Providers' Data Privacy Commitment, Social Policy Working Paper 02-25, London: LSE Department of Social Policy.

In the Fine Print: Investigating EdTech Providers' Data Privacy Commitment

K. Barud[1], T. Henne[1], V. Hillman[2], E. Radkoff[3], C. Saillant[1], E. Zdravevski[4]

[1] University of Vienna (AUSTRIA)
[2] London School of Economics and Political Science (UNITED KINGDOM)
[3]TOSDR (UNITED STATES)
[4]Faculty of Computer Science and Engineering, University Ss Cyril and Methodius University in Skopje (MACEDONIA)

ABSTRACT

In the rapidly evolving field of educational technology, maintaining quality assessment processes is essential for effective education governance. Ensuring transparency in data processing and compliance with privacy laws is crucial for building trust among all stakeholders. This study investigates the data protection practices of selected EdTech providers through a mixed-method approach that integrates manual assessments with machine learning techniques. We focus on two main areas: (1) analysing the transparency and legality of information vendors provide to schools, based on the articulation of their data privacy policies (DPPs), and (2) exploring the methodological integration of human and ML-based analyses. The research evaluates how EdTech providers communicate their data processing practices and adhere to privacy regulations outlined in their DPPs. Such practices are vital for fostering trust between schools and EdTech providers. Given the complexity and cost of conducting Data Privacy Impact Assessments, our study aims to develop a user-friendly template for assessing DPPs and test innovative technologies for scaling this process efficiently. Initial findings from ML-supported assessments of ten popular EdTech providers in England reveal varying levels of transparency and compliance and technological limitations. Our innovative methodology identifies current errors in ML use but equally enhances the scalability of our evaluation framework. This research contributes to discussions on the intersection of education, technology, ethics, and policy, advocating for responsible EdTech innovation that prioritises transparency and ethical integrity around their data practices, while examining the role of ML in supporting schools' procurement and assessment processes.

Keywords: data privacy, data protection, GDPR, EdTech, Machine Learning, ChatGPT

1. INTRODUCTION

How EdTech providers implement data privacy obligations with regards to education data is pertinent not only to demonstrating their compliance with the law but also in ensuring that a private sector with growing influence in public education upholds children's fundamental human rights. How EdTech providers' data practices are articulated in their data privacy policies (DPPs) plays an incremental role in creating a trusted, transparent, and accountable digital education environment. While significant scholarship has scrutinized the risks posed by major tech giants like Microsoft and Google (Kerssens, Nichols & Pangrazio, 2023), smaller EdTech providers, which make up a large portion of the market (EU EdTech Alliance, 2022), have received less attention despite their growing influence in education globally.

Over 50,000 apps are dubbed 'educational' (Kucirkova, Campbell & Cermakova, 2023, p. 10), and the projected revenue for the EdTech market is US$239 billion (Statista, 2023). According to HolonIQ, a global market intelligence provider, the entire education sector is valued at US$7.2 trillion, combining total global expenditure from governments, companies, and consumers (HolonIQ). Yet, there is little substantive knowledge about the true state of EdTech providers' data privacy practices, and their governance and procurement also remain highly fragmented (UNESCO, 2023; Hillman, 2022). Instead, data privacy malpractices have been evidenced in recent years (International Digital Accountability Council, 2020), suggesting that many providers lack transparency and misuse student data—not always out of commercial intent, but possibly due to ignorance or negligence (Human Rights Watch, 2022). Protecting children goes beyond mere compliance with the GDPR; it requires a commitment to transparency and trust (Information Commissioner's Office [ICO] 2018). This study aims to analyse the DPPs of EdTech providers subject to UK and European data privacy regulations (ICO, 2018; Regulation (EU) 2016/679).

In the UK, the GDPR contains provisions that aim to enhance the protection of children's personal data (Information Commissioner's Office [ICO], 2018). Education data specifically pertains to the collection and processing of personal data, all of which are regulated under the UK GDPR and the Data Protection Act (2018). Transparency and accountability are crucial where children's data is concerned. Understanding the lawful basis on which EdTech providers process children's personal data should be outlined through their DPPs at a minimum.

This research sits at the intersection of data protection, privacy, and EdTech within the context of legal and educational frameworks. Although DPPs alone cannot determine compliance, our study highlights the need for comprehensive audits, including data processing records and security measures, as well as additional steps such as signed contracts with data processors and measures used to secure data transfers. Our methodology comprises two parts: manual evaluation of the DPPs of selected providers, and subsequent ML processing of these documents to enable policy assessments efficiently and at scale. The paper outlines key concepts, methodology, and recommendations for improving the systematic evaluation of EdTech vendors.

2. THE DIGITIZED EDUCATION LANDSCAPE: KEY CONCEPTS AND RESEARCH RATIONALE

2.1. What we mean by education (and personal) data

Education data encompasses personal and sensitive data about pupils, such as demographics and personally identifiable information like birth date, home address and unique identifiers given to the pupil by their districts or school and other information like parents' marital and income status, whether pupils receive free lunches in school, medical history, disability information, socio-emotional wellbeing, biometric data such as face and voice, special educational needs and others (Barassi 2020). Within the category of education data, we distinguish personal data which comprises information about 'natural persons' who can be identified directly from the gathered data or who can equally be indirectly identified from the collected data in combination with other data (ICO 2021: 9).

Education data is processed for the purpose of teaching, learning and assessment, to uphold statutory safeguarding requirements, for managing school processes, and for reporting and accountability purposes. For example, Pearson Inc, the publishing and technology company, provides assessments, and various other applications and content, including Q-global, a ‘system [that] organises examinee information, automates the scoring process, and generates score reports’ (Pearson Inc. n.d.: 1). The web-based system collects data such as test results and raw scores from assessments, student demographics and parental information, including living conditions and more.

Both education and personal data can be collected by EdTech products, therefore it is automatically rendered sensitive. Additionally, all this data may be exchanged between schools and EdTech providers and between the school and third-party organisations and authorities, adding to the challenge of governing the privacy, safety and security of the data - and therefore of all pupils.

2.2. The importance of privacy and children’s rights

Substantial literature highlights the importance of data privacy and the risks of harm from data privacy loss (Zeide 2017; Citron & Solove 2021). To the growing child, privacy is a deterministic condition to exercise their basic freedoms and rights. Privacy plays a critical role in the development of one’s identity, ideas, and personhood. Privacy enhances individual autonomy and is an incubator to the development of thought, speech and association. In Neil Richards’ words, intellectual privacy is a ‘zone of protection that guards our ability to make up our own minds freely’ (2008: 95). In increasingly digitised education, children are at risk of losing this zone of protection if data privacy loss is also at risk. The loss of privacy leads to a wide range of risks that can include ‘unknowable’ (Citron & Solove 2021: 817) and future harms. With the fast-advancing generative artificial intelligence which requires the processing and repurposing of granular data, the risks are beyond matters of privacy (Mantelero 2022).

In the UK, children have rights under the United Nations Convention of the Rights of the Child, by the Data Privacy Act (2018) and the UK GDPR among other stipulations. The UNCRC came in 1989 and for the first time a global effort was made for governments to agree to recognise the rights of children. The Convention incorporates 54 rights which are interlinked. Articles between 41 and 54 address adults’ and governments’ responsibilities to ensure that children and young people enjoy these rights. At their core, they stipulate that every child has the right to live and develop; to equality and non-discrimination; to be heard and to participate in society. Additionally, every child is entitled to an education to support their development and achieve their full potential; to play; to freedom of thought; to voice and peaceful protest; to protection from harm; to equality and non-discrimination; to participation; to identity; to enjoy their own cultures and practise their religion and use their own languages if they belong to an indigenous or minority group. However, the UNCRC states:

> In all actions concerning children, whether undertaken by public or private social welfare institutions, courts of law, administrative authorities or legislative bodies, the best interests of the child shall be a primary consideration (UNCRC 1989: 3).

As Van Der Hof et al. emphasise (2020), this statement argues for *a*, not *the* primary consideration, which means they include those of commercial entities, which in practice may ‘be weighed against the interests of powerful companies that may very well be the direct opposite to those of children.’ (841). Crucially, these fundamental rights are also recognised in the digital environment (UNCRC 2021). This means that digital service providers catering for children must adhere to children’s digital rights.

2.3. Data brokering and the commercial value of data

Schooling systems inadvertently facilitate education data brokering through their data practices. As a result, Kemp (2020) highlights the risks and costs of concealed data practices on end users. Manwaring (2022) further argues that the current consent regulations inadequately protect consumers, despite ample evidence emanating from data privacy loss. Stoilova et al. (2020) posit that children and schools struggle to comprehend and manage online privacy in the face of commercial data collection.

Zuboff (2015) has long exposed digital surveillance capitalism’s exploitation of personal data without explicit consent, which has led to valuing data as capital, while van Dijck (2014) highlights the pervasive nature of dataveillance and data collection, which has now been normalised in education. Lupton and Williamson (2017) have argued that children have limited understanding of how corporations exploit their data, while more recently scholars also demand that independent audits and assessments must be carried out on digital technologies especially as they are fast evolving with powerful algorithmic and automation capabilities that can exacerbate existing inequalities and injustices (Broussard 2023). These insights show that while education data has multiple uses, it also offers commercial opportunities, necessitating stricter protection to safeguard children's rights.

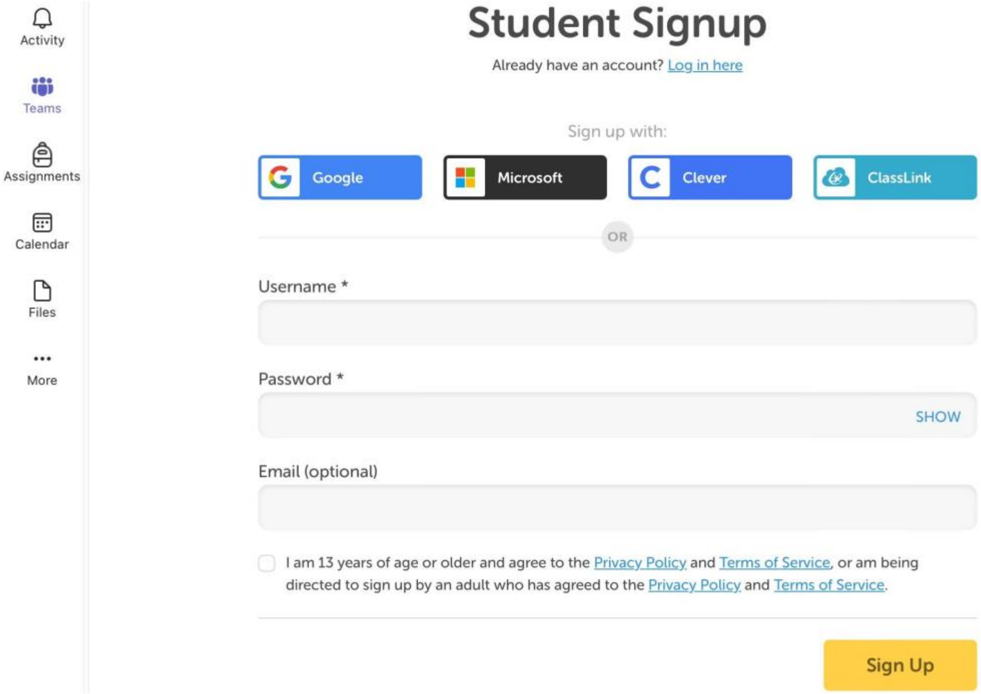
2.4. EdTech and their privacy obligations

Public education systems face two major challenges following the rapid adoption of digital technologies, especially in the aftermath of the global health pandemic. On one hand, there is the platformisation and growing dependence on global corporate ecosystems and their complex architectures, data capitalisation and global infrastructure capture (Zuboff 2019). On the other hand, schools are dealing with everyday aspects of data privacy and cyber insecurity due to the thousands of EdTech products and services, many of which are themselves integrated into larger platforms through socio-technical strategies such as cloud hosting, services integration, single sign-ons, and APIs. This results in a complex 'digital journey' for students (Livingstone et al., 2024). While they may use one EdTech app, their digital learning experience often leads them through several others, with their data trickling to various parties with unknown consequences for how it is protected or used.

For example, a student from Year 8 in a UK public secondary school might be asked to use an app called Typing, which she would access via Microsoft Teams, the Big Tech that is used by millions of schools worldwide. However, to subscribe to Typing, the student can use Google, Microsoft, Clever or ClassLink login credentials - all commercial platforms. The pupil also has to opt in to Typing's Privacy Policy and Terms of Service and to agree that she is 13 years of age (Figure 1)[1]. Thus, from Microsoft Teams (provider

[1] This is based on a real example of pupils from an English secondary school, where the app is used.

#1), the pupil moves to another app (*Typing*, provider #2) and by signing up, say with a Google email, the pupil also enters a third provider (#3), all of which subject her to different DPPs. None of these documents are easy to read and understand as none of them articulate simple legal requirements and how they meet these.



Student Signup

Already have an account? Log in here

Sign up with:

Username *

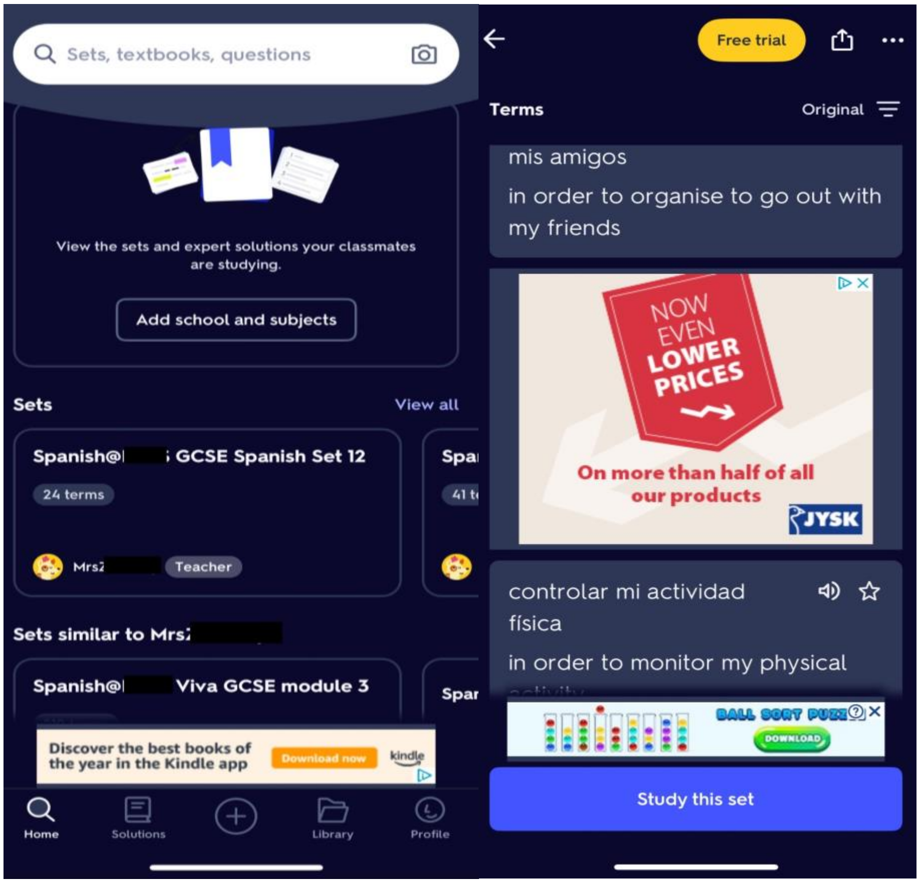
Password *

Email (optional)

□ I am 13 years of age or older and agree to the Privacy Policy and Terms of Service, or am being directed to sign up by an adult who has agreed to the Privacy Policy and Terms of Service.

Figure 1. A student's digital journey often spans multiple providers -- signing up to Typing via Microsoft Teams can lead to varying DPPs and levels of data protection commitment

Another example is the popular app *Quizlet* (Figure 2) a US-based company that provides quiz-based learning and content for core and non-core subjects. The company claims that every other pupil in the US and one in seven K-12 pupils in the UK use their product (Quizlet 2019). The free version, which many pupils and teachers use in the UK, is sprinkled with adverts. The company offers advertisers attractive deals to 'put [their] brand in front of 60 million achievers globally where they are most engaged'(Quizlet n.d.).



View the sets and expert solutions your classmates are studying.

Spanish@[redacted] GCSE Spanish Set 12

Mrs[redacted]

Sets similar to Mrs[redacted]

Spanish@[redacted] Viva GCSE module 3

mis amigos

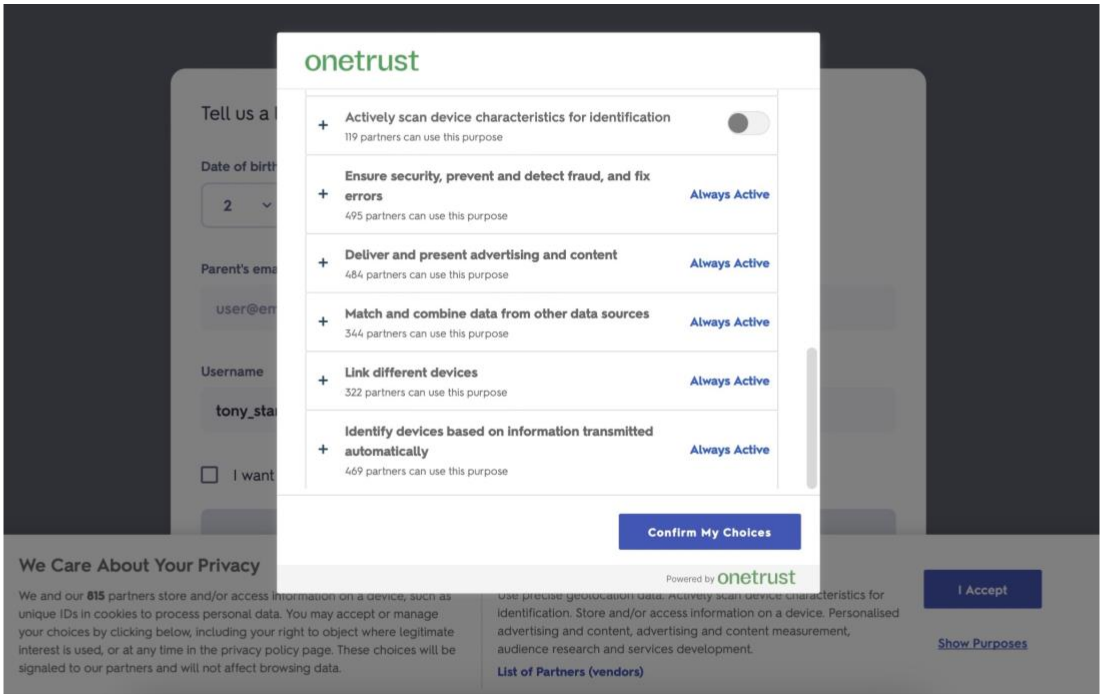
in order to organise to go out with my friends

controlar mi actividad física

in order to monitor my physical activity

Figure 2. Amazon, JYSK and even gambling sites are advertising on the learning app Quizlet.

There are hundreds of apps that use the same business models. Additionally, their conditions to data collection from third parties (Figure 3) is just as complex, difficult to follow, and often unethical and even illegal (see Federal Trade Commission 2023).



onetrust

- Actively scan device characteristics for identification — 119 partners can use this purpose
- Ensure security, prevent and detect fraud, and fix errors — 495 partners can use this purpose — Always Active
- Deliver and present advertising and content — 484 partners can use this purpose — Always Active
- Match and combine data from other data sources — 344 partners can use this purpose — Always Active
- Link different devices — 322 partners can use this purpose — Always Active
- Identify devices based on information transmitted automatically — 469 partners can use this purpose — Always Active

We Care About Your Privacy

We and our 815 partners store and/or access information on a device, such as unique IDs in cookies to process personal data. You may accept or manage your choices by clicking below, including your right to object where legitimate interest is used, or at any time in the privacy policy page. These choices will be signaled to our partners and will not affect browsing data.

Use precise geolocation data. Actively scan device characteristics for identification. Store and/or access information on a device. Personalised advertising and content, advertising and content measurement, audience research and services development.

List of Partners (vendors)

Figure 3: To use Quizlet, DPP requests from users to consent many other companies' own privacy terms

Many schools in the UK rely on data privacy impact assessments (DPIAs) to ensure that the EdTech products they use comply with privacy laws. The problem is further complicated as often it is conflicting and hard to discern who is the data processor and who - the data controller. Often, the EdTech providers view themselves as simply data processors; they do with data what schools tell them to and not as data controllers, which is typically the school.[2] This of course puts all the liability and responsibility over the data on schools.

2.5. Data privacy requirements

The EU and UK GDPR establish a comprehensive legal framework, aligning with human rights principles.[3] These govern how entities process personal data. In educational settings these are pertinent to children's and teachers' safety, wellbeing, and respect for

[2] See e.g. Turnitin Services Privacy Policy.

[3] In the context of the EU GDPR especially Article 8(1) of the Charter of Fundamental Rights of the European Union.

their fundamental rights. These regulations also impact global data transfers, extending protections beyond the EU and UK to individuals worldwide (EDPB 2020).

EdTech providers who process children's data for their own purposes become data controllers and, in this role, they must adhere to the GDPR's principles of transparency, legality and fairness, and importantly – clearly articulate how they do so. Certainly, data controllers may opt to include extra details about data processing, such as security measures, any DPIA outcomes, and so on. Although not explicitly required by Articles 13 and 14 of the GDPR, these disclosures reflect transparency and accountability that breathes more trust to education stakeholders.

Edtech vendors must also adhere to child protection measures, mandated by the UNCRC, which call for technologies to support cultural and linguistic diversity and to be used ethically. This includes protecting children's data from commercial exploitation and manipulation. The Age-Appropriate Design Code (AADC) by the UK's Information Commissioner's Office (2023) for example, outlines guidelines for child-focused services in line with UNCRC, GDPR, and the UK Data Protection Act 2018. While not legally binding, non-compliance with the AADC can suggest violations of these laws.

Moreover, the AADC can be seen as an additional layer of requirements is provided for DPPs. When developing EdTech DPPs, the fourth AADC principle, 'transparency,' should be prioritized, aligning with GDPR Articles 12-15 and the fairness, transparency, and lawfulness requirements of Article 5 UK and EU GDPR. That is, policies must distinguish between essential and optional data processing activities, provide clear and regularly updated information, and be accessible and age-appropriate for children. EdTech providers should use 'just in time' notifications for data use changes and adapt privacy information based on age, employing visuals like cartoons for younger children and detailed explanations for parents and educators (ICO, 2020).

This brief review highlights the importance of protecting children's rights while navigating complex compliance requirements and emphasises that while vendors use DPPs to demonstrate adherence and commitment to the data privacy of children, these documents are often not user-friendly. That said, while circumstantial, from a child's rights perspective, EdTech vendors' DPPs should, at a minimum, be clearly articulated to demonstrate transparency, accountability and commitment to children's wellbeing, needs and best interests.

3. METHODOLOGY

3.1. Research design and questions

Our research aimed to investigate whether EdTech providers are transparent about their data practices and compliant with privacy laws as outlined in their DPPs. The rationale was two-fold: we aimed to streamline a requirement that schools have for assessing and ensuring the EdTech providers they contract with are legally and ethically committed – and clearly articulate that – and to foster learning support to providers themselves as they navigate the complex and evolving landscape of laws and standards.

The research consisted of two parts: one was explorative and bottom-up, aimed to directly engage with policies and terms of services and see how clear these statements are and what they say about companies' data privacy practices; and a second one - empirically structured and top-down as it was informed and developed directly from the existing laws of what companies should articulate and be transparent about.

Starting with the bottom-up method, we engaged with existing grassroots measures - the community project (2012), called *Terms of Service; Didn't Read* (ToS;DR)[tosdr.org] and the Ranking Digital Rights public services project (RDR 2020), which evaluates and ranks some of the biggest digital technology companies in the world across governance, privacy and other measures of accountability. While neither of these specifically target the EdTech sector, exploring these existing grassroots efforts informed our methodology development. Additionally, these existing efforts highlighted the gap and need for a scalable method to evaluate and rank EdTech providers based on key data privacy and child-focused regulations.

Our sample initially started with 800[4] EdTech companies' DPPs and Terms of Services (ToS)[5], which operate in the UK and globally. The focus was on their capacity to

[4] This is an ongoing process and in this working paper we present our initial findings.

[5] The full list of providers, GDPR cases created for this research and relevant code are available on GitHub https://github.com/admin-magix/EdTech-policies

serve children in UK public schools and, as such, how they address and adhere to UK and EU data privacy laws and standards. Our guiding research questions in this work included:

Q1 Do DPPs effectively and in a clear manner indicate the legality, safety, and ethics of managing children's educational data? What are the major shortcomings in these DPPs?

Q2 Can ML techniques adequately and thoroughly assess DPPs and their compliance with the (UK and EU) GDPR requirements? What are the major shortcomings of ML techniques when performing this task?

Q3 What are the learning points from this process for future compliance assessment by means of ML techniques?

3.2. Exploring existing efforts: bottom-up community-based ML tools

Our explorative phase included the experimentation with the community project ToS;DR, which provides easy-to-understand takeaways of DPPs and ToS for the most popular web services and apps. We also drew knowledge from the RDR project (2020) and its corporate accountability index, whose aim was to publicly display levels of transparency and accountability by detailing the biggest companies DPPs and ToS and ranking them based on international human rights standards. While we only share initial findings from the ToS;DR with whose ML we engaged, RDR is mentioned here to highlight their invaluable drive to promote rights-respecting digital services online is valuable in identifying ways to promote children's rights' respecting EdTech.

ToS;DR[6] came to fruition in 2012 with the explicit goal of strengthening the data and digital privacy rights of everyday people by providing summaries of DPPs and other fine print that they agree to without having the time to read. Nearly 1000 volunteers have contributed summaries over the years, and the project is now transitioning to a ML assisted approach. Although not tailored specifically for children's data or EdTech, this collective work initially started by applying the ToS;DR's methodology to a 10 EdTech providers from our subset (GitHub 2024) with the following research objectives:

[6] https://tosdr.org

1. Evaluate the extent to which ToS;DR's framework overlaps with the EdTech-specific data rights issues motivating our research.
2. Examine the ToS;DR summaries of our EdTech companies of focus, including A-E privacy grades, and compare them to other industries.
3. Validate the semi-automated approach of ML-driven analyses with human verification, to gauge feasibility of applying it to our larger corpus of ~800 companies and beyond.

3.2.1. The ToS;DR model

ToS;DR has come up with a taxonomy of privacy-related statements that one might want to know about a service, called *Cases*. Examples follow:

- Your personal data is not sold
- You must provide your legal name, pseudonyms are not allowed
- Tracking cookies refused will not limit your ability to use the service
- Your data is processed and stored in a country that is less friendly to user privacy protection
- The terms for this service are easy to read

There are over 100 Cases in total. These are sorted into four sentiment classes, depending on their perceived desirability for users of the service — positive, neutral, negative, and blocker (very negative). The Cases provide high-level takeaways from DPPs. For an even higher-level summary, ToS;DR provides grades for each service, ranging from A to E, depending on how the practices described in the DPPs suggest the company operates as a custodian of user data. Grades are determined according to the following formula. A weighted sum is used to turn counts of verified Case statements into grades, A through E. An ideal A rating means that the service's stated data handling practices are deemed very respectful to users' privacy.

3.2.2. Automation methodology

Manual extraction of each Case statement across 10 services would be possible, but tedious, and scaling beyond that – not practical. To aid in our study and explore what scaling to hundreds or thousands of services would look like, we utilised ML models developed internally by ToS;DR. provide an overview of the training methodology here.

For each Case a binary classification ML model was trained that takes as input one or more sentences from a DPP, and outputs a 0.0-1.0 score where 1.0 means there is extremely strong evidence that the Case statement is true.

128 Case models were trained in total. Each started as a copy of the uncased BERT language model (Devlin, 2019), and were fine-tuned with LoRA adapters (Hu, 2021) and a binary classification head using a dataset of human judgements. Positive training instances included sentences affirming evidence of a particular Case, first identified and submitted by ToS;DR volunteers, and later double-checked by trusted curators on the ToS;DR platform. Negative training instances included volunteer submissions rejected by ToS;DR curators, sentences that border positive instances, sentences providing evidence for other related Cases, and a larger pool of random sentences pulled from DPPs.

Training was accelerated by a GeForce RTX 3090 GPU, using the `transformers`[7] and `peft`[8] libraries from HuggingFace. Early stopping was used to halt training, selecting models that maximised F1 on a hold-out test set. Training code is open source[9].

[7] https://github.com/huggingface/transformers/

[8] https://github.com/huggingface/peft

[9] https://github.com/tosdr/DocBot

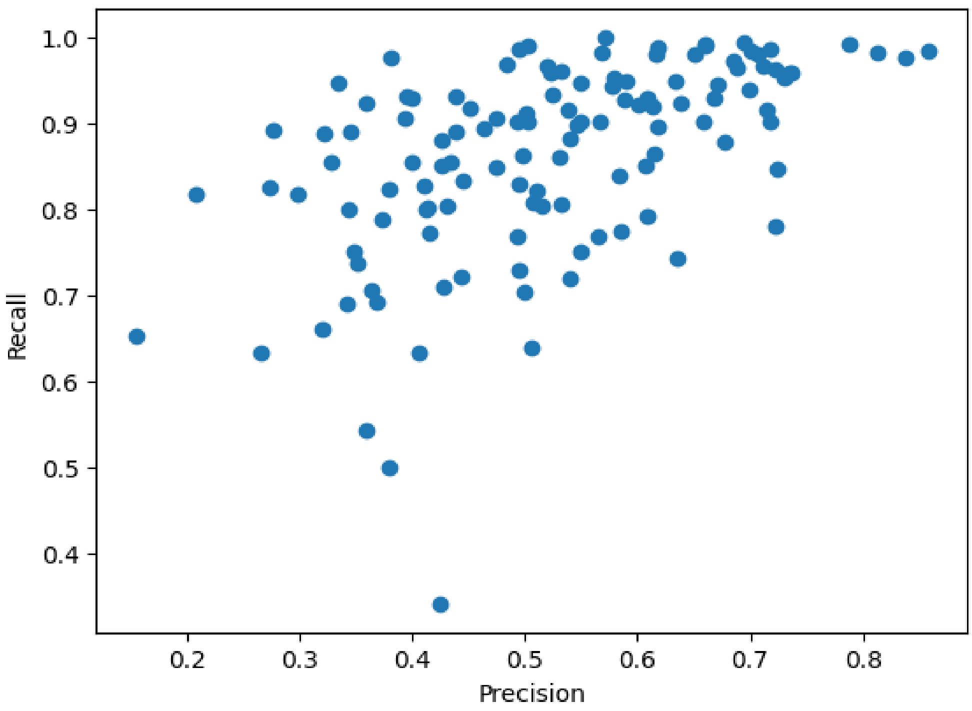


Figure 4. Precision and recall achieved on the test set of all 128 trained Case models

While the models are limited to processing 512 tokens at once, we analyse full DPPs by applying each Case model to all individual sentences, interpreting the maximum score achieved by any one sentence as the best available evidence that the Case statement might be true. After testing variations with additional sentences prepended or appended, we apply a score threshold to determine whether the most convincing evidence is sufficient to assert the Case. Thresholds were chosen to maximise f-score on a hold-out test set of DPPs. Figure 4 shows the precision and recall achieved on the test set of all 128 trained Case models.

Following these automated assessments, we incorporated human oversight to ensure where false positives/negatives may have occurred in the results. This meant that volunteers went through individual cases and compared with the existing DPP and ToS texts to confirm if the result was accurate. This also helped us to assess the robustness of the ML assessment.

3.3. Developing a top-down framework

As a result of the preliminary experiments with the ToS;DR tool, the research process was reviewed and the following several distinct steps were defined:

1. **Primary data collection:** Conducted human assessment and analysis of DPPs from vendors, focusing on legal and child-related standards.
2. **Scorecard framework:** Developed an assessment framework with prompts based on EU and UK GDPR requirements for DPPs, cross-checked for accuracy.
3. **Secondary review:** Verified initial assessments through cross-referencing with five legal researchers.
4. **Enhanced assessment:** Used ML to automate the analysis of DPPs for long-term sector-wide evaluation (see **Figure 5**).
5. **Human oversight:** Legal team reviewed ML results to assess accuracy and indicate false positives/negatives.
6. **Review and thematic analysis:** Combined manual and ML assessments to refine prompts and the assessment framework. (Following the review and thematic analysis step, two more are considered as part of the methodology: [7]. **Vendor feedback and corrective action on aspects flagged by the assessment**; and [8]. **Processing vendor feedback** which is taken into consideration before the final evaluation and ranking. In this paper we only present the process of work from 1-6.)

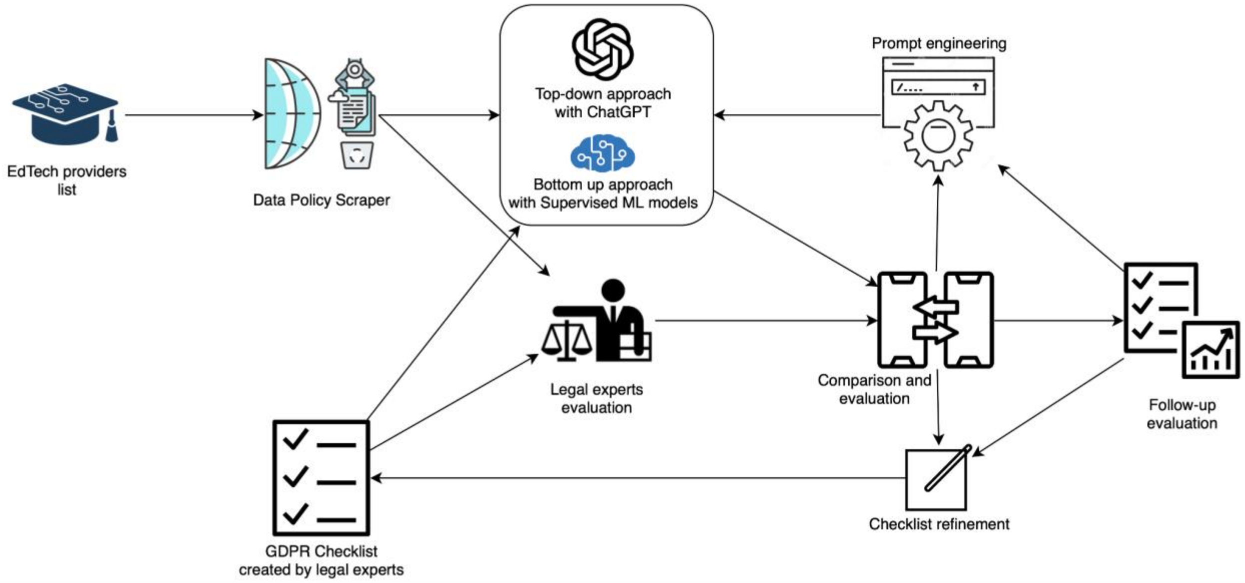

EdTech providers list
Data Policy Scraper
Top-down approach with ChatGPT
Bottom up approach with Supervised ML models
Prompt engineering
/....
Legal experts evaluation
Comparison and evaluation
Follow-up evaluation
GDPR Checklist created by legal experts
Checklist refinement

Figure 5. Evaluation methodology

3.3.1. Sample Selection and Operationalisation of Privacy Requirements

Following the initial exploration with bottom-up approaches, we aimed to develop an innovative ML-based method to assess the DPPs of any EdTech provider basing our research on 10 EdTech vendors popular in the UK. We selected them randomly, to ensure a mix of company sizes and user base. We gathered and analysed their publicly available DPPs and evaluated them against specific content criteria relevant to our study. Our focus was on how these documents address UK and EU data privacy standards specifically.

Next, to test the solution, we created a list of 44 questions/prompts aligning with key requirements under the EU GDPR (Regulation (EU) 2016/679) and UK GDPR and tried to assess whether the privacy policies provided by the selected Edtech providers address these aspects. In other words, the analysis sought to establish whether legally- required information is clearly articulated in the policies and whether other desirable information is also disclosed even though companies are not legally bound to do so but nevertheless explicitly provide to enhance trust. In Table 1, we outline some of the main themes and prompts without being exhaustive (due to space limitation). For a full list see GitHub

(2024). In the right-hand column, we outline which information is required by law and which is desirable to be demonstrated.

Overarching themes as per GDPR and related questions (not an exhaustive list of all our queries)	**Is compliance articulation obligatory? [Yes/No]**
Data processing purpose and lawfulness	
Do you provide the information about the identity and the contact details of the controllers and, where applicable, of the controller’s representative/DPO?	Yes
Is the purpose of processing the data identified in the DPP?	Yes
Does the privacy policy identify the lawful basis for processing personal data under the GDPR?	Yes
If your legal basis is legitimate interest (Article 6(1)(f)), do you define what the legitimate interest pursued by you or by a third party is?	Yes
Does the privacy policy clearly state that processing involves special categories of personal data?	Yes
Does the privacy policy outline the personal data collected and stored in connection with the intended purposes for processing?	Yes
Rights of data subjects	
Does the privacy policy outline the data subjects' rights to erasure and rectification of the data?	Yes
Does the privacy policy outline the data subjects’ right of access in accordance with Article 15?	Yes, if cookies are processed
Does the privacy policy outline the right to restriction of processing where the accuracy of personal data is contested?	Yes
Does the privacy policy outline the right to restriction of processing where processing is no longer necessary or	Yes

lawful? Does the privacy policy outline how data subjects can exercise this right?	
Consent and Notices	
When applicable, does the privacy policy indicate whether the provision of personal data is a statutory or contractual requirement?	Yes
When consent is the legal basis for processing, does the privacy policy outline the right to withdraw consent at any time?	Yes
When consent is the legal basis for processing children's data, does the privacy policy outline that the consent of the holder of parental responsibility is required?	Yes
Do you provide appropriate safeguards to secure international data transfers? If so, do you provide the information what kind of?	No, but desirable
Does the privacy policy outline the intention to transfer personal data to third countries or international organisations?	Yes

Table 1. Examples of prompts produced for the manual assessment and further ML-based assessment.

These questions serve as the foundation for our evaluation of the privacy policy documents' effectiveness in addressing regulatory requirements. We also identified other desirable aspects which should be considered in the DPPs, which stem from the AADC and, according to the UK ICO, should be considered by EdTech providers (ICO 2023). The AADC offers 15 flexible standards, emphasising children's rights, age-appropriate designs and transparency measures. Nevertheless, at this stage of our study, we focus on the compulsory requirements, therefore we assess the DPPs of selected vendors in line with the GDPR-oriented prompts and aim at further assessment including desirable aspects from the AADC in the future.

Furthermore, we explore four distinct approaches for leveraging OpenAI's API to analyse these documents - from direct API calls to frameworks like LangChain and RAG systems. Each approach offers unique advantages in processing and interpreting documents content (Filipovska et al., 2024; Filipovska et al., 2024). Through systematic evaluation, we make sure that the LLM/ML generated output has consistent format and that each 'Yes' answer, denoting that the policy at hand complies with a specific aspect represented by the question, also provides an extract of the policy that is most relevant to it.

By focusing on the documents themselves, we aim to uncover nuanced insights that inform decision-making and optimise outcomes in document governance and compliance. An alternative and very successful approach in analysing legal documents, such as terms of services is presented in (Binns & Matthews 2014). That said, the central point is on testing AI technology for scaling otherwise a manual and highly nuanced assessment of specialised literature that typically is costly to do and takes a long time, rather than encouraging or suggesting that AI technology alone can justify if DPPs adequately reflect companies' compliance, transparency or accountability obligations.

4. FINDINGS AND DISCUSSIONS

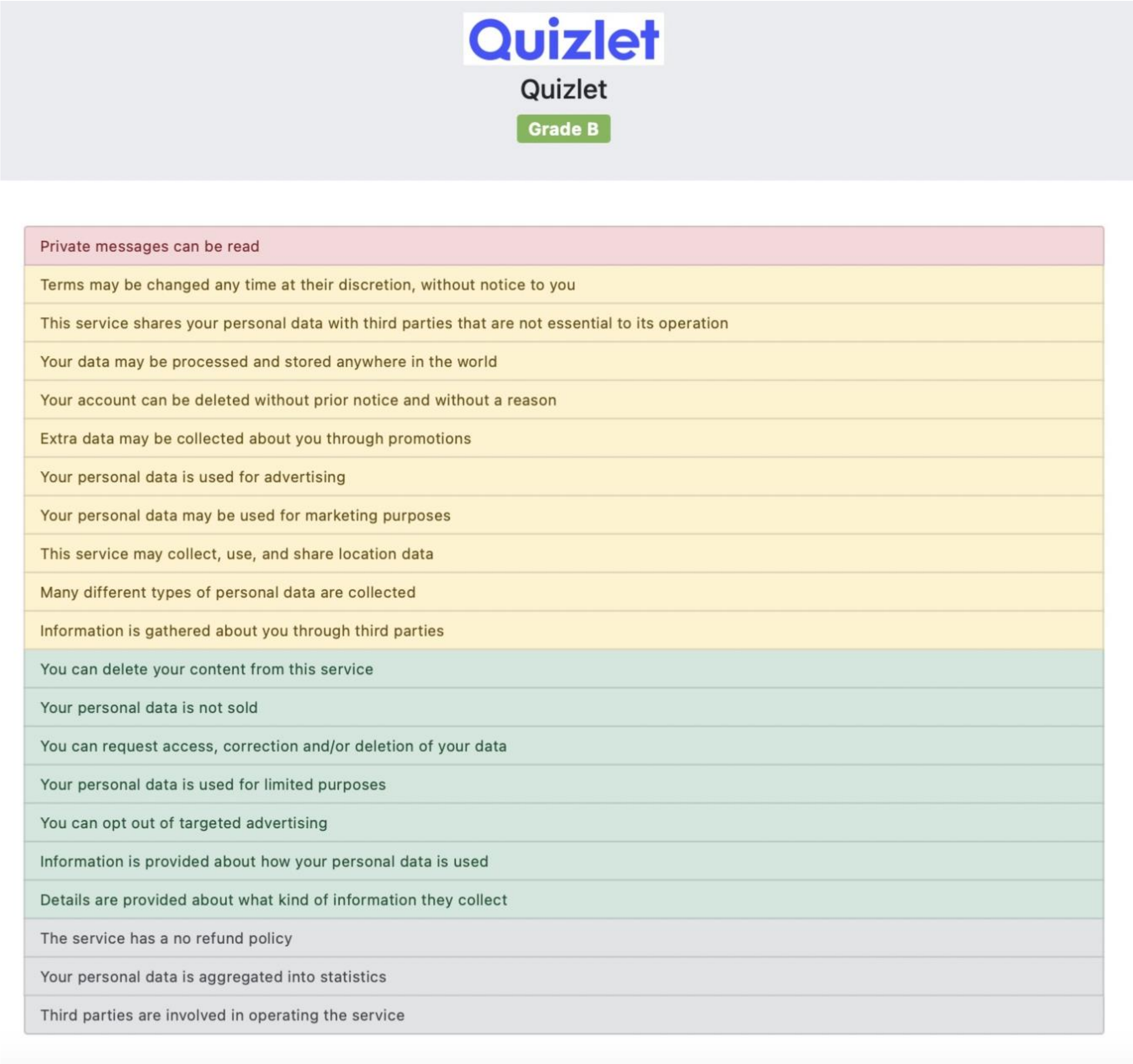
4.1.Results from the ToS;DR bottom-up tests

In this section, we briefly outline the findings from using the ToS;DR bottom-up community approach to semi-automating the DPP assessments and provide a couple of examples. The table below shows counts for the number of positive, neutral, negative, and blocker points found in each EdTech's DPP, along with a final grade using the ToS;DR formula, based on all 128 ToS;DR Cases.

Service	**Grade**	**Positive Cases**	**Neutral Cases**	**Negative Cases**	**Blocker Cases**
Century Tech	E	15	17	27	0
Doodle	D	12	10	19	0
Educake	E	12	28	25	1
Lexia	C	15	13	14	0
Quizlet	E	20	29	35	2

Table 2: Some initial results from the ToS;DR bottom-up community assessments

Additionally, the results are presented to public communities visually (Figure 6) with the relevant statements that affect the final grade. See example with Quizlet (the app mentioned earlier in Figures 2 and 3). Full list of the EdTech providers assessed via ToS;DR can be found on GitHub (2024).



Quizlet

Quizlet

Grade B

Private messages can be read

Terms may be changed any time at their discretion, without notice to you

This service shares your personal data with third parties that are not essential to its operation

Your data may be processed and stored anywhere in the world

Your account can be deleted without prior notice and without a reason

Extra data may be collected about you through promotions

Your personal data is used for advertising

Your personal data may be used for marketing purposes

This service may collect, use, and share location data

Many different types of personal data are collected

Information is gathered about you through third parties

You can delete your content from this service

Your personal data is not sold

You can request access, correction and/or deletion of your data

Your personal data is used for limited purposes

You can opt out of targeted advertising

Information is provided about how your personal data is used

Details are provided about what kind of information they collect

The service has a no refund policy

Your personal data is aggregated into statistics

Third parties are involved in operating the service

Figure 6: Quizlet's DPP assessments through community endeavour and Machine Learning

4.1.1. Limitations from using the ToS;DR model

Considering the community response and the maturity of the ToS;DR model, these initial findings provided support for progressing with the methodology and framework development (see next section). However, the lack of specialisation in the unique needs and challenges of the EdTech and education sector specifically, the ToS;DR model had its limitations.

First, its grading framework was originally developed for general-purpose web-based services and apps. Second, as a community-based bottom-up approach, it has no strict academic basis for scoring methodology. Rather, the Cases have been primarily based on intuition (including from some domain experts), with adjustments to achieve a wider balance of grades across all services on the ToS;DR platform. Lastly, the model outputs were checked for false positives, but there was no audit for true negatives. Following this initial test, we advanced in our top-down methodology, as described in the next section.

An additional limitation to the technical approach of fine-tuning BERT base models is the requirement of a training dataset. Our generally high accuracy was achieved in part thanks to the years of ToS;DR volunteers submitting analyses. Extending this methodology to new EdTech-specific Cases would also require a training set of human annotations.

4.2. Findings from the top-down approach: initial insights

To support education stakeholders in selecting Edtech vendors who are transparent about their data processing activities, we created a list of questions which should be answered in the DPPs. Such a list can facilitate the assessment of the Edtech vendors' compliance by users who are not familiar with legal matters and legal language. Based on the initial manual analysis of selected cases, prior to the automated assessment, which was conducted in the next step, we identified several issues which are common in the scanned policies:

a. The information provided is very general and does not present relevant details on how specific mandatory data protection requirements are addressed by the Edtech vendor.

    - ***Example:*** Selected DPPs include the statements that the data minimisation is ensured by the providers, nevertheless, they do not describe how this principle is applied in practice.

b. Not every Edtech provider that was assessed shares clear information on the data retention periods and what will happen with the data after the retention period expires.
    - ***Example:*** Math-Whizz provides that the data will not be kept for any longer than is necessary in the light of purposes of processing but does not provide the information on a specific data retention period. Renaissance Learning provides rather generic clauses in regard to data storage and no mention of the circumstances under which stored personal data will be deleted.

c. The information on data recipients, i.e. third parties who may receive the data from the Edtech provider, are not always clearly and transparently defined.
    - ***Example:*** Turnitin indicates that personal data will be shared with third-party service providers for services such as data hosting, analytics, content delivery or maintenance, however, the data recipients and the type of personal data shared are not explicitly mentioned.

d. The DPPs are not clear on whether users are subject to solely automated decision-making, whereas it is relevant for users, especially children, to transparently reveal whether such practices are applied or not.

e. The information on the transfers of the data to third countries is very often provided in a very generic manner and without further details on the countries considered and the legal basis for data transfers.

f. Not all DPPs provide clear information on the sources from which the data are obtained by the providers, if not from users directly.

Moreover, in selected cases the mandatory information, required by Articles 13 and 14 of the EU and UK GDPR, is not provided in one place, but users need to go through several documents to understand all details around vendors' data practices. Such an approach doesn't appear to be user-friendly, and anyone with a lower level of Internet literacy may have issues with understanding the full scope, risks and benefits of applied data processing. It should also be stressed that the above-mentioned aspects have an impact on a smooth automated, ML-based assessment of privacy policies. In addition, unlike for our bottom-up exploration using the ToS;DR ML methodology, fine-tuning base language models was not an option because we did not have a dataset of human annotations with which to train. Instead, we explored zero-shot and few-shot prompting of LLMs, specially OpenAI's, where a training dataset is not necessary. This also provides an opportunity to explore how the two ML approaches compare in terms of accuracy and limitations. Further details on the results are presented below.

4.2.1. ML assessment results from the top-down approach

Each of the 44 GDPR queries that we developed (see Table 3 for examples) were further taken up as prompts used in the assessment of the selected Edtech DPPs facilitated by ChatGPT, and assigned the IDs from GDPR1 to GDPR44. The final results were provided in the following manner:

policyid	**GDPR1_answer**	**GDPR1_extract**
Link to the respective file assessed by ChatGPT	Yes / No	Extract from the policy which the answer was based on.

Table 3. Results provided by ChatGPT-based crawling.

Based on whether the provider meets each requirement or not a score (answer) is given. The final score is evaluated across a spectrum of how transparent the company is (note: not how ethical, lawful, and secure it is) and where the companies' communications through DPPs lag in terms of transparency, accuracy, readability and in relation to various fundamental requirements.

4.2.2. Limitations

While ML models provide numerous opportunities which we took advantage of in our research, the current capabilities of the likes of ChatGPT models are not there to replace legal experts yet. The automated evaluations of Edtech vendors’ privacy policies were far from reliable. In order to understand better, what kind of mistakes ChatGPT makes and what questions and prompts among the ones required by law and desirable, stemming from the GDPR, pose the biggest challenge to ChatGPT, we developed a typology of errors (see Table 4). Errors could appear in the “score” assessing whether a score was met (Yes/ No) or in the justification provided by ChatGPT depicting which evidence the score is based on.

Error Code	Name	Explanation
E1	Relevant passage not found	The score is incorrect, and no justification or evidence is provided.
E2	Justification irrelevant	The score is correct or incorrect, and the provided justification or evidence does not relate to the topic in question.
E3	Arbitrary evaluation	Very similar policies are evaluated/scored differently. For example, the same information on how to exercise data subject rights is provided for each right, and nevertheless the scores differ.
E4	Interpretation of abstract concepts required	The evaluation of the GDPR requirement requires the interpretation of an abstract concept, such as fairness, lawfulness, etc. The justification is not convincing, incomplete or vague.
E5	Justification too short	The justification or evidence provided is too short or incomplete. For example, a justification mentions only parts of the relevant sections in the DPP.
E6	Justification misleading	The evidence or justification provided was misleading, e.g. because important elements were missing
E7	(Legal) reasoning	There is an error in the reasoning of the justification, for example saying the UK is in the EU
E8	Information is not obligatory	The lack of information is marked as incompliant, although the information provision is (or may not) be obligatory

Table 4: Typology of errors included in the ChatGPT assessment

The accuracy of the automated evaluation of policies differed. Below we present selected examples:

a) In ChatGPT's assessment of the DPP of Doodle Learning[10] 28 of 44 scores were correct. Two scores were false positives (ChatGPT set a better score than the privacy experts), twelve scores were false negatives (ChatGPT set a worse score than the privacy experts), and four scores were correct but the evidence/ justification of ChatGPT was incomplete or misleading.

b) In the assessment of Whizz Education DPPs[11], only 4-8 scores of 44 scores were correct, depending on the DPP Whizz Education privacy policy is divided into several files: a general Data Protection Policy and further separate Privacy Statements for different types of users: teachers[12] parents[13] and students[14]). The number of false positive and false negative results was almost equal, between 9 and 16 such scores per assessed file, and there was a third score identified - "false in a different manner". The third type of score was assigned in the cases where the score provided by Chat GPT was positive, but the errors E2, E4 or E6 appeared, i.e. the evidence/justification provided by ChatGPT was incomplete or incorrect, however, the file did include the right justification which was not correctly detected by the AI model.

Based on the assessment of the DPPs of 10 EdTech providers, we identified that the most common type of error was that ChatGPT provided excerpts from the DPP that were irrelevant for the question as evidence for the score, leading to the assumption that even a correct score was not necessarily an indicator for mastery of the task. Additionally, evaluating a DPP requires the interpretation of abstract concepts such as "legitimate, fair and transparent data processing" and normative judgments on whether a question is sufficiently answered. In these cases, it required a justification of the score that

[10] Doodle Learning - Privacy Notice. Available at: https://www.doodlemaths.com/privacy-policy/

[11] Whizz Education Data Protection Policy. Available at: https://whizz.com/wp-content/uploads/Data-Protection-Policy-1.2.23.pdf.pagespeed.ce.FXzmF18GCb.pdf

[12] Privacy Statement for Teachers and Staff in Maths-Whizz Schools. Available at: https://whizz.com/wp-content/uploads/Privacy-Statement-Teachers-2024-25-EEF-Study-Only-3.pdf

[13] Privacy Statement for Home Subscribers and Parents/Guardians of School Subscribers to Maths-Whizz. Available at: https://www.whizz.com/wp-content/uploads/2022/12/Privacy-Statement-for-Home-Subscribers-1-2-22.pdf

[14] Privacy Statement for Maths-Whizz Students. Available at: https://whizz.com/wp-content/uploads/Privacy-Statement-Students-2024-25-EEF-Study-Only-3.pdf

oftentimes remained vague and didn’t provide sufficient proof, so that one could trust the score.

A struggle for ChatGPT were also requirements that only relate to certain types of data processing. Hence, ChatGPT must understand first if a requirement is compulsory based on the information provided, e.g. whether sensitive data types are processed. Only then can further be assessed if relevant information pertaining to this type of processing is revealed in the policy or not, as some requirements are compulsory to be addressed only with regard to a specific type of data. A general problem when evaluating privacy policies stems from the fact that one has no insight in the actual data practices of the company. Therefore, if no information is provided one cannot evaluate based on the policy, whether a certain requirement is obligatory and the company is not compliant, or whether the company is providing no information, because they are not conducting any processing of that type.

Furthermore, ChatGPT did not consider that there may be additional policy statements e.g. pertaining to Cookies, and the content of those could be provided as a hyperlink, and not directly included in the text of the DPP. ChatGPT did not reason that there are dependencies between different files (i.e. pieces of privacy policies) made available by an Edtech provider on separate subpages, which should be read and understood in conjunction, and not as standalone information. This was the issue identified e.g. in the context of Math-Whizz’s DPP, which was further complemented by user-oriented privacy statements.

Given these selected examples we can conclude that, although the ML-based assessment would be very beneficial and efficient, it still does not allow for a comprehensive and thorough evaluation of the information included in privacy policies and can lead to a misleading conclusion on the aspect whether a given Edtech provider does comply with the requirements posed by the law in regard to the data processing information with which the user should be equipped.

5. CONCLUSIONS

The above examples of different types of privacy policy assessments - manual and ML-based - provided different findings concerning the transparency and accuracy of the information on data protection practices applied by the Edtech providers.

The manual assessment has proven that the DPPs which were screened could be still improved with regard to the relevant details on data processing activities, especially when it comes to the information on the means and the manner of addressing data protection principles defined in Article 5 (UK and EU) GDPR, such as confidentiality, integrity, data minimisation, storage limitation and accuracy.

The ML-based assessment indicated that a completely automatic and automated assessment can lead to very crucial errors and a misleading and incorrect conclusion that the information shared with the user meets the mandatory criteria imposed by the laws and provides full transparency concerning the data processing practices.

There are two points of contention that we wish to acknowledge. Firstly, there is no legal obligation for entities to share all details about data processing either with the users or publicly and for that we take a more balanced view by considering the legal obligations alongside practical considerations like various, often conflicting interests (e.g., businesses' vs. children's) and confidentiality.

Secondly, another pitfall of the discussion is that, in practical sense, there is no way to fully assess complete compliance based on DPPs and the law does not require companies to demonstrate total compliance in these documents. To this end, external audits, which delve into a vast number of an organisation's processes, can help organisations show they are compliant. Various audit frameworks exist already. They provide assurances to users and are beneficial for developing customer trust. What is required however is industry's commitment to undergo scrutiny and drive towards understanding the laws and prioritising children's wellbeing and rights.

In future research, an opportunity will be to expand the current analysis of DPPs by incorporating a more comprehensive examination of compliance with various codes and standards. It will be useful to the education sector to investigate the bilateral and other agreements between Edtech companies and schools to understand the contractual obligations and privacy guarantees provided. Long-term goals can address how Edtech vendors comply with various standards and frameworks, such as the AADC, digital accessibility, cybersecurity and pedagogic frameworks. Leveraging AI with human assessors at scale offers promising potential for future studies. Future research can advance the present work by developing specific prompt templates to streamline the assessment process across diverse requirements, and enhance the rigour of the assessments.

Acknowledgements: The authors express their gratitude to Anwana Tima Otu and Emily Johnson for their valuable support in the preliminary stages of this research.

References

Article 29 Data Protection Working Party. (2017). Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679. Available online: https://ec.europa.eu/newsroom/article29/items/611236.

Barassi, V. (2020). Datafied times: Surveillance capitalism, data technologies and the social construction of time in family life. *New Media & Society, 22*(9), 1545-1560. https://doi.org/10.1177/1360780419857734

Binns, R., & Matthews, D. (2014). Community structure for efficient information flow in 'ToS;DR', a social machine for parsing legalese. In *Proceedings of the 23rd International Conference on World Wide Web* (pp. 881–884).

Citron, D. K., & Solove, D. J. (2021). Privacy harms (GWU Legal Studies Research Paper No. 2021–11). *Boston University Law Review, 102*(2). https://doi.org/10.2139/ssrn.3782222

Committee on the Rights of the Child. (2013). General comment No. 14 (2013) on the right of the child to have his or her best interests taken as a primary consideration (art. 3, para. 1). Available online: https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=CRC%2fC%2fGC%2f14&Lang=en.

Committee on the Rights of the Child. (2021). General comment No. 25 (2021) on children’s rights in relation to the digital environment. Available online: https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=CRC/C/GC/25&Lang=en.

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *North American Chapter of the Association for Computational Linguistics*.

EU EdTech Alliance. (2022). First results from the European EdTech ecosystem map. Available online: https://www.EdTecheurope.org/news/first-results-from-the-european-EdTech-ecosystem-map.

European Data Protection Board. (2021). Guidelines 07/2020 on the concepts of controller and processor in the GDPR. Version 2.1. Adopted on 07 July 2021. Available online: https://edpb.europa.eu/system/files/2023-10/EDPB_guidelines_202007_controllerprocessor_final_en.pdf

Filipovska, E., Mladenovska, A., Bajrami, M., Dobreva, J., Hillman, V., Lameski, P., & Zdravevski, E. (2024). Benchmarking OpenAI's APIs and Large Language Models for Repeatable, Efficient Question Answering Across Multiple Documents. In *Proceedings of the 19th Conference on Computer Science and Intelligence Systems (FedCSIS)*.

Filipovska, E., Mladenovska, A., Dobreva, J., Kitanovski, D., Mitrov, G., Lameski, P., & Zdravevski, E. (2024). Evaluation of vector databases and LLMs in RAG-based multi-document question answering. In *ICT Innovations 2024: Tech Convergence: AI, Business, and Startup Synergy*. Springer.

GitHub. (2024). *EdTech policies and ToS;DR findings*. https://github.com/admin-magix/EdTech-policies

HolonIQ. (2021). Global education technology in 10 charts. Available online: https://www.holoniq.com/EdTech-in-10-charts.

Human Rights Watch (HRW). (2022). ‘How dare they peep into my private life?’ Children’s rights violations by governments that endorsed online learning during the covid-19 pandemic. Available online: https://www.hrw.org/report/2022/05/25/how-dare-they-peep-my-private-life/childrens-rights-violations-governments.

Information Commissioner's Office [ICO]. (2018). Applications - children and the GDPR. Available online: https://ico.org.uk/media/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/children-and-the-gdpr-1-0.pdf.

ICO. (2021). Guide to the General Data Protection Regulation. Available online: https://ico.org.uk/media/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr-1-1.pdf

ICO. (2022). What does it mean if you are a controller? Version 1.0.12. Available online: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/controllers-and-processors/controllers-and-processors/what-does-it-mean-if-you-are-a-controller/#1.

ICO. (2023). The Children's code and education technologies (EdTech). Available online: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/childrens-code-guidance-and-resources/the-children-s-code-and-education-technologies-EdTech/.

International Digital Accountability Council (IDAC). (2020). Privacy in the Age of Covid: An IDAC Investigation of Covid-19 Apps. Available online: https://digitalwatchdog.org/wp-content/uploads/2020/07/IDAC-COVID19-Mobile-Apps-Investigation-07132020.pdf.

Kerssens, N., Nichols, T. P., & Pangrazio, L. (2023). Googlization(s) of education: Intermediary work brokering platform dependence in three national school systems. *Learning, Media and Technology*, 1–14. https://doi.org/10.1080/17439884.2023.2258339.

Kucirkova, N. I., Cermakova, A. L., & Vackova, P. (2024). *Consolidated benchmark for efficacy and effectiveness frameworks in EdTech*. https://doi.org/10.31265/USPS.270

Livingstone, S., Pothong, K., Atabey, A., Hooper, L., & Day, E. (2024). The Googlization of the classroom: Is the UK effective in protecting children's data and rights? *Computers and Education Open*, 7, 100195. https://doi.org/10.1016/j.caeo.2024.100195

Lupton, D., & Williamson, B. (2017). The datafied child: The dataveillance of children and implications for their rights. *New Media & Society, 19*(5), 780-794. https://doi.org/10.1177/1461444816686328

Manwaring, K., Kemp, K., & Nicholls, R. (2021). (Mis)Informed consent in Australia. UNSW Law Research. Available at SSRN: https://ssrn.com/abstract=3859848 or http://dx.doi.org/10.2139/ssrn.3859848.

Ranking Digital Rights. (2020). *2020 Ranking Digital Rights Corporate Accountability Index*. https://rankingdigitalrights.org/index2020/

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (United Kingdom General Data Protection Regulation). Available online: https://www.legislation.gov.uk/eur/2016/679/contents.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), OJ L 119, 4.5.2016, p. 1–88. Available online: https://eur-lex.europa.eu/eli/reg/2016/679/oj.

Statista. (2023). Online education worldwide. Available online: https://www.statista.com/outlook/dmo/eservices/online-education/worldwide#:~:text=Revenue%20in%20the%20Online%20Education,US%24239.30bn%20by%202027.

Stoilova, M., Livingstone, S., & Nandagiri, R. (2020). Digital by default: Children's capacity to understand and manage online data and privacy. *Media and Communication, 8*(4), 197-207. https://doi.org/10.17645/mac.v8i4.3407

The Office of the High Commissioner for Human Rights. (2020). OHCHR Dashboard - Ratification of 18 International Human Rights Treaties. Available online: https://indicators.ohchr.org/.

Van Der Hof, S., Lievens, E., Milkaite, I., Verdoodt, V., Hannema, T., & Liefaard, T. (2020). The child's right to protection against economic exploitation in the digital world. *The International Journal of Children's Rights, 28*(4), 833-859. https://doi.org/10.1163/15718182-28040003.

Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society, 12*(2), 197-208. https://doi.org/10.24908/ss.v12i2.4776.

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology, 30*(1), 75-89. https://doi.org/10

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power.* Public Affairs.