

Uncovering Privacy and Security Challenges In K-12 Schools

Jake Chanenson
jchanen1@uchicago.edu
University of Chicago
USA

Jason Chee
jchee1@uchicago.edu
University of Chicago
USA

Brandon Sloane
brandon.sloane@nyu.edu
New York University
USA

Navaneeth Rajan
navrajan05@gmail.com
Princeton Day School
USA

Amy Morrill
amymorrill@uchicago.edu
University of Chicago
USA

Danny Yuxing Huang
dhuang@nyu.edu
New York University
USA

Marshini Chetty
marshini@uchicago.edu
University of Chicago
USA

ABSTRACT

Increased use of technology in schools raises new privacy and security challenges for K-12 students—and harms such as commercialization of student data, exposure of student data in security breaches, and expanded tracking of students—but the extent of these challenges is unclear. In this paper, first, we interviewed 18 school officials and IT personnel to understand what educational technologies districts use and how they manage student privacy and security around these technologies. Second, to determine if these educational technologies are frequently endorsed across United States (US) public schools, we compiled a list of linked educational technology websites scraped from 15,573 K-12 public school/district domains and analyzed them for privacy risks. Our findings suggest that administrators lack resources to properly assess privacy and security issues around educational technologies even though they do pose potential privacy issues. Based on these findings, we make recommendations for policymakers, educators, and the CHI research community.

CCS CONCEPTS

- **Security and privacy** → **Social aspects of security and privacy; Domain-specific security and privacy architectures;** •
- **Human-centered computing** → **Empirical studies in HCI.**

KEYWORDS

student data privacy, EdTech, K-12, educational technologies

ACM Reference Format:

Jake Chanenson, Brandon Sloane, Amy Morrill, Jason Chee, Navaneeth Rajan, Danny Yuxing Huang, and Marshini Chetty. 2023. Uncovering Privacy and Security Challenges In K-12 Schools. In *Proceedings of the 2023 CHI '23, April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 28 pages. <https://doi.org/10.1145/3544548.3580777>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI '23, April 23–28, 2023, Hamburg, Germany

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-9421-5/23/04...\$15.00

<https://doi.org/10.1145/3544548.3580777>

Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany. ACM, New York, NY, USA, 28 pages. <https://doi.org/10.1145/3544548.3580777>

1 INTRODUCTION

Schools face privacy and security challenges with educational technologies, often referred to as “EdTech.” Alongside the learning benefits, educational technologies often include expanded data collection features such as monitoring for student engagement or tracking academic performance [71, 91]. This data collection however, has consequences on schoolchildren whose data may be harvested for commercial purposes, may encounter data leaks, or who may face future employment discrimination based on an accrued record [43, 78, 80, 81, 87]. Furthermore, educational technologies have not faced rigorous scrutiny or expanded regulation from policymakers, even as their scope and use has expanded over the years [10, 93]. Finally, schools themselves may similarly incorporate tracking capabilities to gather data on students which could again, lead to data leaks or oversharing of information [10, 47, 79, 99]. For these reasons, schools must be more intentional with and take care to ensure that the technologies they employ are safeguarding the privacy and security of their students.

Moreover, cybersecurity and privacy incidents in K-12 school districts are increasing in occurrence [8, 9, 14, 27, 32, 42, 46, 57, 62, 62, 70, 70, 76]. Many incidents go undisclosed due to a lack of mandatory federal disclosure laws [42] even when these incidents directly concern student data [6, 15, 20, 53, 66, 82, 83, 98]. Yet, the US Government Accountability Office (GAO) found that there were 99 student data breaches, affecting hundreds of school districts, between July 1, 2016 and May 5, 2020 [32]. For instance, a data breach in January 2022 at educational technology vendor Illuminate Education compromised the personal information of 820,000 former and current students of the New York public school district [83, 98]. In addition to personal identifying information (PII) such as names and birthdays, Illuminate Education also stored highly sensitive information including student’s free-lunch and special-education status. These incidents suggest that further research into EdTech privacy and security implications for K-12 is needed. Prior work has already examined privacy and security vulnerabilities in a range of

EdTech products around the world [5, 37, 88, 94]. However, there are still no quantitative measures of what EdTech products are used across K-12 public schools in the United States (US) or what kinds of privacy issues—such as online tracking [50]—they may pose at scale.

Additionally, school district officials and Information Technology (IT) personnel act as gatekeepers for bringing in new technologies and setting up school/district websites to support students. Yet it is unclear what privacy and security considerations they make when procuring educational technologies and deciding which technologies to link to on school/district websites. Inside the school setting, existing literature has investigated the privacy and security awareness of elementary school teachers and children [34, 38, 43–45, 58, 96, 97, 101]. Other work has studied the criteria teachers and district officials employ before purchasing EdTech and the degree of IT personnel's involvement in EdTech procurement [55, 62], but the privacy and security considerations made by these stakeholders are not well known.

To address these gaps in the literature, we posed four research questions:

- **RQ1:** What are the privacy and security issues around EdTech products in K-12 school districts as perceived by key decision makers such as school officials and IT personnel?
- **RQ2:** How do school officials and IT personnel choose EdTech products for their districts, and what privacy and security considerations do they make for these products?
- **RQ3:** What are the main EdTech products that K-12 school districts likely use or endorse?
- **RQ4:** What potential privacy risks can we identify from K-12 school/district websites and third-party EdTech websites?

To answer these questions, we conducted a two-part mixed methods study. First, we interviewed 18 school district officials and IT personnel to see what EdTech-induced privacy and security issues schools face, how they acquire EdTech, what EdTech their districts use, and what other EdTech factors they consider that may affect students' privacy and security with technologies used in schools. Second, given that school districts typically list approved software on their websites [38], we scraped 15,573 K-12 public school/district domains to ascertain which EdTech products are the most commonly endorsed by schools beyond the 11 districts our interviewees hailed from. Using these results, we explored potential privacy issues on school websites, EdTech websites, and the login pages leveraged by these sites to provide access to users.

We have four main findings: 1) schools experience technology-related privacy and security incidents but lack the resources and knowledge to handle these challenges; 2) privacy and security training and awareness about EdTech potential issues for students is limited in schools; 3) schools have clear pipelines to bring EdTech into the classroom but these pipelines often do not adequately assess potential privacy and security issues for students; and 4) schools link to many products that collect student data but are not typically thought of as EdTech and school/district websites link to a non-trivial amount of domains with potential privacy issues such as session recorders. Based on these findings, we recommend that further work is needed to investigate other stakeholder perspectives

in this space and provide additional insight into what data EdTech products collect and share on students. At a minimum, schools need improved training and resources for handling privacy and security challenges that come with EdTech and student data, and K-12 regulation needs to better account for the expanding EdTech space. This will require a coordinated effort from policymakers, the CHI research community, and educators.

Our contributions are as follows:

- Evidence of school district officials' and IT personnel lack of capabilities for managing privacy and security challenges with EdTech from 11 US public school districts, adding to a growing body of work on privacy and security challenges for school children and teachers such as [38, 43, 45, 48].
- New evidence of how school officials and IT personnel bring EdTech into classrooms with points for potential improvement to privacy and security considerations.
- A categorization of the top 300 educational technology domains that are linked from, thus implicitly endorsed by 15,573 public school/district domains.
- Evidence of potential privacy issues for students logging into 15,573 public school/district domains and educational technologies from these domains such as the use of session recorders and the Meta pixel.
- Recommendations to better help schools overcome privacy and security risks in childrens' use of educational technologies.

Next, we describe the related work and background, our methods, findings, and discussion and conclusions.

2 BACKGROUND AND RELATED WORK

2.1 Technologies Used in Schools

K-12 schools are heavily reliant on technology to manage and operate their programs, most of which are considered EdTech and facilitate *“learning and improving performance by creating, using, and managing appropriate technological processes and resources”* [33]. Recently, the EdTech market was valued at over 85 Billion USD in 2021 [68] with EdTech usage increasing over the last few decades particularly during the COVID-19 when these products supplemented or replaced in-person learning [21, 63]. Specifically, before the COVID-19 pandemic, EdTech was primarily used for educational games, communication, collaboration, formative assessment, student feedback, content creation, and delivery of instructional content [37]. However, the pandemic caused a shift in students' needs, including an increased need for video conferencing software, to aid with remote learning [62]. Although there is some anecdotal evidence of the range of EdTech products [36, 37], there is no quantitative evidence of the range and types of EdTech products currently used in US public schools. Our work addresses this gap to better inform policymakers on EdTech regulations and privacy and security researchers about some of the most popularly endorsed EdTech vendors and categories in US public schools that could create challenges for students' data.

Many EdTech products have increased their surveillance capabilities, collecting and analyzing student data for reasons unrelated to learning outcomes [37, 95]. This increases the potential for misuse of collected data. Without an understanding of how schools acquire

EdTech products, it is unclear how protected students are from privacy and security harms, which our work seeks to address. Amidst these unknowns, existing federal and state privacy laws do regulate how EdTech products collect and use student data; Family Education Rights and Privacy Act (FERPA) directly regulates student data in an education context by dictating when and how students' data can be disclosed [93] and Children's Online Privacy Protection Act (COPPA) gives parents control over the use of data from their kids younger than 13. However, FERPA has not been updated to account for the expanding scope of EdTech technologies with student data and COPPA requires active parent involvement—making it hard to enforce strong protections over student data collection and use [23, 24]. There are also at least 40 states with privacy laws that directly affect student data in K-12 schools [88] but these laws vary widely in content and do not protect data from all EdTech uses or potential misuses necessarily [29, 60]. Our work seeks to provide evidence to inform improved EdTech regulations.

Aside from potential privacy and security challenges for schools with EdTech usage, it is also unclear whether schools set up links from their school district/school websites to EdTech websites in privacy-preserving ways. For instance, cookies and trackers can provide persistent session information which can extend beyond a single session [25, 50, 77], meaning that if a student arrived on that site, information about them could be recorded. Information tracking is sometimes further extended by the Meta Pixel [52] by linking persistent session information with social media account information, or across different sites that use the Facebook Javascript, which again could be problematic for tracking students visiting these school/district websites [3]. Mouse and keyboard behavior event listeners and session recorders, on the other hand, capture detailed information about what a user does during a particular session [86]; in particular, if a user enters their email address into a login form with third-party mouse/keyboard listeners or session recorders, the third-party could potentially exfiltrate the email address—or any other sensitive information—even without the user submitting the login [75]. A website that leverages persistent session mechanisms will have the ability to gather information about its users over time and build an extensive profile. Similarly, a website that leverages listener and recording mechanisms will have the ability to gather detailed information about each user session [1, 2]. Both of these categories of information gathering mechanisms present privacy risks and opportunities for misuse, particularly when K-12 children visit these sites. Our study helps to determine if any of these privacy issues occur based on publicly available information on school/district websites.

We note that the web is not the only place where EdTech could infringe upon students' privacy. For example, prior work has identified proctoring software, standalone applications that run on students' computers rather than in the browser, that could monitor students' screen activities and network traffic, sometimes continuing monitoring after exams are completed [12]. We are aware of the invasive nature of such EdTech, although it is beyond our scope of analysis and we focus on web-based EdTech in this paper.

2.2 Technology-Related Privacy and Security Issues in K-12 Schools

Given that EdTech often collects data on students, researchers have sought to understand the privacy implications of these technologies. Many of these studies have focused on the privacy implications of using educational technologies in higher educational institutions, not the K-12 context [16, 17, 35, 39, 61, 65, 67, 72, 85, 100]. There is also a robust body of literature from CHI researchers that examines children and privacy and security more broadly [22, 34, 38, 44, 48, 58, 96, 97, 101] and in the context of schools, the CHI and related communities have studied children, parents, and teachers and how each of these stakeholders perceive privacy and security issues. For example, Tazi et al. polled educators and parents of students about their perceptions of distance learning's privacy and security challenges [89]. To complement this broad swath of work, we interviewed stakeholders that are understudied in the school context and are critical players in safeguarding children's privacy and security at school; that is key decision makers at the district level, school officials and IT personnel, to understand privacy and security issues in the classroom that arise from EdTech. There are also several studies focused specifically on technologies used in K-12 schools, the subject of our research. Researchers have investigated decision-making processes around EdTech selection either school district-wide [55, 62, 69] or for individual teachers [11, 31], but these studies do not focus on how schools consider privacy and security issues for students when selecting EdTech. For example, these studies investigated how US educators, including teachers, principals, and district officials, choose educational technologies and found that the decisions are usually based on how well a product supports student learning and ease of use [55] and that IT professionals typically take responsibility for these tasks [62].

At least one academic and several non-academic studies have examined the privacy issues around EdTech to understand how student data is utilized and protected [10]. The Electronic Frontier Foundation (EFF) and Common Sense Media both investigated the privacy policies of roughly 150 educational technologies and found that many EdTech lack disclosures about key data practices such as student data retention, security, and de-identification of PII [5, 36, 37]. These studies also found EdTech often collects sensitive student information such as sexual orientation and shares data with third parties. Similarly, Human Rights Watch examined the privacy and security vulnerabilities of 164 EdTech products endorsed by 49 governments during the COVID-19 pandemic [94]. Using mixed methods such as technical analysis of technologies and interviews with students, parents, and teachers, the researchers found that 89% of the technologies they studied violated students' privacy through excessive surveillance and sharing data with third parties for advertising purposes. Our study builds on these works to examine how school district officials and IT personnel consider privacy and security for students when acquiring EdTech for their schools in the US.

3 METHOD

To answer our research questions, we conducted a two-part study. In part one, we conducted semi-structured interviews to understand

user perspectives around K-12 educational technologies and security and privacy issues they believe exist (**RQ1**), decision-making processes that drive EdTech selection, and what EdTech school districts use (**RQ2**). To complement our human-centered investigation, in part two we scraped known web domains of K-12 school districts in the US to identify the most commonly linked third-party EdTech domains from these websites, building on the initial list gathered from interviewees (**RQ3**). We then examined a subsection of the top third-parties identified as EdTech and the first-party K-12 district websites in our scraping, for potential privacy issues to see what risks actually exist (**RQ4**).

3.1 Interviews With School District Officials and IT Personnel (RQ1 and RQ2)

To understand school districts' security and privacy challenges, we interviewed 18 district officials and IT personnel between August 2021 and December 2021. Our study was approved by our Institutional Review Board.

3.1.1 Interview Guide. We created our interview guide based on our literature review, research questions (**RQ1 and RQ2**), and our knowledge of US public schools. For instance, US public schools are grouped into school districts, with each school in a district being under the same administrative umbrella. These school districts are typically governed by a set of elected local community members, called a school board [59], while the day-to-day operations of a school district are run by the superintendent—who is the chief administrator of a school district—and their staff. There is no set number of schools that make up a school district, but school districts typically have at least one upper, middle, and elementary school to educate the children within the district from kindergarten or 1st grade through the end of their public school education in 12th grade [59]. Based on this information, we focused our study on school district officials since they run multiple schools and IT personnel since they are likely involved in selecting EdTech and maintaining school privacy and security. We then created questions for these stakeholders and to gather their input on other stakeholders we identified in EdTech and privacy and security, namely teacher, parents, and students themselves.

We iterated on and revised our guide through team discussion and feedback. We also conducted four pilot interviews with education personnel to further refine our guide: a teacher, a district's IT director, a school IT director, and a K-12 security expert. Note, we do not include the pilot interviews in our final data set. The final version of the interview guide, (see Appendix A), consisted of three sections:

- First, we asked for background information including demographic information about the school district, what hardware and software the school district uses, the role of the participant within the district (e.g., superintendent, IT director, etc.), and how the district selects software to support educating its students in addition to any privacy and security considerations made for student data.
- Next, we asked about perspectives on teachers' use of educational technologies within the district, school policies regarding how teachers can introduce software in their classrooms,

what training and oversight teachers receive in privacy and security, and any privacy or security incidents that may have occurred with student data.

- Lastly, we asked participants about the extent of parents' and students' awareness and control over educational technologies and student data. For instance, we asked about how parental consent is obtained for data collection performed by educational technologies and if the data collected on students was viewable by parents or students.

3.1.2 Recruitment and Data Collection. To recruit our target population, we leveraged contacts at partner organizations that directly interface with school districts and their stakeholders to generate a list of contact leads. We also attended relevant EdTech conferences to solicit volunteers. Finally, we used snowball sampling to find additional participants.

We used a screener survey, Appendix B, to check if potential participants were school district officials or IT personnel. Participants passing the check were directed to a Qualtrics survey, Appendix C, to fill out a consent form, demographic information—such as age, gender, and job title—and preference for a 30 or 60 minute Zoom interview. We conducted 23 interviews in total but excluded five interviews because these participants disclosed that they were (1) not a district official or IT personnel or (2) they were not working in a public school; *i.e.*, contrary to their affirmations in the screener survey. Despite the exclusions, we only stopped interviewing once we reached data saturation [74], that is we did not hear substantially new or novel data points, bringing our final sample size to 18¹. All but two interviews were one-on-one interviews—those two were group interviews. Participants were compensated with a \$20 Amazon gift card.

3.1.3 Data Analysis. We performed a thematic analysis on the transcribed interviews [73]. Two researchers developed a codebook through several iterations where each researcher independently read the same transcript, created potential codes, and then compared their work for a subset of transcripts. The research team then met multiple times to discuss the codes before finalizing the codebook. Table 1, our final codebook, had a total of 27 codes which consisted of 6 parent codes, each with 3-6 child codes. Examples of parent codes include: "Tech Acquisition" and "Student Data" and child code examples include: "Tech-Contracts" and "Data-Hygiene."

Each transcript was coded twice, first by a primary coder and then by a secondary coder. For instance, "Tech Acquisition" was applied to all excerpts in transcripts pertaining to discussions related to how EdTech is procured or brought into schools, with subcodes used to indicate the specifics of each discussion. *E.g.*, if it pertained to contracts with EdTech, "Tech-contract" was also applied. After coding all the transcripts, the research team extracted interview excerpts from each associated structural code. Two coders then performed a round of axial coding on the excerpts and wrote thematic summaries for each structural code. The research team met regularly to discuss these thematic summaries, resolve any discrepancies, and finalize the emergent themes discussed in this paper such as technology acquisition and privacy and security awareness.

¹Note, this is larger than the median sample size for user studies at CHI which is 12 [13]

Code	Code Description
Demographics Role School Facts Hardware/Devices Digital Divide COVID-Tech IT-Role	What role is this person in? IT/Admin/School board School size, district composition, location, funding What types of devices does the school have? Do the students take them home? Mentions how school has had to compensate for the digital divide in their district Discussion of how COVID changed the school's tech policy Discussion of the role that IT plays in the EdTech space
Technology Acquisition Teacher-Tech District-Tech School-Tech Software-Info Tech-Contracts	Discussion of how teachers can (or cannot) bring technology into their classroom Discussion of how the school district purchases technology Discussion of how individual schools purchase technology, separately from district Discussion of types of software and if they are paid for or not Discussion of EdTech contracts
Student Data Data-Hygiene Data-Permanence Data-Regulation Data-Parent Data-Student	Discussions of the school's current policy related to data hygiene Discussions of what happens to student data after graduation, after vendor contract termination, etc Discussions of data regulation/privacy laws Discussions of parental control over student data Discussions of student control over student data
Privacy and Security Awareness Educator-Awareness Parent-Awareness Student-Awareness Privacy and Security Incidents Incident Protocol	Discussion of how aware teachers, IT people, and admin are about privacy/security issues with student data Discussion of parental awareness of EdTech both software and hardware Discussion of how aware students are about privacy/security issues with student data Discussion of privacy and security incidents with student data in the district Discussion of the district's protocol for dealing with privacy and security incidents
Technology Oversight Teacher-Oversight Teacher-Guideline Teacher-Training	Discussion of any oversight or compliance the district enforces for teachers' use of tech Discussion of any guidelines that the districts provide in introducing technology to the classroom Discussion of any training that teachers receive about cyber security or student privacy
Miscellaneous Interesting Quotes Data-Collection-Mental-Model	Something interesting that does not fit into other codes Relevant quotes The mental model of school officials about why and how data is being collected

Table 1: Codebook used to code interview transcripts before extracting the coded excerpts for thematic analysis. Primary structural codes are bolded. For instance, “Privacy and Security Awareness” was applied to any excerpts covering discussions of various stakeholder awareness, privacy and security incidents, and mentions of incident protocols.

Since we performed coding as input to thematic analysis, we did not calculate the inter-coder reliability score as outlined in [51].

3.2 Participants

Table 2 summarizes our participant's data. We had 18 participants: six identified as women, 12 as men, and none as non-binary. All of our participants hailed from 11 different school districts located in nine states spanning the Northeast, Mid-Atlantic, Midwest, Northwest, and Southern regions of the United States. 7/18 of our participants had a role as a district official, including: principal, director of secondary teaching and learning, associate superintendent, superintendent, and school board member. Their responsibilities varied both by position and school district. For example, in a larger school district such as SD2, one of the associate superintendents we talked with, P4, was responsible for running the business side of the district, including budget operations and facilities management. At the same time, the other associate superintendent, P5, was responsible for running the teaching side of the district, including managing student assessments and the technology department. The seven district officials came from four unique school districts.

11/18 participants had a role in information technology (IT), including director of IT, head of technology, chief information officer, network manager, or technology resource person. Our participants who were in leadership roles in IT described their responsibilities as overseeing the implementation and security of any technological

component in school, including: end-user devices, servers, access points, websites, and so forth. Our participants in non-leadership IT roles reported that their responsibilities consist of keeping pre-existing technology working. All but one of our IT personnel were from unique school districts.

We were only able to interview both a district official and an IT person in three of the school districts. 6/11 school districts belonged to a privacy group or general consortium such as the Student Data Privacy Consortium [4] or the Illinois Education Technology Leadership group [28] that often helped them understand or assess privacy and security issues with technologies.

3.3 Web Scraping of US Public School/District Websites (RQ3)

Having reached data saturation point in our interviews, and with an initial list of educational technologies being used in school districts from our interviewees, we switched to a technical analysis to augment our findings. We wanted to see whether the initial list of EdTech was also frequently endorsed by public schools across the US and how our participant's experiences and expectations of EdTech's privacy and security risks for students inter-played with EdTech's *actual* online privacy threat landscape. Towards these goals, we scraped all public school districts within the US to see what educational technologies they link on these sites to provide

Participant Code	School District	No of Schools In District	Role	State	Data Protection Legislation	Privacy Group Member	Gender	Age
P1	SD 1	20+	IT	MA	Yes	Yes	Male	35-44
P2	SD 2	1-10	IT	IN	No	N/A	Male	35-44
P3	SD 2	1-10	District Official	IN	No	N/A	Male	35-44
P4	SD 2	1-10	District Official	IN	No	N/A	Male	35-44
P5	SD 2	1-10	District Official	IN	No	N/A	Female	45-54
P6	SD 3	1-10	IT	CA	Yes	No	Male	35-44
P7	SD 3	1-10	District Official	CA	Yes	No	Female	55+
P8	SD 4	1-10	IT	WI	No	No	Male	45-54
P9	SD 5	1-10	IT	IL	Yes	Yes	Female	45-54
P10	SD 6	11-20	District Official	CT	Yes	No	Female	55+
P11	SD 7*	20+	IT	OR	Yes	Yes	Male	35-44
P12	SD 8	1-10	IT	TX	Yes	N/A	Female	35-44
P13	SD 9	1-10	IT	PA	No	No	Female	55+
P14	SD 9	1-10	District Official	PA	No	No	Male	35-44
P15	SD 9	1-10	District Official	PA	No	No	Male	45-54
P16	SD 5	1-10	IT	IL	Yes	Yes	Male	35-44
P17	SD 10	1-10	IT	MA	Yes	Yes	Male	45-54
P18	SD 11	1-10	IT	MA	Yes	Yes	Male	45-54

Table 2: Participant Role and School District (SD) Breakdown. *P11 works for an education service district that supports all of the public school districts in his assigned region. Meaning, SD7 is an amalgam of multiple school districts. The “Data Protection Legislation” column indicates if there is legislation that protects student data in a given participant’s state. The “Privacy Group Member” column indicates if the participant mentioned that their school or district was a member of an education-focused privacy group; N/A means that the participant did not mention a privacy group during their interview.

evidence of the most commonly endorsed technologies across the US.

3.3.1 Scraping the Seed URLs of School/District Websites. We generated a target school list by querying the National Center for Education Statistics (NCES) database [54] for each state to generate a list of URLs for 61,235 K-12 unique public schools in the US across a reported 13,244 school districts. For each school, the NCES database lists the URL for the school’s or its school district’s website URL; there is no additional data to distinguish whether a URL is for a school or a district. Because of this ambiguity, we generally refer to a school or district website as simply “school/district website.”

In total, there were 39,208 distinct URLs across the NCES database. For each URL, we extracted the *registered domain*; for instance, if a URL is <http://classroom.springisd.org> or <http://homepage.springisd.org>, then the registered domain both is springisd.org. We discovered 15,376 registered domains from amongst the 39,208 distinct URLs.

In addition to the NCES database, we also leveraged information from [84] to further expand our inventory of school/district websites. K12 SIX is a “nonprofit threat intelligence and best practices sharing community” and maintains a “Cyber Incident Map which tracks publicly disclosed school cyber incidents from 2016 to present.” Schools and their respective websites that showed up on the K12 SIX Cyber Incident Map were almost certainly active and therefore good targets to include in our inventory. Given the currency of their data, adding school/district websites from the K12 SIX incident map could only enhance our coverage of active

domains. The K12 SIX Cyber Incident Map generated a list of 926 URLs which when filtered yielded 886 registered domains.

As a final step in the domain generation process, we performed a set union operation on the 15,376 NCES and 886 K12 SIX registered domain which gave us 15,573 registered domains in total. These domains could be associated with individual districts (e.g., springisd.org for the Spring Independent School District in Texas) or individual schools (e.g., ivyhawnschool.org for a school in Florida). We shall refer to these URLs and registered domains as the *seed URLs* and *seed domains* respectively.

We ran a Python web scraper that visited each of the seed URLs. For each URL, the scraper downloaded the HTML and extracted all href-based links. If the link shared the same registered domain as the seed URL, the scraper would consider it a first-party link and recursively visit the link. Otherwise, the link would be considered a third-party link; the scraper would visit it once and not follow any subsequent links.

We ran the scraper for a week. From the seed URLs, the scraper discovered a total of 16,693,933 first-party URLs (i.e., which share the same registered domain as the seed URLs) and 3,530,623 third-party URLs. These third-party URLs are associated with 176,115 unique registered third-party domains. For example, among the top domains are google.com (linked from 10,908 seed domains) and facebook.com (linked from 9,519 seed domains); scholastic.com, ranked the 10th, is linked from 3,184 seed domains.

This inventory of school/district websites served as our basis for collecting information about third-party educational technologies listed by US public schools. Although being listed on a

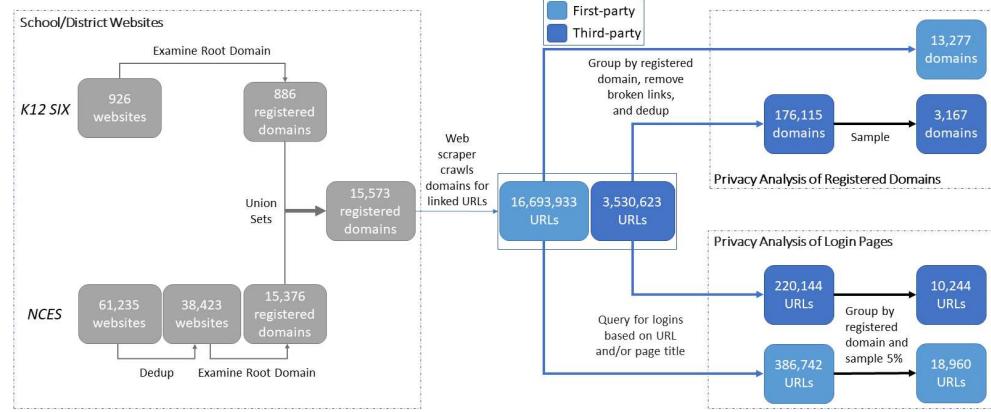


Figure 1: A diagram visually representing how we scraped and processed the seed URLs of school/district websites for analysis. See section 3.3 and section 3.4 for full details.

school/district website does not constitute “use” per se, we take this to mean an endorsement by the school district since prior works have shown that school districts typically list district-approved educational technologies on their sites [38]. We also defined endorsement in this way as we argue that if a link appears on a school/district’s website in some capacity, they effectively drive traffic to that external site. We discuss limitations of this definition and the inventory of educational technologies we gathered in section 8. The process described in this section can be seen in the “School/District Websites” box of figure 1.

3.3.2 Identifying EdTech Sites. Once we derived the list of all third-party links scraped from all the school districts in our master list, we filtered this down to the 300 most frequently appearing third-party domains on these websites. Next, we excluded seven domains for redirecting to other domains within the top 300. *e.g.*, youtube.be redirects to youtube.com. To keep our analysis at 300 domains, we expanded our list to the top 307 domains. As outlined below, we then qualitatively coded the 300 most frequently linked third-party domains found in the scrape to determine which of these domains were EdTech.

For the purposes of our analysis, we defined EdTech as: any website, app, or software that has collected student data—personal information, academic performance, etc.—and markets an educational solution primarily to K-12 schools and their students. For example, although Facebook may gather data on K-12 students above the age of 13, it would not be considered EdTech because Facebook is not primarily marketed as a tool for K-12. However, Google Classroom collects data on students and is specifically marketed as an educational solution for K-12 schools, so it would fit into the EdTech category. To determine if a product qualified as EdTech, we examined its privacy policy to determine if it collected data on students and its “Solutions” or “About Us” page to understand its primary market.

We further categorized domains using EdTech categories included in Common Sense’s 2019 State of EdTech Privacy Report [37]: communication, content creation, delivery of instructional content,

formative assessment, and educational games. During a preliminary exploration of the data, we decided to include nine additional categories to better represent the domains in our data: administrative management, behavior management, college/career, financial, learning management system (LMS), online class textbooks, school merchandise management, study aides, and video communication. Table 3 lists all educational technology categories and definitions.

We discovered these EdTech domains because each of them was linked from one or more school/district websites. Because of this linkage, we assume that the schools or districts endorsed the EdTech service and *potentially* used it. To provide further evidence for usage, we checked whether the linked EdTech webpage included a login page; presumably, a login page would allow users, such as students and teachers, to subsequently interact with private content, thus implying usage—rather than mere endorsement—by schools or districts.

After finalizing these 14 categories, a primary and secondary coder used qualitative data coding to independently code each of the 300 most frequently linked domains as “EdTech” or “non-EdTech” [73]. For all domains coded as EdTech, these coders then applied the 14 additional subcategories to indicate the EdTech product’s primary purpose. The two coders then met to discuss and resolve disagreements. After resolving as many disagreements as possible, the Cohen’s kappa for this qualitative coding was 0.93. In case of disagreements, we report the results of the primary coder.

3.4 Discovery of Potential Privacy Leaks In Listed EdTech Sites (RQ4)

We consider privacy leaks as the transmission of potentially sensitive information from users to third-parties unknown to users, such as advertising and tracking companies [56]. Such sensitive information can include persistent identifiers that are both anonymous (*e.g.*, user IDs) and personally identifiable (*e.g.*, email address), along with any user behavioral data, such as mouse clicks, which third-parties could take advantage for the purpose of cross-site tracking, profiling, and advertising [1, 75].

Category	Definition	Example	Counts	Percentage
Delivery of Instructional Content	Software that allows teachers to deliver instructional content	edmodo	36	32.73%
Formative Assessments	Software that allows students to register for or take assessments	Renaissance	27	24.55%
Administrative Management	Platforms to register students, classes, etc.	PowerSchool	26	23.64%
Educational Games	Software that supports educational games for students	Kahoot	21	19.09%
Communication	Software to facilitate communication between teachers and others	Remind	19	17.27%
Study Aides	Services and resources which help students study	Quizlet	13	11.82%
Online Textbooks	Platforms that provide digital versions of textbooks	ProQuest	12	10.91%
College/career	Services for career development and/or college	Naviance	8	7.27%
Financial	Services to manage school payments	RevTrak	7	6.36%
Behavior Management	Software that monitors or manages student behavior	ClassDojo	7	6.36%
Learning Management Systems	Platforms that help organize courses, homework, etc.	Canvas	6	5.45%
Content Creation	Software that allows students to create content	Prezi	6	5.45%
Video Conferencing	Software that enables video communication	Zoom	2	1.82%
School Merchandise Management	Services that facilitate student purchase of merchandise	Jostens	2	1.82%

Table 3: EdTech categories, their definition, and their prevalence in the EdTech vendors captured in the Top 300 domains (n=110) linked from 15,573 School/District Websites. Note that EdTech categories are not mutually exclusive, and much of the software has been attributed to multiple categories. As such, the counts of each category exceed the total number of software sampled. The “Percentage” column represents the percentage of all EdTech (count/110).

3.4.1 Measuring Privacy Leaks with Blacklight. To quantify privacy leaks, we measure a number of aspects of any first-party (*i.e.*, school/district) and third-party webpages (*e.g.*, EdTechs or other vendors linked from school/district websites): (i) third-party cookies and ad trackers, typically for tracking user behaviors across different sites [25]; (ii) mouse and keyboard behavior event listeners, often used to record how users moved the mouse cursor on a page and what users typed [1, 75]; (iii) Meta Pixel events, usually for linking user activities on and off Facebook and also across different non-Facebook sites [52]; and (iv) session recorders, often used to record/reply user interactions on a webpage [1] (see Section 2). To conduct this measurement, we used Puppeteer [26], a headless browser, and a series of custom scripted tasks developed by Markup, which have been packaged into a suite known as Blacklight [49]. A headless browser is one that does not have a graphical user interface being instead controlled through a command-line interface and in our case, script. These scripts performed automated actions and, paired with our headless browser, extracted tracking information from the browsing session. We utilized a headless browser instead of scripted cURL commands because many of these tracking technologies—such as session recorders—use Javascript, a client-side language that is not rendered by more primitive tools like cURL which only retrieves text.

When we perform our cookie and tracker counts, we generate only unique counts by registered domain. For example, we count two or more cookies from Google Analytics as only a single instance of cookie usage per site. This generally constitutes only a single instance of a contextual integrity transaction (*i.e.*, same source and recipient) [56]. We calculate the median number observed across all sites with particular registered domains. Our goal here is to try and enumerate the number of different parties which might be obtaining information on visiting users. Representing multiple cookies from the same provider allows us to generate a more accurate picture of the number of parties involved. Similarly, median results across sites within a registered domain allows us to generate a more accurate picture of what a particular domain is doing across all of its traffic. Since we do not measure all tracking technologies, our results are a lower bound of what tracking exists on these sites.

3.4.2 Analyzing Registered Domains. We first conducted our privacy discovery process against a sampling of the registered domains identified by our web scraper. Leveraging the 16,693,933 first-party URLs and 3,530,623 third-party URLs, we extracted 13,277 first-party and 176,115 third-party registered domains. From those, we filtered the third-party registered domains down further by selecting a random sample of 3,167 domains. Figure 1 outlines this process. We had to take a sample because it typically took between 5 and 10 seconds to analyze a URL in Puppeteer; going through the entire list of URLs would take significant time.

We then fed the full list of 13,277 first-party and 3,167 third-party registered domains samples to our headless browser. Parsing the results for each of these domains, we extracted information on unique third-party trackers (as defined by the registered domain issuing the tracker), unique cookies (as defined by the registered domain issuing the cookie), keyboard and mouse behavior event listeners, Meta Pixel events and session recorders. This allowed us to construct some lower-bound insights into their respective privacy practices.

3.4.3 Analyzing Login Pages. We next turned to an examination of login pages both on first-party links and third-party links. We focused our attention on login pages as we wanted to ascertain the degree to which these sites gathered information about visitors. As shown in previous work [75], if a webpage asks users to enter the email address (typically through a login form), and if a third-party Javascript is present on the page, the third-party has visibility into the email field and can potentially exfiltrate this data.

To generate our inventory of login pages, we iterated through all first-party and third-party links that our web scraper discovered (Section 3.3.1). For each link, we searched for the term “login” in the URL and/or within the `<title>` tags of the page HTML. We then divided these results into first and third-party links.

We utilized random sampling to select 5% of the login page links for each first and third-party registered domain giving us a list of 18,960 first-party login URLs and 10,244 third-party login URLs. Figure 1 outlines this process. Feeding these two lists into our privacy discovery process, we again extracted information

on unique third-party cookies, ad trackers, mouse and keyboard behavior event listeners, Meta Pixel events, and session recorders.

To confirm the accuracy of our database query for login pages, we randomly sampled 100 pages from our query results and manually visited each website. We visually scanned the website's contents to determine if the page contained a login prompt or if it was a false positive. Testing confirmed that 93% of the manually visited sites had a login prompt. Thus, roughly 7% of the pages we scanned for our login page privacy results may have been false positives.

4 FINDING 1: SCHOOLS EXPERIENCE PRIVACY AND SECURITY INCIDENTS BUT LACK RESOURCES TO HANDLE THEM [RQ1]

To understand the scope of EdTech use in K-12 schools, we investigated how our participants spoke about and handled student privacy and security with technology use in their schools and school districts (as summarized in Table 4).

4.1 School Privacy and Security Incidents Occur With Varying Responses

Under half of the participants (7/18) told us that no privacy or security incidents involving student data have occurred in their current district. However, two of these eight participants did mention viruses and phishing attacks. At least 7/18 participants mentioned that a security incident occurred in their school district, but only 5/18 participants were willing to divulge incident details. For instance, P2, P3, P4, and P5—all from school district SD2—described a ransomware attack in their district. Despite assistance from the Federal Bureau of Investigation (FBI), this school district could not recover the information taken. P2 explained, *"I believe we lost most of that data because at that time we did not have a proper backup and restore solution in place or disaster recovery. Right after that attack, we implemented both of those things, so now we should be covered in that event."* Similarly, P11, whose role has him working with multiple districts, said that two of the districts he works with had experienced ransomware attacks in recent years.

Some of the incidents explicitly involved EdTech: 4/18 participants reported having a privacy incident that served as a “*wakeup call*” to get their EdTech data practices in shape. For instance, in a typical example from our data, P9 and P16, from SD5, described a data breach in an EdTech product formerly used by their district, compromising information on 35 former students. As a result, the school terminated their contract with the EdTech vendor. In another example of a school-based privacy incident, two participants representing two different school districts, P10 and P15, told us that secure files had been accidentally leaked over email in their past or current school districts. In P10’s case, an EdTech company sent secure files to a district staff member which contained sensitive information about multiple students; the staff member then sent these files to parents, unknowingly leaking data in the process. P10 said, *"we actually dropped the company, to be honest, because we felt like they were so lax in such an important area."* By contrast, in P15’s school, a teacher accidentally attached the incorrect documents to an email, thereby leaking sensitive information.

4.2 K-12 Privacy and Security Incidents Response Protocols Are Not Well Developed

Only 7/18 participants—representing 6/11 school districts—reported that they have an incident protocol in their school district for dealing with technology-related privacy and security incidents. The details of those protocols varied widely. Four participants said that EdTech companies typically agree to specific protocols for data breaches in contracts. For example, P11 explained that for technologies complying with the Student Data Privacy Consortium, *"there are very specific requirements for notification post breach and, at a bare minimum, keeping us in the loop on their remediation efforts."* Three participants mentioned that informing parents of the data breach is part of their district’s incident protocol, and two participants said that the authorities, possibly including the FBI, may be informed as part of their protocol depending on the severity of the incident. In other cases, school districts said they conduct a forensic analysis to determine why an incident occurred. For instance, P15 explained that a large part of their protocol is about understanding why the incident happened so they can learn from their mistakes. At least 2/18 participants reported that their district does not have a protocol for security and privacy incidents. Notably, P6 explained that, while his district has technological systems to recover lost data, such as hosting student information on the cloud, they do not have a specific protocol for dealing with security and privacy incidents. Responding to privacy and security incidents requires timely responses, knowledge on how to contain the incident, and knowledge on how to notify those affected and introduce mitigations. However, our findings suggest participants lacked these basic safeguards, particularly around EdTech for students.

5 FINDING 2: PRIVACY AND SECURITY AWARENESS AND TRAINING IS LIMITED IN K-12 [RQ1]

5.1 Low Privacy and Security Awareness in Decision Makers Who Purchase or Utilize EdTech Products

Our participants told us that school officials, IT personnel, and teachers are all—in varying degree and scope—responsible for deciding what EdTech to use in school classrooms for students; although, different stakeholders’ decisions could affect technologies used in all the schools in a district, one school, or just one classroom. Participants had views on how aware each stakeholder is about student data privacy. All 18/18 participants we spoke with were concerned about the privacy of their students’ data in the hands of EdTech vendors when we asked about EdTech usage in their school districts. Only two participants made allusions to their own awareness, however, the participants told us that they felt that teachers are the least aware of the potential threat to student privacy caused by EdTech as compared to IT personnel and district officials. For instance, P18 commented that *"teachers are oblivious to all that."* It was also evident in our interviews that district officials relied heavily on IT officials to properly assess and safeguard student’s privacy with technology usage in schools.

Participants felt that teachers’ lack of awareness is a particular issue when teachers want to bring new EdTech into the classroom.

Data Reported In Interviews	# Participants	% Participants	# School Districts	% School Districts
Concerned about student data privacy	18/18	100.00%	11/11	100.00%
Privacy and Security Incidents				
No privacy or security incidents	7/18	38.89%	5/11	45.45%
Experienced a security incident	7/18	38.89%	4/11	36.36%
Willing to divulge incident details	5/18	27.78%	2/11	18.18%
Privacy incident improved data practices	4/18	22.22%	3/11	27.27%
Incident Protocol				
Have an incident protocol	7/18	38.89%	6/11	54.55%
Does not have an incident protocol	2/18	11.11%	2/11	18.18%

Table 4: Summary of the data reported by participants (n=18) and the school districts (n=11) they represent regarding concerns about student privacy and security, schools experiences with privacy and security incidents, and whether they have security and privacy incident protocols. If nothing was mentioned, it is not included in the table.

Without an understanding of the potential harms of EdTech's data harvesting, teachers are unable to gauge the trade-off between purported functionality of an EdTech product and the product's privacy risks. P3 summed up the trade-off situation well, explaining that a teacher should not introduce an EdTech product whose sole purpose is *"just to tell people, 'hey, you need to bring in your subtraction homework tomorrow."* P3 felt that added value of an additional communication channel was not worth the additional risk of adding student data to yet another EdTech product and thought that it may not be easy for a teacher to come to the same conclusion. In the same vein, P8 told us how he decided that his school would use Microsoft Teams instead of Zoom in the spring of 2020, early in the COVID-19 pandemic, because he noticed that Zoom had serious security and privacy issues. As a result, he received what he felt was harsh feedback from teachers who wanted to use Zoom and were unaware of the security and privacy risks that the software posed. This participant later felt vindicated for his decision when Zoom came under fire for Zoom bombing issues [90]. In these instances, participants indicated that teachers' goals and student data privacy were not always well aligned because they prioritize learning objectives and lack an awareness of privacy and security risks to students.

5.2 Privacy and Security Training for Teachers Does Not Address EdTech Privacy Issues

Participants from all 11 districts told us that they used professional development sessions to help teachers learn new technology-related and EdTech-related skills but that these trainings often lacked information about student data privacy and security. Instead, the vast majority of that training—for 9/11 school districts—was focused on immediate needs such as how to use specific pieces of software such as Screencasts, SMART Notebook, and PowerPoint to enhance their teaching. Two participants also stated that they do not offer privacy training at all in their respective districts. However, participants from half of the school districts explicitly mentioned that they did have some form of cybersecurity training, with privacy included as a small component. P11 explained that one training session is *"a good way to knock out both at the same time. And even though they're separate concepts and very different, they're often interrelated in most people's minds."* Participants spoke of how cybersecurity

topics get more emphasis because schools are concerned about their teachers falling for malicious tactics such as phishing attacks. The privacy elements that are discussed in training sessions varied from school district to school district, but our participants were able to only concretely describe the social media related trainings which discussed what data teachers can release about students online. Student data privacy issues around EdTech use, such as how to treat student data, who to provide access to, and general privacy practices around student data, were not covered.

6 FINDING 3: CURRENT METHODS TO ACQUIRE EDTECH FOR SCHOOL DISTRICTS DO NOT FULLY CONSIDER PRIVACY AND SECURITY [RQ2]

Participants spoke of two different pathways that new educational technologies can enter the classroom: a top down or bottom up approach. However, with both approaches, it was evident that privacy and security considerations for students' privacy were limited in these EdTech acquisition processes. 17/18 participants—which hail from 11/11 school districts—told us they used a top down approach for paid EdTech solutions where a district license is required for a subset of its students to use an EdTech product. For example, participants mentioned that if a teacher finds a product that they find interesting and requests it from the school district, the district would then acquire a license for this teacher's students to use that EdTech. In contrast, 11/18 participants told us that they also employ a bottom up approach for free technologies—or have seen it used in their school district—where an individual teacher directly introduces an educational technology to their classroom without making a request to their school district for permission to use that product. In these cases, the EdTech is not formally vetted and there are no checks that the EdTech product complies with school policies for student data privacy and security.

The root cause of these bottom up processes—as many participants explained—is the lack of teacher training on privacy and security around EdTech, not malicious violations of school policies. Summarizing this viewpoint from our data, P1 explained that a remedial training approach to correct the issue was not fool proof because *"there's no way you're going to 100% avoid an issue when*

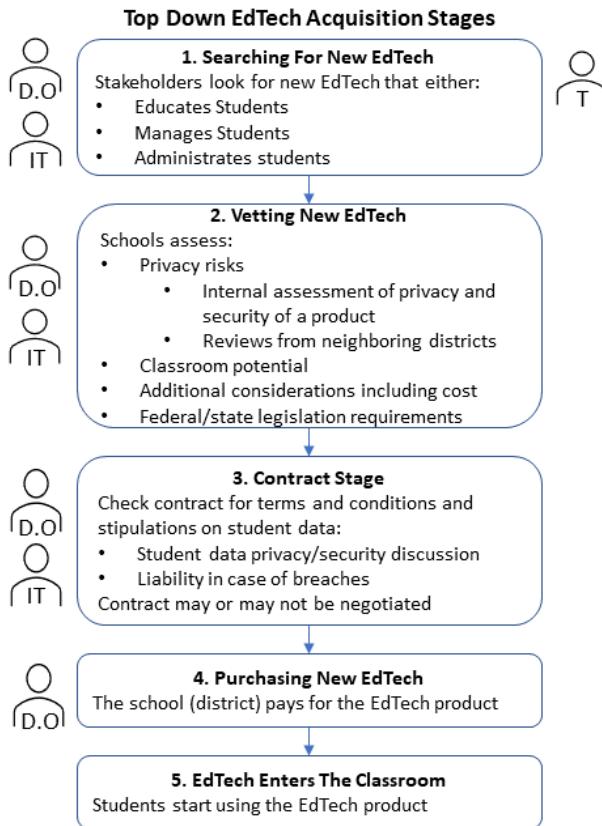


Figure 2: Our study identified 5 major stages for paid EdTech acquisition. We explain relevant considerations at each stage. The figures next to each stage denote which stakeholders are involved at a given stage. D.O. stands for District Official, IT stands for IT Personnel, and T stands for teacher.

you're dealing with the number of people we have in the district. So it's really just a matter of raising as much awareness as you possibly can and then keeping an eye out, just make sure that people aren't doing things that they shouldn't be doing." And P12 spoke of trying to educate teachers about the issues: *"It has to go through the checks and it has to be approved because if it's a free product, they're making money off of you somewhere. If you're not paying for it, then they're selling the data."*

IT personnel did speak of how they minimize potential issues by: (1) locking down school-issued devices so that technologies cannot be installed without prior approval, (2) refusing access to student information systems by unapproved software, (3) requiring all student accounts to use single-sign on that only works for approved applications, and (4) blocking unapproved web applications. In other cases, participants wanted to give teachers freedom to choose EdTech that fit their learning goals and wanted to minimize red tape processes so they were not unhappy about bottom up EdTech acquisitions. Given that our participants do not participate in the bottom up process, we will focus on the paid top down process as shown in Figure 2.

6.1 Top Down Acquisition of Paid EdTech Limits Considerations for Student Privacy and Security

Participants described which stakeholders are involved in identifying, vetting, and purchasing paid EdTech; what criteria are used to identify EdTech of interest; and the nuances of educational technology contracts. They also highlighted when, if at all, student privacy and security issues are considered and how much leeway they have to choose technologies that safeguard K-12 student privacy.

6.1.1 Searching for New EdTech. The first stage of top down acquisition of paid EdTech involves finding new EdTech for students to use. Our participants told us that they had a diverse team of district affiliated stakeholders involved in finding and vetting new EdTech. The majority of participants (10/18)—representing 7/11 school districts—have a process where teachers find new EdTech products and submit a form to start the technology acquisition process. In addition, 6/18 participants—representing 5/11 school districts—have a designated committee that is responsible for finding and vetting new EdTech. These committees are comprised of various stakeholders, including IT personnel, district officials, and teachers. A few of the participants mentioned that they have both a process for teachers to find new EdTech and a designated committee. At this point, privacy and security is rarely a primary consideration.

6.1.2 Vetting New EdTech. The next stage is vetting EdTech to see if the product is useful and or has any considerations that would affect its use or purchase by the school. At this point, schools consider privacy and security issues for students but are limited by their knowledge and training on these issues. District officials, IT, and even the school board are involved in vetting EdTech to various degrees depending on the district. At least 7/18 participants spoke of how the technology director or IT department advises technology purchasing decisions, typically with the responsibility of vetting privacy policies and ensuring that the new technology will mesh with existing infrastructure. Fewer participants, 5/18, told us that their superintendent or another administrator is involved in EdTech purchasing decisions, primarily to approve the cost. Only one participant said they needed school board approval to purchase new EdTech, which was for budgetary reasons. To aid in the vetting process, 7/18 participants said they rely on peers' suggestions—*i.e.*, neighboring school districts—when vetting an EdTech product. In a typical response, P18 said that discussing an EdTech product with peers *"gives you all the pros and cons before you buy."* The remaining participants were either part of a group interview or did not explicitly state who was involved with vetting new EdTech.

Participants discussed four factors for choosing and vetting which EdTech product to buy for their school or school district: (1) privacy risks, (2) classroom potential, (3) additional considerations and constraints not related to students, and (4) legal requirements. First, 16/18, said privacy was a consideration for EdTech acquisition but few participants could provide concrete details. At least 4/18 participants told us they use a privacy alliance or organization to evaluate the security and privacy of an EdTech product. For instance, P1 described that a student data privacy agreement created through a privacy alliance *"is a nationally recognized document that [...] essentially outlines all the things that [we] would want in terms*

of regulations for privacy of student data" and how Schoology, one of the popular EdTech vendors, had signed it. Others were not specific on how their districts evaluate these issues for student data privacy, and at least one felt that it was not always necessary to vet all educational technology. For example, P7 does not evaluate security and privacy concerns of acquired EdTech products because her district uses the vendor's reputation as a proxy. Second, 11/18 and 7/18 participants respectively mentioned classroom potential for educators and students as the second most important factor. Third, participants expressed assessing factors not related to instruction including cost (7/18), quality of product support (3/18), and COVID-19 induced need (2/18). Finally, at least 5/18 participants discussed how state-level data privacy legislation created a more stringent vetting process for new EdTech entering their schools and districts. For example, in Illinois, an amendment of Illinois's Student Online Protection and Privacy Act (SOPPA) imposed more onus on schools to keep student data safe starting July 1, 2021 [30]. When discussing the examination of data usage and retention in their vetting process, P16 explained that "*with SOPPA it's explicit. Before that, it wasn't necessarily something [we considered] when we did our [internal] review. It may not have always been top of mind to have that explicitly laid out.*" These participants sometimes viewed their privacy legislation-induced vetting process as alleviating the need to vet an EdTech product. For example, in Connecticut, state data privacy protections enabled P10's school district to, without vetting, sidestep Zoom security and privacy woes that plagued school districts nationwide [90]. Since Zoom had not pledged compliance with Connecticut's privacy law, she relayed that "*Zoom was far superior [compared to other videoconferencing products] in the spring of '20 [...] and in the State of Connecticut [Zoom] couldn't meet our compliance laws for that first semester, so we couldn't use them.*"

6.1.3 Contract Stage. After vetting an EdTech product, our participants told us that a school district typically needs to enter into a contractual agreement between the district and the EdTech vendor. This contract specifies details such as how long the software service will last, how much the software will cost, and what rights and responsibilities each party has. Only a few participants spoke of how multiple school district leaders—such as a business manager or the superintendent—and IT reviewed EdTech contracts before signing off. Some participants, 4/18, said that the IT department—or specifically the IT director—vets contracts for security and privacy of student data but expressed reservations about the IT department's ability to review contracts effectively. The remaining participants did not explicitly state who was involved with reviewing EdTech contracts. Participants also made it clear that often the people reviewing the contracts do not necessarily know what to look for regarding student data privacy or security issues.

Student Privacy and Security Rarely Considered Outside Of Boilerplate Language. Most participants, 11/18, reported that contracts typically specify what data can be collected, how that data is used/stored, and who can access that data. Overall, however, participants' responses varied regarding the specificity of information in contracts. 7/18 participants specifically mentioned sharing data with third-parties as an important topic covered in contracts as summed up by P2: "*I would like to know exactly who we're going to be sharing our data with [...] and if they're trying to share our data with*

somebody else, [you] never really know what you're getting into." Indeed, in another district, P12 fastidiousness on this issue has caught potential red flags with prospective EdTech vendors in the past. For example, she vetted a math enrichment program which asked the district to "*include social security numbers [and students'] free and reduced lunch status*" which she surmised was not necessary to help students learn math. The requested data was a deal breaker in this instance. However, this sentiment does not translate to a zero-tolerance approach to student data sharing. Three participants said they expect vendors to share data with companies that audit or review EdTech companies, while one participant said that they allow an EdTech vendor to share with third-parties as long as they list them in the contract. Finally, three participants mentioned that liability is an important piece of the contract because it outlines the specifics of data breach disclosure and liability for paying for identity theft protection. For instance, P12 told us they look for contractual language that "*spell[s] out where there liability is or where their liability is limited. That way we know what we're getting ourselves into.*"

Limited Ability To Negotiate EdTech Contracts. 10/18 participants discussed whether a district can negotiate what software companies can and cannot do once they have direct access to students. These participants described that contracts tend to be "boilerplate" and lack customization with some negative consequences as summed up by P12: "*[T]he hardship [with vendor contracts] is still the same place. There's no option for negotiation. You either accept what they're doing, or you go find a different vendor.*" Participants also told us that companies do not want to create bespoke agreements with schools because, unlike other organizations' business-to-business transactions, they lack purchase volume.

For participants involved in larger privacy groups—such as the Student Data Privacy Consortium—which create general agreements that vendors can sign and use to replace or supplement their contracts with multiple schools, boilerplate contracts are a positive because these agreements verify that a vendor meets a certain safety standard. Moreover, P6 described that districts can modify a boilerplate contract and negotiate what data they agree to share with the vendor. However, only two participants—from affluent districts—mentioned negotiating EdTech contracts to address student data protection. In summary, although top down acquisition of EdTech products can have some safeguards for student privacy and security, the checks are not extensive and teachers can still bring technologies into the classroom outside of this process that do no undergo checks.

7 FINDING 4: DISTRICT WEBSITES TACITLY ENDORSE MANY DOMAINS WITH POTENTIAL PRIVACY ISSUES (RQ3 AND RQ4)

To complement our previous findings regarding our participants limited privacy and security awareness, and an initial list of EdTech vendors reported by interviewees (see Appendix D), we sought to understand, at scale, what EdTech products schools were likely using or endorsing based on what the school/district websites linked

to online. Combined with privacy scanning, we are able to understand how widespread the lack of privacy and security awareness is in the K-12 space.

7.1 What Are the Top EdTech Vendor Domains Linked From School/District Websites? [RQ3]

Table 3 reflects the results of our categorization of the top 300² most frequently listed third-party domains on 15,573 public school/district websites in the US, which matched and expanded on EdTech vendors mentioned by our interviewees. Note that, given the overlap in categories, some EdTech can fall into multiple categories; hence, the individual percentages fail to total 100%. In the Communication category, for instance, several EdTech products did not have communication as the platform's sole purpose. Rather, communication was often an additional feature to enrich the underlying use case of the product. For example, Schoology's primary purpose is a learning management system, but it also boasts functionality to communicate with parents about their student's grades, assignments, and class announcements.

Within the top 300 third-party domains found, 110/300 fit our definition of "EdTech" as defined in the methods section, 16/300 were broken URLs (*i.e.*, unreachable site at the time of scraping), and the remaining 174/300 domains we classified as not "EdTech" because they either did not collect student data or were not primarily marketed to K-12 schools and students. For instance, as seen in Table 5, only the last two entries—zoom.us and scholastic.com—of the top 10 linked domains qualified as EdTech. We also note that 28/110 of the EdTech links were also mentioned by participants in our interviews as being used in their school districts. Meaning, that our participant's use of popular services and their struggles with EdTech vendors, outlined in section 6.1.3, may be mirrored by the hundreds of districts who also tacitly endorse the web-based EdTech product. The full data set is included in the Appendix.

The 110/300 EdTech domains fit into the established subcategories found in Table 3, the three most common EdTech purposes were 32.7% Delivery of Instructional Content, 24.6% Formative Assessments, and 23.6% Administrative Management which likely reflect a school's most pressing needs. The least prevalent categories for EdTech were Video Communication and School Merchandise Management, both accounting for 1.8% of links. Our findings corroborate other market reports about EdTech prevalence, for instance, Delivery of Instructional Content/Classroom Engagement & Instruction is the top EdTech product category [40] and add quantitative evidence to complement other qualitative examinations of EdTech [5, 37, 94].

All 110/300 EdTech domains were linked from websites of schools and districts which we obtained from the NCES list and K12 SIX list, as described in section 3.3.1. We assume that schools and districts linking to these domains implies endorsement and even potential usage of these EdTech services. To provide further evidence for usage, we identified which of these linked EdTech pages asked users to login. For instance, kidsa-z.com, an online learning portal, was linked from 625 seed domains (out of a total of 15,573 seed domains); the linked webpage asked for login information on 606 of

²We limited our analysis to the top 300 to make it tractable.

the 625 domains (which we manually verified). In other words, if a school or district linked to kidsa-z.com, there is a $606 / 625 = 0.97$ probability that the linked page was a login page. This probability is the highest for kidsa-z.com, followed by i-ready.com (0.93 probability), an online learning platform, as well as safeschools.com (0.93 probability), a safety compliance platform. Across the 110 EdTech domains, the mean probability that the page linked from schools/districts asked for login information is 0.16 ± 0.26 .

The 174/300 non-EdTech domains that school/district websites list vary wildly. Many non-EdTech domains linked are general audience productivity and document creation software, such as Adobe and DropBox, which are used by schools despite not directly marketing to them. Additionally, school/district websites link to social media platforms such as YouTube, Instagram, and Blogspot which can be used for outreach and education, but do not meet the definition of EdTech. Non-EdTech domains also included US-affiliated institutions of higher education, US government websites (top-level domain .gov), and multiple physical and mental health resources, *e.g.*, domains for the American Academy of Pediatrics, the American Psychological Association, and the Child Mind Institute. Of the remaining domains, many were educational resources not marketed to schools, including PBS Kids, Common Sense Media, and news sites, *e.g.*, the New York Times and CNN.

60/174 non-EdTech domains were coded as being potentially in possession of student data based on their privacy policies. Some of the sites which had student data but were non-EdTech include social media sites such as twitter.com, career-oriented sites such as linkedin.com, health resource sites such as crisistextline.org, and survey platforms such as surveymonkey.com. Sites that were non-EdTech and did not have student data were primarily the government and university sites which hosted information that did not require a login to access, *e.g.*, stopbullying.gov. Thus 20% of the top 300 most linked domains are not necessarily EdTech but still handle student data.

7.2 What Are the Potential Privacy Issues of EdTech Vendor Domains Linked From School/District Websites? [RQ4]

We analyzed both school/district (first-parties) and vendor (third-party) websites for potential privacy issues. As described in our methodology (Section 3.4), for every subset of analyzed links, we extracted usage information concerning third-party cookies, third-party trackers, mouse and keyboard behavior event listeners, Meta Pixel events, and session recorders. We have summarized our results across first and third-party websites as well as for both registered domains and logins in Table 6.

We note here that many school websites are actually subdomains on school/district websites. For example, we find both altaview.canyonsdistrict.org and albionmiddle.canyonsdistrict.org appear as subdomains of the canyonsdistrict.org school district website. We also note here that keyboard and mouse behavior event listeners are specific forms of session recording however the existence of one does not necessitate the other. A JavaScript that registered as a keyboard event listener for example, might listen for keyboard movement but then not actually transmit any data. A session recorder on the

All Domains	# Referring Schools	EdTech Domains	# Referring Schools
1. google.com	10913	1. zoom.us	3290
2. facebook.com	9337	2. scholastic.com	3131
3. youtube.com	7910	3. collegeboard.org	3079
4. twitter.com	6943	4. powerschool.com	2707
5. instagram.com	4869	5. khanacademy.org	2669
6. ed.gov	4406	6. frontlineeducation.com	2513
7. cdc.gov	4347	7. act.org	2361
8. apple.com	3589	8. smore.com	2076
9. zoom.us	3290	9. starfall.com	1973
10. scholastic.com	3131	10. clever.com	1914

Table 5: Top 10 domains linked overall and top 10 EdTech domains linked from 15,573 school/district websites. The 3 bolded domains were also mentioned by our interview participants as being used in their school district.

other hand, is designed to capture and send mouse and keyboard activities to third-parties. We therefore make the claim session recorders present a potentially greater privacy risk than keyboard or mouse event listeners. Further, while our analysis might have detected these sorts of scripts, companies typically only record a sample of website visits so it is unclear what percentage of users are actually being recorded.

7.2.1 Potential Privacy Leaks on Registered Domains. In performing our privacy analysis, we first looked at registered domains of both school/district websites and linked vendor websites. Table 6 calls out some high level privacy statistics and tells us that nearly all of the top sites discovered through our scraping are making extensive use of tracking technologies. Cookies and trackers are not necessarily malicious or harmful in nature however they do still record information about the visiting party. A potentially more alarming discovery was the number of sites that utilized a session recorder (7.4%). Anyone visiting those sites would have their entire session captured which includes information such as which links they clicked on, what images they hovered over, and even data entered into fields but not submitted. This could include data that users might otherwise consider private such as the autofilling of saved user credentials or social network data [1, 75].

7.2.2 Potential Privacy Leaks on Login Pages. Having examined registered domains, we next turned to the subset of login pages from both school/district websites and third-party sites. As stated in section 3.4.3, we searched for the term “login” in the URL and/or within the `<title>` tags of the page HTML. We then divided these results into first and third-party links and performed our privacy analysis against each subset. At a high level, Table 6 shows that first-party login pages utilized more mechanisms which are traditionally associated with session persistence (e.g., cookies) while third-party login pages leveraged more mechanisms which are traditionally associated with information gathering (Meta Pixel and session recorders) but not by a large margin. Further, we find that both categories of login pages made use of both mouse and keyboard event listener technologies, although we saw more of these employed by the first-party websites than we did the third-party sites. When paired with the observed session recorders on third-party login pages, this calls into question the confidentiality of

login data and, logically, any further information shared during the remainder of the session. This could then lead to any number of situations for students including credential compromises, data loss events, or even the unauthorized capture and subsequent sale of private information.

While all of our interview participants were concerned about the privacy of their students’ data in the hands of EdTech vendors (Section 4.2), our technical findings indicate that those concerns are warranted. It is clear from our results that there is a lot of tracking which occurs both on school/district websites and those third-party sites which students and other users are directed to visit. The sites which contain login pages present a particularly interesting subset of results as we can argue that they represent a stronger degree of “use” by the school population. For the login pages we examined, we saw widespread use of tracking technologies which potentially represents significant privacy risk to students.

8 DISCUSSION: RECOMMENDATIONS, LIMITATIONS, AND FUTURE WORK

Based on our findings, we discuss methodological limitations and suggestions for improvements as well as recommendations for researchers, school districts, and policymakers.

8.1 The Need for Empirical Evidence on EdTech Privacy and Security Implications in K-12

This paper scratches the surface of EdTech privacy issues, as our interview study sample size is small, and our privacy analysis is restricted to publicly-facing sites. We need to collect and analyze more data on EdTech privacy, so that the research community could have a deeper understanding of the issues at play. In this way, districts could make data-driven decisions (e.g., on training), and policymakers could regulate the EdTech industry and/or education institutions based on real-world evidence. To achieve these goals, we list limitations and proposed next steps.

8.1.1 Less Biased Sampling. Our recruitment strategy for our user study may have had a self-selecting bias since those interested in privacy and security may have been more likely to participate. Our interview sample is also non-exhaustive since it does not consider all school districts in the United States. Future work could gather

Tracking Measure	First-party (School/District) Websites		Third-party (Vendor) Websites	
	Registered Domains	Login Pages	Registered Domains	Login Pages
Third-party Cookies	86.91%	95.58%	92.44%	90.7%
Third-party Trackers	74.06%	68.49%	84.89%	50.39%
Mouse Event Listeners	84.3%	84.24%	88.0%	69.51%
Keyboard Event Listeners	69.29%	81.16%	74.96%	55.81%
Session Recorders	0.55%	0.17%	7.41%	3.1%
Meta Pixel	6.83%	1.51%	26.52%	6.46%

Table 6: Privacy statistics from sampled registered domains and login pages of school/district and vendor websites. Each percentage indicates the prevalence of a given tracking measure for each category. Notably, the use of session recorders and the Meta pixel on vendor websites can indicate enhanced tracking of students private data and activities across the web.

a more extensive dataset through a large-scale survey to capture greater variation in schools' funding, location, and size. Future work could also further investigate privacy and security issues with teachers' bottom up acquisition of free EdTech.

8.1.2 Understanding Data Flows in Private Services. Our web scraping discovers publicly facing school/district websites and linked EdTech sites. It is unclear what EdTech services (including websites and mobile apps) schools/districts use that are not publicly linked (e.g., behind authentication screens), what privacy implications these non-public technologies present, and how the observation is similar to or different from public-facing services. Furthermore, for both public-facing and private EdTech services, it is unclear how sensitive data flows, e.g., in terms of contextual integrity for privacy [56]: what data is transmitted, how it is transmitted, and to whom it is transmitted. Our current analysis simply shows that, for instance, sensitive data, such as the username or email address, *could potentially* be transmitted to session-recording third parties on login forms, but we do not have evidence that this data is *actually* transmitted. Similarly, the Meta Pixel *could potentially* be used to track students across the web but we need further evidence that this *actually* occurs and what utility this has for pedagogy if at all. Additionally, some private data may not be transmitted over the web. For instance, certain EdTech products (such as Lightspeed, an online safety platform [92]) are integrated with schools' existing EdTech accounts (e.g., Google Education) through backend APIs, such that no user data is directly transmitted from the user to these integrated platforms. Such behind-the-scene data sharing cannot be discovered by web scraping alone (even when conducted on private networks).

To peer behind the walled gardens of non-public EdTech services, we suggest the CHI research community develop new methods to collect evidence from a user's and/or IT administrator's vantage point. Leveraging crowdsourcing, researchers could develop and deploy extensions similar to prior work to understand how Google Education accounts are sharing data with EdTech products [7], or how sensitive information is actually being transmitted to third-party webpages [75].

8.2 The Need for Training on Privacy and Security Around EdTech in K-12

The schools in our study often lacked the knowledge and resources to handle the privacy and security incidents they experienced, and their key decision-makers lacked privacy and security awareness

and training on EdTech. Moreover, our technical analysis suggests that the privacy risks such as enhanced tracking of students by EdTech vendors could be problematic. Future work could investigate how to help schools develop standardized privacy and security incident response protocols. This effort could be supplemented with more transparency into, and more systematic understanding on, actual versus perceived privacy issues at schools/districts. Improved EdTech-related standards at a national level with accompanying resources for teacher and IT professional development could better equip school districts to manage privacy and security challenges that arise with increasing technology use. For instance, policymakers could require, and fund, high-quality privacy and security training for all school employees and students to help keep student data safe. Encouraging or making schools aware of privacy-and-security-focused foundations and consortia, such as the Common Sense District Privacy Program [18], can also raise awareness—the 6/11 districts who were members of such groups, for instance, were more knowledgeable in general about EdTech and potential vulnerabilities in their schools.

In addition, researchers could create privacy checklists for various stakeholders, such as district officials or teachers, which would scaffold how to assess the privacy risks of an EdTech product. For district officials, the checklist would be more oriented to EdTech acquisition issues such as “appropriate third-parties to receive your students' data” while for teachers the checklist could focus on operationalizing specific privacy hazards of using EdTech products day to day—e.g., check to ensure that students are not unintentionally sharing their work on the open web—and considerations for privacy settings on those products.

8.3 The Need for Enhanced Regulation of EdTech in K-12

Participants used a wide variety of heuristics to vet EdTech products in the top down approach. Some of these seemingly more robust voluntary heuristics—e.g., checking to see if the vendor signed on to Future of Privacy Forum's Student Privacy Pledge—may still be problematic since there are no repercussions for violating the pledge [19, 64]. However, our findings do suggest that state data privacy legislation increases district official and IT privacy and security awareness and enhances vetting protocols for EdTech products. Ideally, improved student data privacy legislation at the federal level, such as enhancements to FERPA or new legislation to cover new types of EdTech products and to curb unnecessary

data collection on students, would create a national standard of protection for all students.

Although this is a long term and difficult goal to achieve, we believe that improved transparency into actual and perceived school/district privacy challenges, per section 8.1, could potentially motivate more grassroot actions toward legislation and/or influence the policy making process. Improved privacy laws at the state level, such as California Privacy Rights Act [41]—even when not EdTech specific—can also go a long way to enhance privacy protections and raise awareness about these issues around technologies students use. We recommend that future work builds on our categorization of what EdTech is being endorsed tacitly by school/districts to determine how these EdTech products collect and use data on students to better inform legislation improvements.

9 CONCLUSION

In this paper, we examined privacy and security challenges in K-12 public schools across the United States. We found that school officials, IT personnel, and teachers lack resources to deal with privacy and security incidents more generally and around EdTech, given that limited privacy and security training is offered on these issues. Even at procurement time, key decision makers in school districts do not fully consider the potential privacy and security implications of EdTech products for their students and have little room to negotiate with companies around these issues. Additionally, we found preliminary evidence that the EdTech space is expanding with many technologies holding student data that are not typically considered as such by current legislation. Finally, we also uncovered potential privacy issues for student EdTech users with the extensive usage of third-party trackers and cookies on school/district websites and third-party domains, such as the use of session recorders. Based on these findings, we recommend that the CHI community engage in future explorations of what data EdTech products collect on students in K-12; that policymakers consider revisions to existing federal laws, or enacting state privacy legislation, to account for the potential privacy issues EdTech products entail and their creeping scope; and that more resources be allocated at the school district level to improve training on privacy and security around EdTech for school district officials, IT personnel, teachers, and students.

ACKNOWLEDGMENTS

We are grateful to our study participants. We thank our reviewers whose suggestions helped improve our paper. This work was funded by the National Science Foundation CAREER Award #2047465.

REFERENCES

- [1] Gunes Acar, Steven Englehardt, and Arvind Narayanan. 2020. No boundaries: data exfiltration by third parties embedded on web pages. *Proceedings on Privacy Enhancing Technologies* 2020, 4 (2020), 220–238.
- [2] Gunes Acar, Christian Eubank, Steven Englehardt, Marc Juarez, Arvind Narayanan, and Claudia Diaz. 2014. The web never forgets: Persistent tracking mechanisms in the wild. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. 674–689.
- [3] Gunes Acar, Brendan Van Alsenoy, Frank Piessens, Claudia Diaz, and Bart Preneel. 2015. Facebook tracking through social plug-ins. *Technical report prepared for the Belgian Privacy Commission* (2015), 1–24.
- [4] Access4Learning. 2022. Student Data Privacy Consortium. <https://privacy.a4l.org/>.
- [5] Frida Alim, Nate Cardozo, Gennie Gebhart, Karen Gullo, and Amul Kalia. 2017. *Spying on Students: School Issued-Devices and Student Privacy*. Technical Report. Electronic Frontier Foundation. 49 pages. <https://www.eff.org/files/2017/04/13/student-privacy-report.pdf>.
- [6] Amelia Pak-Harvey. 2019. Nevada Students' Information Exposed in Data Breach. *Las Vegas Review-Journal* (Aug. 2019). <https://www.reviewjournal.com/local/education/nevada-students-information-exposed-in-data-breach-1817032/>.
- [7] David G. Balash, Xiaoyuan Wu, Miles Grant, Irwin Reyes, and Adam J. Aviv. 2022. Security and Privacy Perceptions of Third-Party Application Access for Google Accounts. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, Boston, MA, 3397–3414. <https://www.usenix.org/conference/usenixsecurity22/presentation/balash>.
- [8] Gabriel Basset, C. David Hylander, Philippe Langlois, Alexandre Pinto, and Suzanne Widup. 2021. 2021 Data Breach Investigations Report. <https://www.verizon.com/business/resources/reports/2021/2021-data-breach-investigations-report.pdf>.
- [9] Eileen Belastock. 2022. "Our Biggest Nightmare Is Here": Cyberattacks are targeting school districts. How can schools respond to keep data and systems secure? *Education Next* 22, 2 (2022), 44–50.
- [10] Faith Boninger and Alex Molnar. 2016. Learning to be watched: Surveillance culture at school. *The eighteenth annual report on schoolhouse commercialism trends. National Center for Education Policy at the University of Colorado at Boulder*. <http://nepc.colorado.edu/files/publications/RB%20Boninger-Molnar%20Trends.pdf>. Accessed 17 (2016).
- [11] Ferry Boschman, Susan McKenney, and Joke Voogt. 2014. Understanding decision making in teachers' curriculum design approaches. *Educational technology research and development* 62, 4 (2014), 393–416.
- [12] Ben Burgess, Avi Ginsberg, Edward W. Felten, and Shaanan Cohn. 2022. Watching the watchers: bias and vulnerability in remote proctoring software. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, Boston, MA, 571–588. <https://www.usenix.org/conference/usenixsecurity22/presentation/burgess>.
- [13] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 981–992. <https://doi.org/10.1145/2858036.2858498>.
- [14] Irene L Chen and Libi Shen. 2016. The cyberethics, cybersafety, and cybersecurity at schools. *International Journal of Cyber Ethics in Education (IJCEE)* 4, 1 (2016), 1–15.
- [15] Catalin Cimpanu. 2019. Over 500 US Schools Were Hit by Ransomware in 2019. <https://www.zdnet.com/article/over-500-us-schools-were-hit-by-ransomware-in-2019/>.
- [16] Shaanan Cohn, Ross Teixeira, Anne Kohlbrenner, Arvind Narayanan, Mihir Kshirsagar, Yan Shvartzshnaider, and Madelyn Sanfilippo. 2021. Virtual Classrooms and Real Harms: Remote Learning at {US} Universities. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. 653–674.
- [17] Shaanan Cohn, Ross Teixeira, Anne Kohlbrenner, Arvind Narayanan, Mihir Kshirsagar, Yan Shvartzshnaider, and Madelyn Sanfilippo. 2021. Virtual Classrooms and Real Harms: Remote Learning at {U.S.} Universities. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. 653–674.
- [18] CommonSenseMedia. 2022. Common Sense District Privacy Program. <https://privacy.commonsmense.org/resource/for-districts-and-schools>.
- [19] Sophia Cope, Jason Kelley, and Bill Budington. 2021. FPF's 2020 Student Privacy Pledge: New Pledge, Similar Problems. <https://www.eff.org/deeplinks/2021/09/fpf-s-2020-student-privacy-pledge-new-pledge-similar-problems>.
- [20] Joseph Cox. 2017. Hacker Steals Millions of User Account Details from Education Platform Edmodo. <https://www.vice.com/en/article/ezbwe/hacker-steals-millions-of-user-account-details-from-education-platform-edmodo>.
- [21] Shivangi Dhawan. 2020. Online Learning: A Panacea in the Time of COVID-19 Crisis. *Journal of Educational Technology Systems* 49, 1 (Sept. 2020), 5–22. <https://doi.org/10.1177/0047239520934018>.
- [22] Anirudh Ekambaranathan, Jun Zhao, and Max Van Kleek. 2021. "Money Makes the World Go Around": Identifying Barriers to Better Privacy in Children's Apps From Developers' Perspectives. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3411764.3445599>.
- [23] Federal Trade Commission. 2013. Children's Online Privacy Protection Rule. <https://www.ftc.gov/system/files/2012-31341.pdf>.
- [24] Federal Trade Commission. 2020. Complying with COPPA: Frequently Asked Questions. <http://www.ftc.gov/business-guidance/resources/complying-coppa-frequently-asked-questions>.
- [25] Richard Gomer, Eduarda Mendes Rodrigues, Natasa Milic-Frayling, and M.C. Schraefel. 2013. Network Analysis of Third Party Tracking: User Exposure to Tracking Cookies through Search. In *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*. IEEE, Atlanta, GA, USA, 549–556. <https://doi.org/10.1109/WI-IAT.2013.77>.
- [26] Google. 2022. Puppeteer | Tools for Web Developers. <https://developers.google.com/web/tools/puppeteer>.

[27] Holly Hobbs and Nick Marinos. 2021. As Remote Learning Increased, So Did the Cyber Threats Facing K-12 Schools. <https://www.gao.gov/podcast/remote-learning-increased%2C-so-did-cyber-threats-facing-k-12-schools>.

[28] IETL. 2022. Illinois Education Technology Leaders. <https://www.iletl.org/>.

[29] Illinois General Assembly. 2015. HB3527. <https://www.ilga.gov/legislation/publicacts/fulltext.asp?Name=099-0460>.

[30] Illinois General Assembly. 2019. Student Online Personal Protection Act. <https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=3806&ChapterID=17>.

[31] Jonathan Jackson and Laura Zieger. 2012. Teacher Acquisition of Educational Technology: An Activity Theory Perspective. In *Society for Information Technology & Teacher Education International Conference*. Association for the Advancement of Computing in Education (AACE), 2332–2339.

[32] Jacqueline M. Nowicki, Sherri Doughty, Jennifer Gregory, and Jessica Mausner. 2020. *Recent K-12 Data Breaches Show That Students Are Vulnerable to Harm*. Technical Report GAO-20-644. United States Government Accountability Office. <https://www.gao.gov/assets/gao-20-644.pdf>.

[33] Al Januszewski and Michael Molenda. 2013. *Educational technology: A definition with commentary*. Routledge.

[34] Rebecca Jeong and Sonia Chiasson. 2020. 'Lime', 'Open Lock', and 'Blocked': Children's Perception of Colors, Symbols, and Words in Cybersecurity Warnings. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13.

[35] Georgios Kambourakis. 2013. Security and Privacy in m-Learning and Beyond: Challenges and State-of-the-art. *International Journal of u-and e-Service, Science and Technology* 6, 3 (2013), 67–84.

[36] G. Kelly, J. Graham, J. Bronfman, and S. Garton. 2018. *2018 State of Edtech Privacy Report*. Technical Report. Common Sense Media, San Francisco, CA.

[37] G. Kelly, J. Graham, J. Bronfman, and S. Garton. 2019. *2019 State of Edtech Privacy Report*. Technical Report. Common Sense Media, San Francisco, CA.

[38] Priya C. Kumar, Marshini Chetty, Tamara L. Clegg, and Jessica Vitak. 2019. Privacy and Security Considerations for Digital Technology Use in Elementary Schools. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300537>

[39] Celeste Lawson, Colin Beer, Dolene Rossi, Teresa Moore, and Julie Fleming. 2016. Identification of 'at risk' students using learning analytics: the ethical dilemmas of intervention strategies in a higher education institution. *Educational Technology Research and Development* 64, 5 (2016), 957–968.

[40] LearnPlatform. 2022. *EdTech Top 40: Fall 2022 Report*. Technical Report. 8 pages. <https://learnplatform.com/top40>.

[41] California Legislature. 2020. California Privacy Rights Act. <https://cpra.gtlaw.com/cpra-full-text/>.

[42] Douglas A. Levin. 2022. *The State of K-12 Cybersecurity: Year in Review – 2022 Annual Report*. Technical Report. K12 Security Information Exchange (K12 SIX). 30 pages. <https://www.k12six.org/the-report>.

[43] Sonia Livingstone, Mariya Stoilova, and Rishita Nandagiri. 2019. Children's data and privacy online: growing up in a digital age: an evidence review. *London School of Economics and Political Science, Department of Media and Communications* (2019).

[44] Sonia Livingstone, Mariya Stoilova, and Rishita Nandagiri. 2019. Children's Data and Privacy Online: Growing up in a Digital Age: An Evidence Review. London School of Economics and Political Science, London. <https://www.lse.ac.uk/media-and-communications/assets/documents/research/projects/childrens-privacy-online/Evidence-review.pdf>.

[45] Sonia Livingstone, Mariya Stoilova, and Rishita Nandagiri. 2021. *Data and privacy literacy: the role of the school in educating children in a datafied society*. 219–236. <https://doi.org/10.5771/9783748921639-219>

[46] Gil Abraham Lopez. 2021. INVESTIGATING THE RANSOMWARE INFECTION RATE OF K12 SCHOOL DISTRICTS DURING THE COVID PANDEMIC. (2021).

[47] Jamie Manolev, Anna Sullivan, and Roger Slee. 2019. The datafication of discipline: ClassDojo, surveillance and a performative classroom culture. *Learning, Media and Technology* 44, 1 (2019), 36–51.

[48] Sana Maqsood and Sonia Chiasson. 2021. "They Think It's Totally Fine to Talk to Somebody on the Internet They Don't Know": Teachers' Perceptions and Mitigation Strategies of Tweens' Online Risks. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–17. <https://doi.org/10.1145/3411764.3445224>

[49] The Markup. 2022. Blacklight – The Markup. <https://themarkup.org/blacklight>.

[50] Jonathan R. Mayer and John C. Mitchell. 2012. Third-Party Web Tracking: Policy and Technology. In *2012 IEEE Symposium on Security and Privacy*. 413–427. <https://doi.org/10.1109/SP.2012.47>

[51] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 72:1–72:23. <https://doi.org/10.1145/3359174>

[52] Meta. 2022. Meta Pixel - Documentation. <https://developers.facebook.com/docs/meta-pixel/>.

[53] Nader Issa and Lauren FitzPatrick. 2022. Massive CPS Data Breach Exposes Records of 560,000 Students, Employees. *Chicago Sun-Times* (May 2022). <https://chicago.suntimes.com/education/2022/5/20/23132983/cps-public-schools-data-breach-students-employees-records-battelle-kids>.

[54] NCES. 2022. U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. <https://nces.ed.gov>.

[55] NewSchools Venture Fund. 2019. *Education Technology Use in Schools; Student and Educator Perspectives*. Technical Report. Gallup Inc. 64 pages. <https://www.newschools.org/wp-content/uploads/2019/09/Gallup-Ed-Tech-Use-in-Schools-2.pdf>.

[56] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.

[57] Mercy Ikhueoria Nwankwo. 2020. *IT Security Managers' Strategies for Mitigating Data Breaches in Texas School Districts*. Ph.D. Dissertation. Walden University.

[58] Maggie Oates, Yama Ahmadullah, Abigail Marsh, Chelse Swoopes, Shikun Zhang, Rebecca Balebako, and Lorrie Faith Cranor. 2018. Turtles, Locks, and Bathrooms: Understanding Mental Models of Privacy Through Illustration. *Proceedings on Privacy Enhancing Technologies* (2018). <https://petsymposium.org/poops/2018/poops-2018-0029.php>.

[59] U.S. Department of Education. 2008. Organization of U.S. Education: The School Level.

[60] Legislature of Louisiana. 2015. HB 718 (2015). <https://www.legis.la.gov/legis/ViewDocument.aspx?d=960015>.

[61] Abelard Pardo and George Siemens. 2014. Ethical and privacy principles for learning analytics. *British Journal of Educational Technology* 45, 3 (2014), 438–450.

[62] Paula Maylahn. 2022. *edTech Leadership Survey Report*. Technical Report. The Consortium for School Network. 50 pages. <https://www.cosn.org/edtech-topics/state-of-edtech-leadership/>.

[63] Lana Peterson, Cassie Scharber, Amy Thuesen, and Katie Baskin. 2020. A rapid response to COVID-19: one district's pivot from technology integration to distance learning. *Information and Learning Sciences* 121, 5/6 (Jan. 2020), 461–469. <https://doi.org/10.1108/ILS-04-2020-0131>.

[64] Alexi Pfeffer-Gillett. 2018. Peeling Back the Student Privacy Pledge. *Duke Law and Technology Review* 16, 1 (2018), 41.

[65] Paul Prinsloo and Sharon Slade. 2014. Student data privacy and institutional accountability in an age of surveillance. In *Using data to improve higher education*. Brill Sense, 195–214.

[66] Alan Rappeport. 2017. Up to 100,000 Taxpayers Compromised in Fafsa Tool Breach, I.R.S. Says. *The New York Times* (April 2017). <https://www.nytimes.com/2017/04/06/us/politics/internal-revenue-service-breach-taxpayer-data.html>.

[67] Joel R Reidenberg and Florian Schaub. 2018. Achieving big data privacy in education. *Theory and Research in Education* 16, 3 (2018), 263–279.

[68] Research and Markets. 2021. *EdTech Market: Global Industry Analysis, Trends, Market Size, and Forecasts up to 2027*. Technical Report 5401915. Infinium Global Research. 100 pages. <https://www.researchandmarkets.com/reports/5401915/edtech-market-global-industry-analysis-trends>.

[69] Jason Ribeiro. 2016. Educational technology decision-making: Technology acquisition for 746,000 Ontario students. *Canadian Journal of Educational Administration and Policy* 176 (02 2016).

[70] Stan Riley. 2022. *Independent School Districts in Texas: A Focused Ethnography on Cybersecurity Barriers*. Ph.D. Dissertation. Northcentral University.

[71] David Rosen and Aaron Santesso. 2018. School Surveillance and Privacy. In *The Palgrave International Handbook of School Discipline, Surveillance, and Social Control*. Springer, 491–507.

[72] Alan Rubel and Kyle ML Jones. 2016. Student privacy in learning analytics: An information ethics perspective. *The information society* 32, 2 (2016), 143–159.

[73] Johnny Saldaña. 2013. *The Coding Manual for Qualitative Researchers* (2nd ed ed.). SAGE, Los Angeles.

[74] Irving Seidman. 2006. *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. Teachers college press.

[75] Asuman Senol, Gunes Acar, Mathias Humbert, and Frederik Zuiderveen Borgesius. 2022. Leaky Forms: A Study of Email and Password Exfiltration Before Form Submission. In *31st USENIX Security Symposium (USENIX Security 22)*. 1813–1830.

[76] Libi Shen, Irene Chen, and Anchi Su. 2017. Cybersecurity and data breaches at schools. In *Cybersecurity breaches and issues surrounding online threat protection*. IGI Global, 144–174.

[77] Erica Shuford, Tara Kavanaugh, Brian Ralph, Ebrima Ceesay, and Paul Watters. 2018. Measuring Personal Privacy Breaches Using Third-Party Trackers. In *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*. IEEE, New York, NY, USA, 1615–1618. <https://doi.org/10.1109/TrustCom/BigDataSE.2018.00233> <https://ieeexplore.ieee.org/document/8456104/>.

[78] Natasha Singer. 2014. ClassDojo Adopts Deletion Policy for Student Data. <https://bits.blogs.nytimes.com/2014/11/18/classdojo-adopts-deletion-policy-for-student-data/>.

[79] Natasha Singer. 2014. Protecting Data Privacy at School and at Play. <https://bits.blogs.nytimes.com/2014/12/02/protecting-data-privacy-at-school-and-at-play/>.

[80] Natasha Singer. 2015. Privacy Pitfalls as Education Apps Spread Haphazardly. <https://www.nytimes.com/2015/03/12/technology/learning-apps-outstrip-school-oversight-and-student-privacy-is-among-the-risks.html>.

[81] Natasha Singer. 2015. Uncovering Security Flaws in Digital Education Products for Schoolchildren. <https://www.nytimes.com/2015/02/09/technology/uncovering-security-flaws-in-digital-education-products-for-schoolchildren.html>.

[82] Natasha Singer. 2018. For Sale: Survey Data on Millions of High School Students. *The New York Times* (July 2018). <https://www.nytimes.com/2018/07/29/business/for-sale-survey-data-on-millions-of-high-school-students.html>.

[83] Natasha Singer. 2022. A Cyberattack Illuminates the Shaky State of Student Privacy. *The New York Times* (July 2022). <https://www.nytimes.com/2022/07/31/business/student-privacy-illuminate-hack.html>.

[84] K12 SIX. 2022. K12 Security Information EXchange (K12 SIX). <https://www.k12six.org>.

[85] Sharon Slade and Paul Prinsloo. 2013. Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist* 57, 10 (2013), 1510–1529.

[86] Konstantinos Solomos, Panagiotis Ilia, Soroush Karami, Nick Nikiforakis, and Jason Polakis. 2022. The Dangers of Human Touch: Fingerprinting Browser Extensions through User Actions. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, Boston, MA, 717–733. <https://www.usenix.org/conference/usenixsecurity22/presentation/solomos>.

[87] Valerie Steeves, Priscilla Regan, and Leslie Regan Shade. 2018. Digital surveillance in the networked classroom. In *The Palgrave international handbook of school discipline, surveillance, and social control*. Springer, 445–466.

[88] Rachael Stickland and Leonie Haimson. 2019. *State Student Privacy Report Card*. Technical Report. Parent Coalition for Student Privacy. <https://secureservercdn.net/198.71.233.31/18b.b96.myftpupload.com/wp-content/uploads/2019/01/The-2019-State-Student-Privacy-Report-Card.pdf>.

[89] Faiza Tazi, Sunny Shrestha, Dan Norton, Kathryn Walsh, and Sanchari Das. 2021. Parents, Educators, & Caregivers Cybersecurity & Privacy Concerns for Remote Learning During COVID-19. In *CHI Greece 2021: 1st International Conference of the ACM Greek SIGCHI Chapter (CHI Greece 2021)*. Association for Computing Machinery, New York, NY, USA, 1–5. <https://doi.org/10.1145/3489410.3489426>.

[90] The Federal Bureau of Investigation. 2020. FBI Warns of Teleconferencing and Online Classroom Hijacking During COVID-19 Pandemic. <https://www.fbi.gov/contact-us/field-offices/boston/news/press-releases/fbi-warns-of-teleconferencing-and-online-classroom-hijacking-during-covid-19-pandemic>.

[91] J William Tucker and Amelia Vance. 2016. School Surveillance: The Consequences for Equity and Privacy. Education Leaders Report. Volume 2, No. 4. *National Association of State Boards of Education* (2016).

[92] Unknown. 2022. Lightspeed Systems: Leaders in Online Safety and Education Solutions. <https://www.lightspeedsystems.com/>.

[93] U.S. Department of Education. 2021. Family Educational Rights and Privacy Act (FERPA). <https://www2.ed.gov/print/policy/gen/guid/fpco/ferpa/index.html>.

[94] Human Rights Watch. 2022. "How Dare They Peep Into My Private Life?" Children's Rights Violations by Governments that Endorsed Online Learning During the Covid-19 Pandemic. <https://www.hrw.org/report/2022/05/25/how-dare-they-peep-my-private-life/childrens-rights-violations-governments#1436>.

[95] Martin Weller. 2018. Twenty years of Edtech. *Educause Review Online* 53, 4 (2018), 34–48.

[96] Zheng Yan, Yukang Xue, and Yaosheng Lou. 2021. Risk and Protective Factors for Intuitive and Rational Judgment of Cybersecurity Risks in a Large Sample of K-12 Students and Teachers. *Computers in Human Behavior* 121 (aug 2021), 106791. <https://doi.org/10.1016/j.chb.2021.106791> <https://www.sciencedirect.com/science/article/pii/S074756322100114X>.

[97] Jason C. Yip, Kiley Sobel, Xin Gao, Allison Marie Hishikawa, Alexis Lim, Laura Meng, Romaine Flor Ofiana, Justin Park, and Alexis Hiniker. 2019. Laughing Is Scary, but Farting Is Cute: A Conceptual Model of Children's Perspectives of Creepy Technologies. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3290605.3300303>.

[98] NBC New York. 2022. Data of More Than 800,000 NYC Public School Students Compromised in Data Hack. *NBC New York* (March 2022). <https://www.nbcnewyork.com/news/local/data-of-more-than-800000-nyc-public-school-students-compromised-in-data-hack/3616939/>.

[99] Elana Zeide. 2018. Education technology and student privacy. *The Cambridge handbook of consumer privacy* (2018), 70–84.

[100] Elana Zeide and Helen Nissenbaum. 2018. Learner privacy in MOOCs and virtual education. *Theory and Research in Education* 16, 3 (2018), 280–307.

[101] Jun Zhao, Ge Wang, Carys Dally, Petr Slovak, Julian Edbrooke-Childs, Max Van Kleek, and Nigel Shadbolt. 2019. 'I Make up a Silly Name': Understanding Children's Perception of Privacy Risks Online. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300336>.

A SUPPLEMENTARY MATERIALS

A.1 Educational Technologies Interview Guide

A.1.1 Part 1: Opening. My name is [INTERVIEWER NAME] and I'm a researcher from the University of Chicago. Thank you so much for taking the time to talk with me.

I'm part of a research project that is studying the use of educational technologies in schools, and, more specifically, the relationship between technology companies and school districts. I'd like to talk to you about your experiences working with educational technologies as a [DISTRICT OFFICIAL/ MEMBER OF THE DISTRICT IT DEPARTMENT]. It's okay if you don't know the answers to any of the questions, we're not trying to test you.

I'd like to record this Zoom call so that my team can create a transcript of our conversation. Your identity will be kept confidential, and any quotes we use will be attributed to a pseudonym. You can stop the interview at any time.

Do we have your consent to record the call?

A.1.2 Part 2: Background and Paid Educational Technologies. First, I'd like to get a bit of background about you, before asking about the educational technologies used in your district.

1. Please describe both the composition, size and your role in the district.

- What is your device situation? Are your students 1 to 1?
- Do different grade levels use different devices?
- Was this set up different before the COVID-19 pandemic?

2. What educational technologies does your district pay to use?

- Note: We define educational technologies to be any software with data flow from or about students.*
- How does your district choose which technologies to pay for? (ex.: cost? colleague recommendations? expert recommendations? security/privacy considerations?)
- In general, are technologies' privacy policies read as part of this decision making process?
- Do you know if student data protection laws such as the Family Educational Rights and Privacy Act (FERPA) are taken into account as part of this decision making process?
 - Note: FERPA is a federal law protecting the privacy of student education records, which applies to schools receiving funds from the US Department of Education. It gives parents and students certain rights over student data such as the right to inspect and correct data. It also limits who can access and use student data.*
 - More specifically, do you know if the technology companies you work with are generally designated as a school officials under FERPA?
 - Is your district part of a group or alliance (ex.: Massachusetts Student Privacy Alliance) that helps with decision making about educational technologies?
 - In general, what is included in your districts' contracts with educational technology companies?

- i. Do contracts generally specify what types of data can be collected?
 1. In particular, does your district use any technologies that collect sensitive data about students? Is this specifically allowed or restricted by contracts?
- ii. Do contracts generally specify how student data can be used?
 1. Does your district use any educational technologies that share student data with third parties? Why or why not?
- iii. Do contracts generally specify how student data is stored?
 1. Does your district have a standard for how technology companies can store student data (encryption, physical controls, etc.)?
 2. Specifically, what happens to student data after they graduate?
 3. Does your district have a standard for how long technology companies can retain student data and if so, what is that standard?
 4. Do you know the protocol of the technology companies used in your district for dealing with privacy and/or security incidents relating to student data?
- d. Is a privacy policy a deal breaker when assigning a contract with an ed tech company

A.1.3 Part 3: Teachers' Usage of Educational Technologies. Now that we've finished discussing the educational technologies used in your district, I want to learn more about how your district communicates with teachers about using those technologies.

1. **Does your district have a list of technologies that are recommended / approved for teachers?**
 - a. If yes, how are these recommendations / approvals communicated to teachers?
 - i. Is a list of recommended / approved technologies publicly available and if so can we see it?
 - b. If yes, are any of these technologies not paid for by the district?
 - c. How, if at all, do you think the process of approving a technology that the district does not pay for differ from the process for paid technologies?
 - i. Do you think privacy policies are read as part of the process of approving an unpaid technology?
 - ii. Do you think student data protection laws such as FERPA are taken into account as part of the process of approving an unpaid technology?
 - d. How, if at all, do you think your district monitors the collection, storage, and usage of student data collected by unpaid technologies?
2. **Can schools purchase educational technologies separate from the district? Why or why not?**
 - a. If yes, do they need district approval?
 - i. If yes, what is the process for obtaining approval?
3. **Does your district provide guidelines for how teachers should choose and use educational technologies to best protect the privacy and/or security of student data?**

- a. If yes, how and why do you think your district chose those specific guidelines?
- b. If yes, how are these guidelines communicated to teachers?
- c. If yes, how, if at all, do you think your district monitors whether or not teachers are adhering to these guidelines?
4. **Has any teacher or school in your district experienced a privacy or security incident relating to student data collected by educational technologies?**
 - a. If yes, tell me more about how it was handled.
 - b. If yes, what support, if any, was provided to the teacher or school involved?
 - c. If yes, what steps, if any, did the district take to prevent a similar incident from happening in the future?
 - d. If no, what would be the process to deal with a privacy incident if one were to occur?

A.1.4 Part 4: Parents' and Students' Awareness / Control Over Educational Technologies. Next, I want to learn more about how much awareness and control parents and students have over the collection, usage, and dissemination of student data by educational technologies used in your district.

1. **What do you think parent concerns surrounding ed tech are?**
2. **How, if at all, do you think parental consent is obtained for data collection done by educational technologies in your district?**
 - a. Do you think the schools in your district or technology companies are legally responsible to obtain parental consent for data collection done by those technologies?
 - i. Do you know where, if anywhere, this is typically specified? (contracts? informal agreements?)
 - b. What do you think happens if a parent does not give consent for data to be collected about their child? Why?
 - c. Do you think parents can withdraw consent about data collection? Why or why not?
 - i. If yes, do you know what the typical process is to do so?
3. **Do you think students or parents are able to limit data collection done by educational technologies used in your district?**
 - a. If yes, do you know what the typical process is to do so?
4. **Are students or parents able to view, modify, delete, or export/download student data collected by educational technologies used in your district?**
 - a. If yes, do you know what the typical process is to do so?
5. **What happens to student data after they graduate?**

A.1.5 Part 5: Closing Questions.

1. Is there anything else that you think would be helpful for us to know?
2. Is there anyone else in your district that you think we should talk to?

B PARTICIPANT SCREENER SURVEY

B.1 EdTech Study Sign-Up

This study is being conducted by the University of Chicago. To participate in the study, you will need to fill out this form, complete

a short demographic survey, and participate in an interview with member(s) of the research team in person or over video conference.

What is the study about?

- We want to learn more about what educational technologies are used in K-12 schools in the US, why those technologies are used, and how those technologies impact students.

What do I have to do to participate?

- You must first complete the form below.
- If you are selected, you will be asked to complete a survey on your demographics that should take no longer than 10 minutes
- You will then participate in a 30-45 minute interview in person or through a video conferencing platform such as Zoom.

What do I get out of it?

- You will receive a link to a \$20 Amazon gift card upon completing the questionnaire and virtual interview.
- You will have the chance to learn more about how educational technologies impact students.
- With your help, we will use develop guidelines and tools to help educators, parents, and students use educational technologies without compromising student privacy or security.

All participation is voluntary and may be stopped at any time.

To find out more about the study, you can visit:
<https://www.k12inspector.org/>

Q1. Do You Work in a K-12 Public School District in the US?

- Yes
- No

B.1.1 Personal Information. Q1. What Is Your School Email?

- Short answer text box

Q2. What Is Your Name?

- Short answer text box

Q3. What Role Do You Have In Your School District??

- IT
- Admin

B.1.2 Schedule An Interview Slot IT. We'd Like To Conduct An Interview With You! Please select either the 30 min, 45 min, or 60 min interview slot.

If you have time for a 30 min interview please click on the following link: LINK REMOVED If you have time for a 60 min interview please click on the following link: LINK REMOVED

Q1. Did you sign up for an interview slot

- I have signed up for an interview slot!

B.1.3 Schedule An Interview Slot Admin. We'd Like To Conduct An Interview With You! Please select either the 30 min, 45 min, or 60 min interview slot.

If you have time for a 30 min interview please click on the following link: LINK REMOVED If you have time for a 60 min interview please click on the following link: LINK REMOVED

Q1. Did you sign up for an interview slot

- I have signed up for an interview slot!

B.1.4 Thank You! We really appreciate you signing up for our study! The Calendly signup should have generated an email with a zoom link.

C DEMOGRAPHIC SURVEY

Q1.1
Please enter your first name

Q1.2
Please enter your last name

Q3
Do you agree to be audio-taped for the interview? ★

Yes
 No

Q4
Do you agree to be video-taped for the interview? ★

Yes
 No

Q2
★

 Skip to
End of Survey if I do NOT agree to participate... Is Selected

 By clicking "Agree" below, you confirm that you have read the consent form, are at least 18 years old, and agree to participate in the research. Please print or save a copy of this page for your records.

I agree to participate in the research
 I do NOT agree to participate in the research

Demographic Questions ...

Q5
What is your gender? ★

Female
 Male
 Non-binary
 Prefer to self-describe
 Prefer not to say

Q6
What is your age? ★

18-24 years old
 25-34 years old
 35-44 years old
 45-54 years old
 55 years or older
 Prefer not to answer

Figure 3: Our demographic survey with consent section and demographic questions.

▼ District Questions

Q7 ? *
What school district do you currently work in?

Q8 ? *
Please briefly describe your role in the district.

Q9 ? *
Approximately how many years have you been working in your current district? (If this is your first year, enter 1.)

Q10 *
Are there any specific regions or schools in your district that you work with?
 Yes
 No

Q11 ? *
Display this question
If: Are there any specific regions or schools in your district that you work with? Yes. Is Selected
Which specific regions or schools in your district do you work with?

[Import from library](#) [Add new question](#)

End of Survey

We thank you for your time spent taking this survey.

Your response has been recorded.

Figure 4: Our demographic survey with questions about school districts.

D TOP EDTECH VENDORS LINKED FROM 15,573 SCHOOL/DISTRICT WEBSITES

Most Frequently Linked EdTech Domains 1-40		Most Frequently Linked EdTech Domains 41-80		Most Frequently Linked EdTech Domains 81-110	
Domains	Counts	Domains	Counts	Domains	Counts
zoom.us	3290	newsela.com	815	abcmouse.com	473
scholastic.com	3131	remind.com	799	studyisland.com	466
collegeboard.org	3079	canva.com	764	gabbart.com	465
powerschool.com	2707	edgenuity.com	757	lexiacore5.com	461
khanacademy.org	2669	mobymax.com	751	overdrive.com	458
frontlineeducation.com	2513	quizlet.com	734	quia.com	458
act.org	2361	worldbookonline.com	734	parentsquare.com	456
smore.com	2076	infinitecampus.org	730	symbaloo.com	429
starfall.com	1973	xtramath.org	728	coolmath.com	423
clever.com	1914	soraapp.com	714	coolmath4kids.com	423
boarddocs.com	1839	edmentum.com	696	proquest.com	416
ixl.com	1804	sheppardsoftware.com	686	maxpreps.com	409
brainpop.com	1680	schoology.com	681	schoolmint.net	406
abcya.com	1594	familyid.com	669	titank12.com	403
myschoolbucks.com	1586	i-ready.com	657	teachingbooks.net	399
fastweb.com	1586	classdojo.com	650	mysteryscience.com	396
apptegy.com	1474	thrillshare.com	650	livebinders.com	395
blackboard.com	1425	eboardsolutions.com	644	prezi.com	392
schoolmessenger.com	1423	myschoolapps.com	637	illuminateed.com	389
anthology.com	1338	typingclub.com	633	flipgrid.com	387
funbrain.com	1331	boardbook.org	622	mrnussbaum.com	373
nutrislice.com	1303	seesaw.me	622	schoolnutritionandfitness.com	369
jostens.com	1276	finalsite.com	621	yearbookforever.com	368
discoveryeducation.com	1275	kidsa-z.com	613	careercruising.com	367
naviance.com	1224	tumblebooklibrary.com	606	imaginelearning.com	367
instructure.com	1125	peachjar.com	600	noodletools.com	352
renlearn.com	1119	collegeboard.com	595	libguides.com	348
getepic.com	1065	gale.com	578	explorelearning.com	336
code.org	986	typing.com	567	schoolcafe.com	335
parchment.com	939	readworks.org	548	duolingo.com	334
thinkcentral.com	927	raz-kids.com	547		
arbookfind.com	915	education.com	544		
hmhco.com	914	multiplication.com	544		
spellingcity.com	900	gonoodle.com	537		
nfhsnetwork.com	874	actstudent.org	536		
safeschools.com	870	revtrak.net	528		
classlink.com	861	pebblego.com	515		
prodigygame.com	842	hrw.com	508		
mathplayground.com	841	rschooltoday.com	498		
commonapp.org	828	nwea.org	481		

Table 7: Top 110 EdTech Domains Linked From 15,573 School/District Websites. Counts refers to the number of times a link was found across the 15,573 school/district websites. The 28 bolded domains were mentioned by our participants as being used in their school district.

E TOP NON-EDTECH DOMAINS LINKED FROM 15,573 SCHOOL/DISTRICT WEBSITES

Most Freq. Linked Not EdTech Domains 1-40		Most Freq. Linked Not EdTech Domains 41-80		Most Freq. Linked Not EdTech Domains 81-120	
Domains	Counts	Domains	Counts	Domains	Counts
google.com	10913	tinyurl.com	875	readwritethink.org	596
facebook.com	9337	bls.gov	874	suicidepreventionlifeline.org	591
youtube.com	7910	scholarships.com	866	commonsense.org	583
twitter.com	6943	bbc.co.uk	845	nctm.org	583
instagram.com	4869	sharepoint.com	813	mit.edu	574
ed.gov	4406	purdue.edu	812	readingrockets.org	573
cdc.gov	4347	usnews.com	806	edjoin.org	569
apple.com	3589	si.edu	797	fcc.gov	562
usda.gov	2932	loc.gov	797	ftc.gov	542
weebly.com	2659	boxtops4education.com	788	cnn.com	540
adobe.com	2469	ny.gov	784	nyc.gov	532
forms.gle	2445	childmind.org	778	netsmartz.org	523
vimeo.com	2139	corestandards.org	768	tasb.org	515
amazonaws.com	1958	mcgraw-hill.com	767	healthychildren.org	512
amazon.com	1907	988lifeline.org	751	qualtrics.com	509
pbskids.org	1614	wikipedia.org	739	thetrevorproject.org	506
bit.ly	1569	factmonster.com	737	mailchi.mp	506
commonsensemedia.org	1548	npr.org	726	easybib.com	504
cloudflare.com	1494	ebscohost.com	703	wisc.edu	501
linkedin.com	1463	arcgis.com	697	educationalnetworks.net	499
surveymonkey.com	1457	pta.org	692	seussville.com	490
office.com	1451	pinterest.com	689	michigan.gov	488
nationalgeographic.com	1441	yahoo.com	683	state.tx.us	487
studentaid.gov	1360	outlook.com	683	careeronestop.org	476
signupgenius.com	1298	jotform.com	675	ted.com	471
pbs.org	1235	padlet.com	675	studentscholarships.org	470
w3.org	1215	gofan.co	674	timeforkids.com	470
microsoft.com	1195	finaid.org	673	irs.gov	467
ncaa.org	1168	constantcontact.com	665	alumniclass.com	467
kidshealth.org	1157	ala.org	652	nagc.org	467
storylineonline.net	1128	pbslearningmedia.org	651	state.nj.us	465
stopbullying.gov	1118	texas.gov	640	understood.org	462
nasponline.org	1101	wordpress.org	628	square.site	461
blogspot.com	1098	paypal.com	625	cappex.com	458
nytimes.com	1058	eb.com	625	loom.com	456
ca.gov	1029	mheducation.com	619	colorincolorado.org	454
wordpress.com	963	eventbrite.com	617	nj.gov	452
nysed.gov	900	bsnsports.com	614	iscorp.com	451
nih.gov	900	samhsa.gov	609	zendesk.com	451
nasa.gov	889	edutopia.org	608	mischooldata.org	449

Table 8: Entries 1-120 of 174 Non-EdTech Domains Linked From 15,573 School/District Websites. Counts refers to the number of times a link was found across the 15,573 school/district websites.

Most Freq. Linked Not EdTech Domains 121-160		Most Freq. Linked Not EdTech Domains 141-174	
Domains	Counts	Domains	Counts
calendly.com	449	myplate.gov	353
hhs.gov	444	psychologytoday.com	352
princetonreview.com	442	affordablecollegesonline.org	347
internetessentials.com	442	healthiergeneration.org	345
crisistextline.org	442	lnks.gd	345
aap.org	440	advanc-ed.org	344
harvard.edu	438	onetonline.org	341
issuu.com	437	whitehouse.gov	341
petersons.com	436	casel.org	340
usa.gov	434	ipl.org	339
enchedlearning.com	433	nextgenscience.org	337
dol.gov	428	nuxtjs.org	337
greatschools.org	423	wonderopolis.org	336
flickr.com	421	mo.gov	335
apa.org	418		
force.com	410		
pbis.org	407		
epa.gov	404		
goodreads.com	402		
salliemae.com	400		
nami.org	398		
niche.com	389		
archives.gov	388		
washingtonpost.com	385		
lifetouch.com	384		
aaamath.com	380		
history.com	379		
citationmachine.net	376		
merriam-webster.com	374		
varsitytutors.com	372		
internet4classrooms.com	372		
stanford.edu	368		
usu.edu	366		
who.int	366		
nps.gov	364		
berkeley.edu	363		
texastransition.org	357		
ffa.org	356		
pacer.org	355		
mn.gov	354		

Table 9: Entries 121-174 of 174 Non-EdTech Domains Linked From 15,573 School/District Websites. Counts refers to the number of times a link was found across the 15,573 school/district websites.

F TOP 300 VENDORS LINKED FROM 15,573 SCHOOL/DISTRICT WEBSITES

Most Freq. Linked Domains 1-40		Most Freq. Linked Domains 41-80		Most Freq. Linked Domains 81-100	
Domains	Counts	Domains	Counts	Domains	Counts
google.com	10913	linkedin.com	1463	nasa.gov	889
facebook.com	9337	surveymonkey.com	1457	tinyurl.com	875
youtube.com	7910	office.com	1451	bls.gov	874
twitter.com	6943	nationalgeographic.com	1441	nfhsnetwork.com	874
instagram.com	4869	blackboard.com	1425	safeschools.com	870
ed.gov	4406	schoolmessenger.com	1423	scholarships.com	866
cdc.gov	4347	studentaid.gov	1360	classlink.com	861
apple.com	3589	anthology.com	1338	bbc.co.uk	845
zoom.us	3290	funbrain.com	1331	prodigygame.com	842
scholastic.com	3131	nutrislice.com	1303	mathplayground.com	841
collegeboard.org	3079	signupgenius.com	1298	commonapp.org	828
usda.gov	2932	jostens.com	1276	sharpschool.com	823
powerschool.com	2707	discoveryeducation.com	1275	newsela.com	815
khanacademy.org	2669	pbs.org	1235	sharepoint.com	813
weebly.com	2659	naviance.com	1224	purdue.edu	812
frontlineeducation.com	2513	w3.org	1215	myschoolbuilding.com	812
adobe.com	2469	microsoft.com	1195	usnews.com	806
forms.gle	2445	follettsoftware.com	1174	remind.com	799
act.org	2361	ncaa.org	1168	si.edu	797
vimeo.com	2139	kidshealth.org	1157	loc.gov	797
smore.com	2076	storylineonline.net	1128		
starfall.com	1973	instructure.com	1125		
amazonaws.com	1958	renlearn.com	1119		
clever.com	1914	stopbullying.gov	1118		
amazon.com	1907	nasponline.org	1101		
microsoftonline.com	1853	blogspot.com	1098		
boarddocs.com	1839	tedk12.com	1087		
ixl.com	1804	getepic.com	1065		
follettdestiny.com	1787	nytimes.com	1058		
renaissance-go.com	1692	ca.gov	1029		
brainpop.com	1680	code.org	986		
pbskids.org	1614	wordpress.com	963		
abcyia.com	1594	parchment.com	939		
myschoolbucks.com	1586	thinkcentral.com	927		
fastweb.com	1586	schoolwires.net	926		
bit.ly	1569	arbookfind.com	915		
commonsensemedia.org	1548	hmhco.com	914		
cloudflare.com	1494	nysed.gov	900		
apptegy.com	1474	spellingcity.com	900		
wixsite.com	1473	nih.gov	900		

Table 10: Entries 1-100 of Top 300 Domains Linked From 15,573 School/District Websites. Counts refers to the number of times a link was found across the 15,573 school/district websites.

Most Freq. Linked Domains 101-140		Most Freq. Linked Domains 141-180		Most Freq. Linked Domains 181-200	
Domains	Counts	Domains	Counts	Domains	Counts
boxtops4education.com	788	typingclub.com	633	tasb.org	515
ny.gov	784	wordpress.org	628	pebblego.com	515
childmind.org	778	paypal.com	625	healthychildren.org	512
corestandards.org	768	eb.com	625	qualtrics.com	509
mcgraw-hill.com	767	ada.gov	622	hrw.com	508
canva.com	764	boardbook.org	622	thetrevorproject.org	506
edgenuity.com	757	seesaw.me	622	mailchi.mp	506
mobymax.com	751	finalsite.com	621	easybib.com	504
988lifeline.org	751	mheducation.com	619	wisc.edu	501
wikipedia.org	739	eventbrite.com	617	educationalnetworks.net	499
factmonster.com	737	bsnsports.com	614	rschooltoday.com	498
quizlet.com	734	kidsa-z.com	613	seussville.com	490
worldbookonline.com	734	samhsa.gov	609	michigan.gov	488
infinitecampus.org	730	edutopia.org	608	state.tx.us	487
xtramath.org	728	tumblebooklibrary.com	606	nwea.org	481
npr.org	726	peachjar.com	600	mapnwea.org	476
soraapp.com	714	readwritethink.org	596	careeronestop.org	476
ebscohost.com	703	collegeboard.com	595	abcmouse.com	473
arcgis.com	697	suicidepreventionlifeline.org	591	ted.com	471
edmentum.com	696	commonsense.org	583	studentscholarships.org	470
pta.org	692	nctm.org	583		
pinterest.com	689	gale.com	578		
sheppardsoftware.com	686	mit.edu	574		
yahoo.com	683	readingrockets.org	573		
outlook.com	683	edjoin.org	569		
schoology.com	681	typing.com	567		
jotform.com	675	fcc.gov	562		
padlet.com	675	readworks.org	548		
gofan.co	674	list-manage.com	547		
finaid.org	673	raz-kids.com	547		
familyid.com	669	education.com	544		
constantcontact.com	665	multiplication.com	544		
i-ready.com	657	ftc.gov	542		
ala.org	652	cnn.com	540		
pbslearningmedia.org	651	gonoodle.com	537		
classdojo.com	650	actstudent.org	536		
thrillshare.com	650	nyc.gov	532		
eboardsolutions.com	644	revtrak.net	528		
texas.gov	640	netsmartz.org	523		
myschoolapps.com	637	itemorder.com	516		

Table 11: Entries 101-200 of Top 300 Domains Linked From 15,573 School/District Websites. Counts refers to the number of times a link was found across the 15,573 school/district websites.

Most Freq. Linked Domains 201-240		Most Freq. Linked Domains 241-280		Most Freq. Linked Domains 281-300	
Domains	Counts	Domains	Counts	Domains	Counts
timeforkids.com	470	force.com	410	mn.gov	354
irs.gov	467	maxpreps.com	409	myplate.gov	353
alumniclass.com	467	pbis.org	407	psychologytoday.com	352
nagc.org	467	schoolmint.net	406	noodletools.com	352
studyisland.com	466	epa.gov	404	libguides.com	348
gabbart.com	465	titank12.com	403	affordablecollegesonline.org	347
state.nj.us	465	goodreads.com	402	healthiergeneration.org	345
understood.org	462	salliemae.com	400	lnks.gd	345
lexiacore5.com	461	teachingbooks.net	399	advanc-ed.org	344
square.site	461	nami.org	398	onetonline.org	341
overdrive.com	458	mysteryscience.com	396	whitehouse.gov	341
quia.com	458	livebinders.com	395	casel.org	340
cappex.com	458	prezi.com	392	ipl.org	339
loom.com	456	illuminateed.com	389	nextgenscience.org	337
parentsquare.com	456	niche.com	389	nuxtjs.org	337
colorincolorado.org	454	archives.gov	388	wonderopolis.org	336
nj.gov	452	flipgrid.com	387	explorelearning.com	336
iscorp.com	451	washingtonpost.com	385	mo.gov	335
zendesk.com	451	lifetouch.com	384	schoolcafe.com	335
mischooldata.org	449	galegroup.com	384	duolingo.com	334
calendly.com	449	aaamath.com	380		
hhs.gov	444	history.com	379		
princetonreview.com	442	citationmachine.net	376		
internetessentials.com	442	merriam-webster.com	374		
crisistextline.org	442	mrnussbaum.com	373		
aap.org	440	pk12ls.com	373		
harvard.edu	438	varsitytutors.com	372		
issuu.com	437	internet4classrooms.com	372		
petersons.com	436	schoolnutritionandfitness.com	369		
usa.gov	434	yearbookforever.com	368		
enCHANTEDlearning.com	433	stanford.edu	368		
symbaloo.com	429	careercruising.com	367		
finalsite.net	428	imaginelearning.com	367		
dol.gov	428	usu.edu	366		
greatschools.org	423	who.int	366		
coolmath.com	423	nps.gov	364		
coolmath4kids.com	423	berkeley.edu	363		
flickr.com	421	texastransition.org	357		
apa.org	418	ffa.org	356		
proquest.com	416	pacer.org	355		

Table 12: Entries 201-300 of Top 300 Domains Linked From 15,573 School/District Websites. Counts refers to the number of times a link was found across the 15,573 school/district websites.