INTERNET POLICY REVIEW
Journal on internet regulation

Volume 5 | Issue 1

Regulating "big data education" in Europe: lessons learned from the US

Yoni Har Carmel
Haifa Center for Law and Technology, University of Haifa, Haifa, Israel

Published on 31 Mar 2016 | DOI: 10.14763/2016.1.402

Abstract: European schools are increasingly relying on vendors to collect, process, analyse, and even make decisions based on a considerable amount of student data through big data tools and methods. Consequently, portions of school's power are gradually shifting from traditional public schools to the hands of for-profit organisations. This article discusses the current and forthcoming European Union (EU) data protection regime with respect to the protection of student rights from the potential risk of outsourcing student data utilisation in Kindergarten-12th grade (K-12) educational systems. The article identifies what lessons can be drawn from recent developments in the United States (US) "student data affair". These lessons can provide a new perspective for designing a balanced policy for regulating the shift in school's power.

Keywords: Education, Learning analytics, Student data, Big data

Article information

Received: 18 Dec 2015 **Reviewed:** 16 Feb 2016 **Published:** 31 Mar 2016

Funding: This research was supported by the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation (1716/20).
Competing interests: The author has declared that no competing interests exist that have influenced the text.

URL:
http://policyreview.info/articles/analysis/regulating-big-data-education-europe-lessons-learned-us

Citation: Har Carmel, Y. (2016). Regulating "big data education" in Europe: lessons learned from the US. *Internet Policy Review*, *5*(1). DOI: 10.14763/2016.1.402

This paper is part of 'Big data: big power shifts?', a Special issue of Internet Policy Review, supported by the Vodafone Institute for Society and Communications.

1. INTRODUCTION

Since the early days of the modern educational system, schools, almost exclusively, have been entrusted with the responsibility and authority to educate children by shaping their skills, values and knowledge.

In this capacity, schools have relied on student information to effectively administer and

improve learning (Polonetsky & Tene, 2014 [PDF]). The rapid development and the ubiquity of digital technologies have dramatically changed the classroom experience and generated an unprecedented “explosion of student data” that opens up a new world of potential evidence on how students learn (Polonetsky & Jerome, 2014).

Nowadays, students can access a wide range of learning resources, interact with a variety of applications, enhance their experience in virtual environments, augment reality and communicate with others through different platforms (Pardo & Siemens, 2014). The interaction with these innovative tools generates a vast amount of granular learning-related data that opens up a new world of potential evidence on how students learn (The White House, 2014 [PDF]).

The generation, accumulation, processing and analysis of student data made available is now being touted as a potential panacea for many current educational challenges and problems (Selwyn, 2015). Indeed, the limitations of conventional applications in utilising the enormous quantities of data, have caused schools and reformers to increasingly explore “big data” techniques, such as learning analytics (LA)[1] and educational data mining (EDM)[2]. Their goal is to make the available data an integral part of planning, designing and assessing the learning experiences (Baker, 2014 [PDF]).[3]

Incapable of independently pursuing these goals, schools have come under increased pressure to turn to the private sector for the provision of big data techniques; relying on vendors of commercial technology in the so-called “EdTech” industry for the mining, collecting, processing and analysing of student data, and even for applying data-driven decision-making processes (Charlton, Mavrikis, & Katsifli, 2013).[4]

The outsourcing of student data utilisation affords many educational opportunities. However, authorising commercial corporations to perform school’s core functions establishes a new power-balance and may run the risk of facing unintended consequences that affect students’ rights and liberties.

Strong emotions and high stakes have created a polarised debate surrounding student data use in the US, which subsequently led to a deluge of state and federal regulatory reforms seeking to protect students and allay the public deep concerns.

This article will discuss the capacity of EU law to address the various concerns raised with regard to the turn towards “big data education”. I begin with a description of the shift in power from traditional schools to for-profit organisations resulting from the introduction of big data techniques into school’s core functions. Thereafter I present the recent public discourse in the US over the protection of students’ rights from the potential risks of outsourcing student data utilisation, and the subsequent regulatory actions that were taken, illustrating well the sectoral approach of US regulation. Then, I give an overview of the principles of the current and the forthcoming EU student data protection regime that follow an encompassing regulatory approach, and I discuss their application in the context of big data technologies in education, pointing out their key shortcomings. Finally, I conclude by identifying what lessons can be drawn from the developments in the US to strengthen the regulatory protection of students’ rights in the EU following the introduction of “big data education”.

2. DATA-DRIVEN SHIFT IN SCHOOL'S POWER

In this article, big data is defined as "*large and complex datasets collected from digital and conventional sources that are not easily managed by traditional applications or processes*" (Reyes, 2015).

In the educational context, big data technologies collect vast amounts of student data from multiple sources and subject them to analysis using data processing algorithms. Student data may include basic academic and administrative information, but also 'data traces' and metadata from students' interaction with educational digital platforms as well as unexpected sources like student ID badges and social media. This digital data can range from online test scores, to session times, to records of where a student has clicked or touched while figuring out a problem (Williamson, 2015).

Realising the profit potential, vendors - such as Pearson, McGraw-Hill and Knewton - have been offering schools (sometimes free of charge) a wide variety of big data-based technologies that can be applied in all aspects of digital education. These can afford technology-enhanced pedagogical applications, such as:

Personalised learning: big data can drive personalised learning that goes beyond tailoring instruction to what students know, but also to how they learn based on needs, preferences, aspirations or cultural background (Mayer-Schönberger & Cukier, 2014).

Adaptive learning: adaptive learning systems can continuously collect and interpret student data to change the learning course and environment based on the individual's needs and abilities (US Department of Education, 2013 [PDF]).[5]

Accurate assessment: with the capability of observing students while they work on an activity, it is possible to deploy new assessment techniques which measure achievements more accurately (Polonetsky & Jerome, 2014).

Effective feedback: big data can provide a more intelligent and effective feedback-loop where students receive information in a short amount of time in response to their input (Weber, 2015).

Performances prediction: students' behaviour, skill and performance can be predicted by analysing various activities performed while interacting with digital platforms so that the instructors can focus on developing underperforming students (Charlton, Mavrikis, & Katsifli, 2013).

Traditionally, schools have been empowered with the capacity to "*manage, administer, discipline, shape, care for and enable*" students through various techniques, including pedagogy (Pykett, 2012). As education is becoming more web-based and data-driven, larger portions of schools' core functions, i.e., teaching, learning, and assessment (Burch & Good, 2015) are gradually being outsourced to vendors. This brings a power shift from traditional schools to the hands of for-profit organisations which now possess the capacity to collect, mine, process, analyse, and even make educational decisions based on student data (Charlton et al., 2013).

3. THE AMERICAN STUDENT DATA PROTECTION "UPROAR"

Using new technologies or services offered by vendors in order to improve learning processes is certainly not new to the education domain (Polonetsky & Jerome, 2014). Nonetheless, the combination of more technology and a reliance on private vendors has raised wide-ranging concerns in the US (Chui & Sarakatsannis, 2015).

In 2011, inBloom, a non-profit data analytics company, designed an advanced secure service offering states and school districts to store data and connect to a personalised learning software. By mid-2013 inBloom provided its services to nearly every public school in New York State. But for many the software got a little too personal.

Although there was no evidence of inBloom misusing the information, parents and privacy advocates raised concerns about the scope of inBloom’s potential data collection. Following some negative campaigns led by privacy activists, parents and teachers’ groups, inBloom announced in April 2014 that it would be shutting down its operations (Bennett & Weber, 2015).

CONCERNS OVER THE POWER SHIFT FROM SCHOOLS TO VENDORS

Big data encapsulates two significant components: huge quantities of varied data and large-scale analytics. To achieve both the touted benefits and anticipated harms of big data in education, a vendor needs to utilise a significant analysable quantity of student data (Young, 2015).

The demise of inBloom, and frequent media reports of data security breaches, gave rise to increasing concerns in the US over student data protection (Young, 2015 [PDF]).

Critics, mainly parents and educational and privacy advocacy groups, have been concerned that the large dissemination of student data to private vendors might risk students' privacy, disclosing sensitive information about children, like data about learning disabilities, disciplinary problems or family trauma (Singer, 2014). Of particular concern is the likelihood that vendors will improperly “mine” or sell student data or otherwise monetise student information through building advertising profiles or marketing (Herold, 2014).[6]

Critics have also been concerned that constant monitoring of students’ online activities may overly limit creativity, free speech and free thought, by creating a “surveillance effect” (Zeide, 2016) and invading their “intellectual privacy” i.e. “*the ability, whether protected by law or social circumstances, to develop ideas and beliefs away from the unwanted gaze or interference of others*” (Richards, 2015).

Another prominent fear concerns big data techniques prematurely and permanently labeling students as underperformers which may "*forestall future opportunities by becoming a modern day version of the proverbial permanent record*" (Zeide, 2015). The identification of students as “at risk”, for example, might not allow them to remove any harmful record of their failures if they improve in the future. Consequently, “*students may see labels as self-fulfilling prophecies and predictive analytics may prime educators to make prior judgments about students’ capabilities and character*” (Alarcon et al., 2014).

Furthermore, critics have been concerned that continuous student data mining coupled with decision-making based on algorithmic models will exacerbate bias and create new forms of discrimination, resulting from the embedment of arbitrary or unfair factors (MacCartney, 2014

[PDF]). Grounding decision-making on objective information retrieved by algorithms from multiple educational sources, and based on students' performance in a wide array of educational contexts, may appear "neutral" and irrefutably scientific. However, the "hidden" algorithms that facilitate educational data-driven decision-making reflect particular norms and values about what educational opportunity and equity means. As such, they may rely on biased data that reflect social inequality and plausibly reinforce present structural inequities and contribute to a problem of cumulative disadvantage (Alarcon et al., 2014).

REGULATORY REFORMS TO PROTECT STUDENT DATA IN THE US

The concerns over the engagement of for-profit third parties in education, through the utilisation of student data, revolve around how and for which educational and non-educational purposes data is collected, processed and analysed.

Legal frameworks that apply to student data held by schools and vendors acting on their behalf exist primarily in three US federal statutes which focus mainly on protecting student privacy by limiting access to and disclosure of data:

a. The Family Educational Rights and Privacy Act of 1974 (FERPA) prohibits the unauthorised disclosure of education records. FERPA applies to any school receiving federal funds and levies financial penalties for non-compliance;
b. The Protection of Pupil Rights Amendment (PPRA) of 1978 governs the administration of surveys soliciting specific categories of information, and imposes certain requirements regarding the collection and use of student information for marketing purposes; and
c. The Children's Online Privacy Protection Act of 1998 (COPPA) which applies particularly to online service providers that have direct or actual knowledge of users under 13 and collect information online.

Enacted over four decades ago, FERPA was not created for a world where data flows freely and where third parties who are not educational actors become integral part of day-to-day information flow. Even since COPPA came into effect in 2000, education technology has changed radically (Krueger, 2014).

While FERPA was groundbreaking privacy legislation when it was enacted over four decades ago, it is inadequate in today's world where data flows freely and where third parties who are not educational actors become an integral part of day-to-day information flow. "*FERPA is so dated that when confronted with a technology that can collect and use big data… the statute practically breaks down*," says Young (2015). For example, the definition of educational record and personally identifiable information (PII) would likely not include unconventional types of student data collected through EdTech, such as a lunch item choice or the subject of an email message. Moreover, FERPA's "*school officials*" exception allows directory PII, such as students' names, addresses, and phone numbers, to be disclosed to third parties who have "*legitimate educational interests*" without parental consent if the school notifies parents of this practice once a year, and parents are given the opportunity to opt-out of this disclosure. The majority of EdTech providers arguably meet the "school official" exception because they are often under contract with a school to perform an institutional service or function.[7] Furthermore, FERPA puts the primary compliance burden on schools themselves, whereas vendors are not required to comply with the law's provisions (Center for Democracy & Technology, 2015 [PDF]).

PPRA requires that schools give notice to, obtain written consent from, and provide an opt-out opportunity to parents before students can participate in commercial activities that involve the collection, disclosure or use of personal information for marketing purposes.[8] Nevertheless, this rule does not apply if a vendor is using student data solely for the purpose of developing,

evaluating, or providing educational products or services to students or schools. Moreover, like FERPA, PPRA does not provide a private right of action, thus students and parents cannot enforce compliance with the statute (Tudor, 2015).

COPPA, as opposed to FERPA and PPRA, was not designed to be a student privacy law. Even though the law does help to ensure vendors collect and use student data responsibly, it is limited only to sites or services that collect information from children under 13 years old, and not information provided by adults about these children. Therefore, a vendor would not have to comply with the law if the data it collects on an under-13-year-old is only obtained from a parent, school, or presumably anyone of 13 years or older (which includes the majority of high school and some junior high school students).

In response to Americans’ persistent concerns, state legislatures began passing laws to fill the gaps in FERPA and other federal laws, as well as to extend privacy protections to other areas. In fact, by September 2015, 46 US states introduced 182 bills addressing student data privacy. Of these bills, 28 in 15 states were enacted into law (Data Quality Campaign, 2015 [PDF]).[9] Many of these bills focus on who can access student information and mandate that private entities only use student data for educational purposes. They often stipulate substantive restrictions on the use of student data for creating advertising profiles and for marketing purposes. A lot of the rules focus on providing more opportunities for notice and choice for parents to consent to particular uses or collection (Center for Democracy & Technology, 2015 [PDF]).

The growing student data-related concerns have also garnered attention from legislators of both houses of Congress, who stood on guard to protect student privacy “from the hands” of private vendors by introducing numerous bills.

In April 2015, the Student Digital Privacy and Parental Rights Act (SDPPRA), was introduced with the support of President Obama. The bill prohibits the use of students' PII for advertising and marketing purposes and seeks to minimise the amount of such information that is transferred from schools to private companies.

The bi-partisan Protecting Student Privacy Act (PSPA) was introduced in May 2015 by senators Ed Markey, Orrin Hatch and Mark Kirk. The bill proposes to amend FERPA to, *inter alia*, require schools to implement policies and procedures that protect students' PII; prohibit schools from knowingly providing access to PII for advertising or marketing purposes; and require states and schools to ensure that outside parties comply with specific requirements.

Perhaps the most rigid bill introduced in the Senate was by senator David Vitter. The Student Privacy Protection Act (SPPA) which would amend FERPA, takes a dramatically different tack than other student-data-privacy legislation that have previously appeared at the federal and state level. SPPA requires educational agencies and institutions to receive parental consent before sharing student data with third parties. It also, for the first time, allows for individual families to receive monetary awards from educational agencies and private actors that violate their children's FERPA rights.[10]

While it is unclear when, or if at all, they will be enacted, it can already be expected that the pending US federal bills will not quell the uproar or diminish the sizzling debate over the expanding role of for-profit companies in education.[11] Parents and privacy advocates have by now vigorously expressed their fears that the bills are inadequate to protect students’ rights.

Representatives of the Parent Coalition for Student Privacy, for example, raised alarms that

SDPPRA does not require any parental notification or consent before schools share personal data with third parties, allowing vendors to target ads to students and to continue collecting and sharing vast amounts of highly sensitive student information (Strauss, 2015). Pasquale (2015b) had also criticised the bill for focusing mainly on privacy issues, while not addressing other issues such as student profiling (e.g. “at risk” students).

PSPA, on the other hand, was criticised by privacy advocates for not holding vendors legally accountable, and for not expanding its definition of “educational records” to include student e-mails and digital metadata created on school provided services, platforms, and equipment (Roscorla, 2014).

Even SPPA, ostensibly the most comprehensive proposed policy change, suffered critique for being too lenient towards schools and vendors. In her critical analysis of the bill, Hoge (2015 [PDF]) argues that SPPA will increase psychological screening and profiling of those with disabilities by allowing special education teams to implement psychological testing, treatment, analysis, and evaluation, without parental consent. In addition, as Hoge argues, the bill will not decrease access to private data by third parties. When referring to PII, the bill creates protections for “student data” and then aligns to the definition already listed in FERPA for PII that allows directory information to be cross matched and used to identify the individual student. Moreover, according to the bill, third parties and “school officials” still have access to data because of written agreements in the original version of FERPA.

The description of the American case shows the variety of concerns over student data use, which go way beyond privacy. The US legislative attempts, however, focus on privacy (e.g., prohibition of ads targeting and disclosure of PII).

The protection of “student privacy” is, of course, a major concern when it comes to the potential risks of student data utilisation by vendors. But much of the debate about “student privacy” is not about privacy, and the term has actually become a rallying cry related to any issue involving data use in education. Some of the concerns are less about information practices than about education policy and pedagogy, including the “privatisation” of the public school system. Parents care most about whether their child receives a good education and they want to ensure that his safety and future opportunities are not compromised in pursuit of conflicting corporate interests (Zeide, 2016). Therefore, regulation should account for other student rights that might be jeopardised by the shift in power from schools to vendors, such as equality, autonomy and freedom of thought.

4. PROTECTING STUDENT DATA IN THE EU

In contrast to the US piecemeal approach to regulating data protection where legislation is sector driven and may be enacted at state and/or federal levels, personal data protection has been regulated in the EU for a long time, applying a comprehensive prescriptive legal approach which focuses on the population as a whole. For the reasons that will be outlined in the following section, it is nevertheless not sufficiently equipped to deal with the pitfalls of big data in education.

CURRENT EU STUDENT DATA PROTECTION REGIME

At present, the most important EU legal instrument on personal data protection is the 1995 Directive 95/46/EC on the protection of individuals with regard to the processing of personal

data and on the free movement of such data (DPD).[12]

Recognising the important role vendors play in processing personal data, the DPD distinguishes between first parties and vendors through the introduction of “data controllers” and “data processors” (art. 2(d)-(e)).

Within this structure, a school acts as a data controller if it decides on (a) outsourcing of student data processing; (b) delegating all or part of the processing activities to an external organisation; and (c) determining the ultimate purpose of the processing. A vendor acts as a data processor if it merely supplies the means and the platform, acting on behalf of the school (Article 29 Working Party, 2012 [PDF]).

Deemed data controllers, schools must abide by data protection legislation and must adhere to basic principles of the DPD. Without entering into a discussion as to the effect of holding schools accountable for the actions of third parties, the DPD has two key drawbacks in protecting student privacy and personal data in the context of “big data education”.

First, the DPD does not protect student data from re-identification. The DPD's definition of personal data is: “*any information relating to an identified or identifiable natural person (‘data subject’)*” (art. 2(a)). If the data is anonymised or aggregated and an individual cannot be identified from the remaining data, it ceases to be personal data, and the provisions of the DPD no longer apply.

When talking about big data, it is questionable whether the personal/non-personal data distinction remains viable and whether anonymisation and aggregation remain effective in protecting users against tracking and profiling (Monreale, Rinzivillo, Pratesi, Giannotti, & Pedreschi, 2014, pp. 1-2 [PDF]). Even if identifiers, such as names and ID numbers, have been removed, one can use background knowledge and cross-correlation with other databases in order to re-identify student data records (Narayanan & Shmatikov, 2008 [PDF]). Therefore, it could be that when student data is anonymised or aggregated the provisions of the DPD will not apply, but the risk of identifying the student - or more precisely: re-identifying - still remains.

Second, setting consent as the DPD’s main legal guide may be ineffective. A key principle in the DPD is the need to obtain personal unambiguous consent before data can be processed (art. 2(h)). Before big data, parents could roughly gauge the expected uses of their children's personal data and weigh the benefits and the costs at the time they provided their consent. Today, the ability to make extensive, often unexpected, secondary uses of student data makes it simply too complicated for the average parent to make fine□grained choices for every new situation (Kay, Korn, & Oppenheim, 2012 [PDF]). Moreover, in many instances vendors do not offer users the option of choosing which data they agree to share and for which purposes, thus users are forced to accept or deny the service as a whole. Consequently, parents could end up unintentionally excluding their children from services necessary for their education just because they are unable or unwilling to parse out complex data policy statements (Polonetsky & Jerome, 2014).

The Directive does not address the fact that opting-out is hardly a feasible alternative for users in the educational context, since most parents do not have the privilege of changing their children’s schools based on the applicable privacy policy (Zeide, 2016). Therefore, student privacy should not be a binary concept that is either on or off and parents should be given the option of choosing which data they agree to share and for which specific purposes, without having to disengage their children from “big data education”.

Furthermore, the DPD presumes that consent is not freely given in situations where the party requesting consent has power over the individual granting it. Since a school, ultimately, has the power to make decisions that can affect a student's life chances, there is a risk that parents will feel compelled to consent (Kay et al., 2012).

STUDENT DATA PROTECTION UNDER THE GENERAL DATA PROTECTION REGULATION

EU data protection law has undergone a long-awaited, rigorous and comprehensive revision. After long discussions in the various committees, on 16 April 2016, the EU Parliament formally approved the General Data Protection Regulation (GDPR or Regulation) and it is set to go into effect in May 2018 in all EU member states.

The GDPR was adopted by the European Commission "*to strengthen online privacy rights and boost Europe's digital economy*", recognising that "*technological progress and globalisation have profoundly changed the way our data is collected, accessed and used*" (European Commission, 2012).

In general, the GDPR does not forsake the basic principles of data protection established by the DPD, including consent as a ground for lawfulness of processing (art. 6), and the definition for "personal data" which is a key for determining the scope of the Regulation (art. 5). However, the GDPR adopts several innovative approaches to data protection which could improve the level of data protection for data subjects by imposing considerable additional duties on data controllers.

For example, the Regulation places notable emphasis on transparency by requiring data controllers to communicate with data subjects "*in a concise, transparent, intelligible and easily accessible form, using clear and plain language, in particular for any information addressed specifically to a child*" (art. 12). According to Burrell (2016 [PDF]), however, attempts to enforce transparency are challenged by the fact that, for several reasons, algorithms of classification that operate on data, and machine learning algorithms in particular, are irremediably opaque. As Burrell argues, a recipient of the output of the algorithm (the classification decision), rarely has any concrete sense of how or why a particular classification has been arrived at from the inputs. Additionally, the inputs themselves may be entirely unknown or known only partially.

In addition to the transparency requirement, the GDPR introduces the new '*data protection by design and by default*' principle (art. 25) which motivates architects of big data analytics to embed good data protection practices, like anonymisation, pseudonymisation, encryption, and protocols for anonymous communications (European Commission, 2015).

Furthermore, the GDPR obligates data controllers to carry out a risk analysis of the potential impact of the intended data processing if it is "*likely to result in a high risk to the rights and freedoms of natural persons*" (art. 35). If a specific high risk is likely to be presented, the controllers should also carry out data protection impact assessment and periodical compliance reviews.

It is clear that the GDPR is intended to address some of the key data protection issues that have been identified in relation to big data analytics. Although no specific provisions are included with respect to the protection of school student data, the GDPR explicitly refers to providing children (i.e. any person below the age of 18 years), with specific protection of their personal data. In this sense, several provisions are stipulated to set out special conditions for the processing of personal data of children.[13]

5. CONCLUSION

A shift in power from schools to vendors of the EdTech industry is most likely inevitable to some degree. As EU schools become more data-driven we can expect the vendor role in the everyday pedagogic and administrative operations of schools to expand.

The GDPR indicates a possible paradigm shift in the approach of considering privacy and data protection as a new collective interest that would require more public regulation than private enforcement. Once in effect, it may re-establish a different power-balance between data subjects and data users (controllers and processors) thus achieving a significant milestone for increasing the actual level of student privacy protection.

Notwithstanding the need for the EU data protection law to enhance the protection of student privacy by increasing transparency and providing users more consent options, the US experience elucidates that although education shows similarities with other areas, such as social networks or e-commerce, data use in K-12 education also has significant differences.

The mounting public discussion over the outsourcing of student data utilisation goes well beyond traditional privacy and data protection concerns. The expanding role of vendors inside and outside the classroom is taken as a threat to autonomy, liberty, freedom of thought, equality and opportunity.

An adequate regulatory protection of students' rights would focus not only on uses of data outside of school premises, but inside it as well (Pasquale, 2015b). EU policymakers should define the potential risks of outsourcing student data utilisation and need, and establish a new power-balance that will safeguard the full scope of students' rights. For example, and as already pointed out, despite the "*aura of neutrality*", the algorithms that facilitate educational data-driven decision-making may rely on biased data and thus may affect low-income and underserved populations. Drawing from Pasquale's (2015a) analysis of the reputation, search, and finance sectors, one arguable regulatory solution for addressing the risk of big data analytics facilitating discrimination, would be to deploy auditing systems that review the algorithms and the data used to detect biases and test for disparate impact in education.

Another vital regulatory effort would be for policymakers to protect students' "intellectual privacy" from the "surveillance effect". Broadly speaking, policymakers should set boundaries between 'private' and 'public' spaces within digital learning environments, that will safeguard students' freedom of thought and belief, right to read and engage in intellectual exploration, and the confidentiality of communications between participants (Richards, 2015).

Parental concerns in the US stem from the unproven and unpredictable outcomes and potential unintended consequences of student data use, thus they seek to avoid uncertainty by limiting who can access student information in the first place (Zeide, 2015). The demise of inBloom is perhaps the best example of the uncompromising backlash from parents and media and it is indicative of the deep anxiety about the use of student information.

Regulatory rules that focus not only on how student data is transferred from schools to vendors, but also on when and where student data is collected, for what purposes, and by which tools, will build trust around and allay the wide-ranging concerns related to the shift in power from European schools to private entities in the contemporary data-infused educational landscape.

REFERENCES

Alarcon, A., Zeide, E., Rosenblat, A., Wikelius, K. Boyd, D., Gangadharan, S. P., & Yu, C. (2014). Data & civil rights: Education primer. Data & Civil Rights Conference (2014, October 30). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2542268

Article 29 Data Protection Working Party (2012). *Opinion 05/2012 on Cloud Computing*. Retrieved from http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp196_en.pdf

Baker, R., & Siemens, G. (2014). Educational Data Mining and Learning Analytics. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences. Second Edition* (pp. 253-274). Cambridge University Press.

Bennett, E., & Weber, A. S. (2015). Cloud computing in New York State education: Case study of failed technology adoption of a statewide longitudinal database for student data. *QScience Connect 2015*(1). Retrieved from http://dx.doi.org/10.5339/connect.2015.2

Booker, E. (2013, April 26). Education data: Privacy backlash begins [Blog post], *Information Week*. Retrieved from http://www.informationweek.com/education-data-privacy-backlash-begins/d/d-id/1109713?

Burch, P., & Good, A. (2015). More important than the contract is the relationship.□*Phi Delta Kappan, 96*(5), 35-39. doi: 10.1177/0031721715569467

Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society, 3*(1). doi: 10.1177/2053951715622512

Center for Democracy & Technology. (2015). Privacy and the digital student. Retrieved from https://cdt.org/files/2015/06/Student-Privacy-White-Paper-v.-9_1.pdf

Charlton, P., Mavrikis, M., & Katsifli, D. (2013). The potential of learning analytics and big data. *Ariadne, 71*. Retrieved from http://www.ariadne.ac.uk/issue71/charlton-et-al#sthash.wainfh00.dpuf

Chui, M., & Sarakatsannis, J. (2015). Protecting student data in a digital world: Proponents of data-enabled education can learn from other industries that have faced concerns about the risks of using personal information. Retrieved from http://www.mckinsey.com/insights/public_sector/protecting_student_data_in_a_digital_world

Data Quality Campaign (2015). *Student data privacy legislation: What happened in 2015, and what is next?*. Retrieved from http://dataqualitycampaign.org/wpcontent/uploads/2015/09/DQC-Student-Data-Laws-2015-Sept23.pdf

Dennen, V. P. (2015). Technology transience and learner data: Shifting notions of privacy in online learning. *Quarterly Review of Distance Education, 16*(2), 45-59. Retrieved from https://sslt.haifa.ac.il/eds/pdfviewer/,DanaInfo=.aeeuCeEkiolys057Os54+pdfviewer?vid=3&sid=bfabd7b7-9c87-4aac-ad8c-1c30b96d0f32%40sessionmgr4003&hid=4110

Dumon, O. (2014, October 13). Big data and education: the power of transformation. *Research*

Information. Retrieved from http://www.researchinformation.info/news/news_story.php?news_id=1720

Effrem, K. R., & Robbins, J. (2015, June 3). Response to concerns about the Student Privacy Protection Act – S. 1341 (full document). Education Liberty Watch. Retrieved from http://edlibertywatch.org/2015/06/response-to-concerns-about-the-student-privacy-protection-act-s-1341/

European Commission (2015). *The EU data protection reform and big data* [Fact sheet]. Retrieved from http://ec.europa.eu/justice/data-protection/files/data-protection-big-data_factsheet_web_en.pdf

European Commission (2012). Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) COM(2012)11 Final

European Commission. (2012, January 25). Commission proposes a comprehensive reform of data protection rules to increase users' control of their data and to cut costs for businesses. Press Release IP/12/46. Retrieved from http://europa.eu/rapid/press-release_IP-12-46_en.htm?locale=en

European Parliament and Council Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281

Green, C. (2015, September 24). Five things you need to know about the proposed EU General Data Protection Regulation. *Information Age*. Retrieved from http://www.information-age.com/it-management/risk-and-compliance/123460223/five-things-you-need-know-about-proposed-eu-general-data-protection-regulation

Herold, B. (2014, January 22). Americans worried, uninformed about Student data privacy, survey finds [Blog post]. *Education Week*. Retrieved from http://blogs.edweek.org/edweek/DigitalEducation/2014/01/american_worried_uninformed_student_data_privacy.html.

Hoge, A. B. (2015). Vitter's bitter bill joins the "no privacy club”: The Student Privacy Protection Act. Retrieved from https://drive.google.com/file/d/0B6zikOSdV-TAUHlZYkx5V3ZvQWM/view?pli=1

International Educational Data Mining Society (n.d.). Home. Retrieved from http://www.educationaldatamining.org/

K-12 school service provider pledge to safeguard student privacy. (n.d.). Retrieved from http://studentprivacypledge.org/?page_id=45

Kay, D., Korn, N., & Oppenheim, C. (2012). Legal, risk and ethical aspects of analytics in higher education. *Cetis Analytics Series 1*(6). Retrieved from http://publications.cetis.org.uk/wp-content/uploads/2012/11/Legal-Risk-and-Ethical-Aspects-of-Analytics-in-Higher-Education-Vol1-No6.pdf

Kiss, A., & Szőke, G. L. (2014). Evolution or Revolution? Steps Forward to a New Generation of Data Protection Regulation. In S. Gutwirth, R. Leenes, &P. de Hert (Eds.), *Reforming European*

Data Protection Law (pp. 311-331). doi:10.1007/978-94-017-9385-8_13

Krueger, K. R. (2014). 10 steps that protect the privacy of student data. *T.H.E. Journal, 41*(6), 8. Retrieved from http://thejournal.realviewdigital.com/?iid=93887#folio=8

LACE (2015). What are the main barriers in making use of learning analytics?. Retrieved from http://www.laceproject.eu/faqs/barriers-to-learning-analytics/

Leong, B. (2013, November 5). Student privacy pledge reaches 200 signatories! [Blog post]. Retrieved from https://fpf.org/2015/11/13/student-privacy-pledge-reaches-200-signatories/

Locke, C. (2015, March 10), Optimism Returns to Student Data Privacy Debate. edSurge. Retrieved from https://www.edsurge.com/news/2015-03-10-optimism-returns-to-student-data-privacy-debate

Long, P. & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review, 46*(5), 31-40

MacCarthy, M. (2014). Student privacy: Harm and context. *International Review of Information Ethics, 21*, 11-24. Retrieved from http://www.i-r-i-e.net/inhalt/021/IRIE-021-MacCarthy.pdf

Manyika, J., Chui, M., Farrell, D., Van Kuiken, S., Groves, P., & Almasi Doshi, E. (2013). Open data: Unlocking innovation and performance with liquid information. Retrieved from http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information

Marketplace.org (2015). Parents’ attitudes toward education technology. Retrieved from http://www.marketplace.org/sites/default/files/Education%20Technology%20-%20APM%20Marketplace%20Report.pdf

Mayer-Schönberger, V., & Cukier, K. (2014). *Learning with big data: The future of education* [Kindle iPad version]. Retrieved from Amazon.com

Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston: Houghton Mifflin Harcourt

Monreale, A., Rinzivillo, S., Pratesi, F., Giannotti, F., & Pedreschi, D. (2014). Privacy-by-design in big data analytics and social mining. *EPJ Data Science, 3*(1). doi:10.1140/epjds/s13688-014-0010-4

Narayanan, A., & Shmatikov, V. (2008), Robust De-Anonymization of Large Sparse Datasets, *Proceedings of the 2008 IEEE Symposium on Security and Privacy*, 111-125. Retrieved from https://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology, 45*(3), 438-450. doi:10.1111/bjet.12152

Pasquale, F. (2015a). *The black box society: The secret algorithms that control money and information*. Cambridge, MA: Harvard University Press.

Pasquale, F. (2015b, January 15). We're being stigmatized by 'big data' scores we don't even know about. *The Los Angeles Times*. Retrieved from

www.latimes.com/opinion/op-ed/la-oe-0116-pasquale-reputation-repair-digital-history-20150116-story.html.

Privacy Technical Assistance Center (2014). Protecting student privacy while using online educational services: Requirements and best practices. Retrieved from http://ptac.ed.gov/sites/default/files/Student%20Privacy%20and%20Online%20Educational%20Services%20%28February%202014%29.pdf

Polonetsky, J., & Jerome, J. (2014). Student data: Trust, Transparency, and the role of consent. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2628877

Polonetsky, J., & Tene, O. (2014). The ethics of student privacy: Building trust for ed tech. *International Review of Information Ethics, 21*, 25-34.

Protecting Student Privacy Act of 2015, S.1322, 114th Cong. (2015-2016)

Pykett, J. (2012). The pedagogical state: education, citizenship, governing. In J.Pykett (Ed.), *Governing through pedagogy: Re-educating citizens* (pp. 1-4). London: Routledge.

Reyes, J. A. (2015). The skinny on big data in education: Learning analytics simplified. *TechTrends, 59*(2), 75-80.

Richards, N. (2013). *Intellectual privacy: Rethinking civil liberties in the digital age*. New York: Oxford University Press.

Roscorla, T. (2014, July 30). U.S. Senate ponders student data privacy bill. *Government Technology*. Retrieved from http://www.govtech.com/education/Senate-Ponders-Student-Data-Privacy-Bill.html

Rubinstein, I. S. (2013).Big data: The end of privacy or a new beginning?. *International Data Privacy Law, 3*(2), 74-87.

Sabourin, J., Kosturko, L., FitzGerald, C., & McQuiggan, S. (2015). Student privacy and educational data mining: perspectives from industry. *Proceedings of the 8th International Conference on Educational Data Mining*, 164-170. Retrieved from http://www.educationaldatamining.org/EDM2015/proceedings/full164-170.pdf

Safe Kids Act, S.1788, 114th Cong. (2015-2016)

Selwyn, N. (2015). Data entry: Towards the critical study of digital data and education. *Learning, Media and Technology, 40*(1), 64-82.

Singer, N. (2014, September 14). With tech taking over in schools, worries rise. *The New York Times*. Retrieved from http://www.nytimes.com/2014/09/15/technology/with-tech-taking-over-in-schools-worries-rise.html

Society for Learning Analytics Research (SOLAR) (n.d.). About. Retrieved from http://www.solaresearch.org/mission/about/.

Solove, D. (2014, May 7). Big data and our children’s future: On reforming FERPA. Linkedin. Retrieved from

https://www.linkedin.com/pulse/20140507051528-2259773-big-data-and-our-children-s-future-on-reforming-ferpa

Solove, D. J. (2013). Privacy self-management and the consent dilemma. *Harvard Law Review, 126*, 1880-1903. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2171018##

Strauss, V. (2015, March 23). Proposed student data privacy bill does little to protect privacy (update). *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/answer-sheet/wp/2015/03/23/new-student-data-privacy-bill-in-congress-does-little-to-protect-privacy-analysis/

Sullivan, J. V. (2007). How our laws are made. U.S. House of Representatives. Retrieved from https://www.gpo.gov/fdsys/pkg/CDOC-110hdoc49/pdf/CDOC-110hdoc49.pdf

The White House (2014). Big Data: Seizing opportunities, preserving values. Retrieved from https://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf

Strauss, V. (2014, March 6). Why a ‘Student Privacy Bill of Rights’ is desperately needed. *The Washington Post*. Retrieved from http://m.washingtonpost.com/blogs/answer-sheet/wp/2014/03/06/why-a-student-privacy-bill-of-rights-is-desperately-needed/

Student Privacy Protection Act, H.R. 3157, 114th Cong. (2015-2016)

Tsukayama, H. (2015,, January 12). More than 70 companies just signed a pledge to protect student data privacy - with some notable exceptions. *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/the-switch/wp/2015/01/12/more-than-70-companies-just-signed-a-pledge-to-protect-student-data-privacy-with-some-notable-exceptions/

Tudor, J. (2015). Legal implications of using digital technology in public schools: Effects on privacy. *Journal of Law & Education, 44*(4), 287-343.

US Department of Education (2015). *Guidance issued on protecting student privacy while using online educational services: Model terms of service offer direction on guarding student data*. Retrieved from http://www.ed.gov/news/press-releases/guidance-issued-protecting-student-privacy-while-using-online-educational-services

US Department of Education (2013). *Expanding evidence approaches for learning in a digital world*. Retrieved from http://tech.ed.gov/files/2013/02/Expanding-Evidence-Approaches.pdf

Weber, A. S. (2016). The Big Student Big Data Grab. *IJIET International Journal of Information and Education Technology, 6*(1), 65-70. doi:10.7763/ijiet.2016.v6.660.

Williamson, B. (2015). Governing software: Networks, databases and algorithmic power in the digital governance of public education. *Learning, Media and Technology, 40*(1), 83-105, doi:10.1080/17439884.2014.924527

Young, E. (2015). Educational privacy in the online classroom: FERPA, MOOCS, and the big

data conundrum. *Harvard Journal of Law & Technology, 28*(2), 549-592. Retrieved from http://jolt.law.harvard.edu/articles/pdf/v28/28HarvJLTech549.pdf

Zeide, E. (2016, January 21). Student privacy & big data: The status quo, implications & considerations [Video file]. Retrieved from http://www.datasociety.net/events/databite-no-66-elana-zeide/

Zeide, E. (2015, September 18). *Parsing Student Privacy: Creating a Parent-Focused Framework for Conversation* [Blog post]. TAP. Retrieved from http://www.techpolicy.com/Blog/Featured-Blog-Post/Parsing-Student-Privacy.aspx

FOOTNOTES

1. The most widely acknowledged definition of learning analytics is “*the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environment in which it occurs*”. Long, P. & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. Educause Review, *46*(5), 31-40.

2. EDM has been defined as “*an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in*” International Educational Data Mining Society (n.d.). Home. Retrieved from http://www.educationaldatamining.org/

3. The Learning Analytics Community Exchange (LACE) project, for example, funded by the EU, aims at bringing together key European players in the field of LA and EDM to promote the effective use of analytics in a wide range of educational settings including schools, higher education establishments and workplace learning environments.

4. According to the Horizon Report Europe: 2014 Schools Edition (2014) [PDF], European schools have already started routinely using the services and products of vendors to make effective use of varied and real-time student data. For example, the report states that hundreds of primary and secondary schools in Norway, the UK and the Netherlands are using the “itslearning” learning management system, offered by a market leading vendor, to get quick assessments of learning inside and outside the classroom.

5. Knewton, for example, one of the most prominent companies in the field, uses big data to develop adaptive learning systems and data analytics for students, teachers, school district and publishers. The data analytics are intended to map students' weaknesses and strong points along time, to enable the teacher to personalise the learning process and the content.

6. In 2014, Google admitted that it mines student data from its Google Apps for Education for targeted advertisement purposes. See Gould, J. (2014, January 31). Google admits data mining student emails in its free education apps. SafeGov.org. Retrieved from http://safegov.org/2014/1/31/google-admits-data-mining-student-emails-in-its-free-education-apps

7. A contractor providing outsourced services to a school is treated as a “school official” if it is (1) performing services for which the school would otherwise use employees; (2) is under the direct control of the school with respect to the use and maintenance of student data; and (3) agrees to abide by FERPA regulations governing use and redisclosure of student data.

8. Under the statute, "personal information" is defined as individually identifiable information including a student or parent's first and last name, a physical address, a telephone number, or a social security number.

9. A bill is the form used for most legislation, whether permanent or temporary, general or special, public or private. A bill does not become law until it is passed by the legislature and, in most cases, approved by the executive (in the US, the President). See Sullivan, J. V. (2007). How our laws are made. U.S. House of Representatives. Retrieved from https://www.gpo.gov/fdsys/pkg/CDOC-110hdoc49/pdf/CDOC-110hdoc49.pdf

10. The Safe Kids Act, an additional federal bill addressing student data protection, was introduced later on in July 2015.

11. In a recent survey it was found that although there is a feel good factor about the growing use of technology in education, 79% of parents reported they are at least somewhat, very or extremely concerned about the security and privacy of their child’s data (See Marketplace.org (2015). Parents’ attitudes toward education technology. Retrieved from http://www.marketplace.org/sites/default/files/Education%20Technology%20-%20APM%20Marketplace%20Report.pdf

12. The DPD applies to all individuals whose personal data is processed in a member state of the EU. Any use of data constitutes processing under the DPD, and anything that is done to the data is considered to be processing the data, ranging from its creation or collection, to its eventual destruction.

13. Article 6 of the draft GDPR indicates that special consideration is to be given to the fact that the interests of children might be at stake. Article 12 establishes that information must be adapted to the data subjects, especially if they are children. Article 40, on ‘Codes of conduct’, states that data controllers and data processors should be encourage to draw up codes of conduct to the proper application of the Regulation taking into account the specific features of the various data processing sectors, such as with regards to “*the information provided to, and the protection of, children, and the manner in which the consent of the holders of parental responsibility over children is to be obtained*”.