BUILDING EFFECTIVE COMMUNICATIONS AROUND STUDENT DATA PRIVACY

AN ANALYSIS OF SELECT K-12 EDTECH COMPANIES

HEINZ COLLEGE SYSTEMS SYNTHESIS | SPRING 2017

PROJECT TEAM MEMBERS
Joseph Babler
Flora Horvath
Daisy Huang
Elizabeth Martin
Mandi Prichard
Niels Smith

FACULTY ADVISOR
Andrew Richman

Carnegie Mellon University
Heinz College

Table of Contents

1. Executive Summary

As part of a capstone project of the Master of Science in Public Policy and Management (MSPPM) program at Carnegie Mellon University Heinz College, our team of six graduate students examined current practices of a select group of education technology (EdTech) startups in the K-12 space around student data privacy. Through a series of semi-structured interviews, we explored how each company develops public-facing communications regarding data use, privacy, and security policies. From analyzing the findings, the team garnered key insights and identified themes, which were used as a foundation for developing recommendations to the industry on building effective communications around student data privacy.

In Fall 2016, the Systems Synthesis Development Committee at Heinz College developed a broad problem statement for the team about student data privacy and commercialism in schools in the United States. This problem statement gave us a starting point for conducting a comprehensive literature review of over 135 articles and refining our project objective. After our initial research and consultation with our Advisory Board of subject matter experts, the team worked toward defining our project's objective, scope, timeline, and deliverables. Employing a qualitative case methodology, we examined the current practices of a select group of emerging startups in the K-12 space, including how they develop public-facing communications regarding their data use, privacy, and security policies. Our intent was to find best practices regarding how emerging EdTech companies relay data privacy practices to the public and achieve meaningful transparency with stakeholders.

Using a merged database of companies that combined startup information from multiple sources, our team filtered 450 known EdTech startups down to a list of 120 EdTech companies and then selected 18 finalists based on criteria such as student data privacy risk, staff size, reputation, revenue growth, customer base, and value proposition. After a few weeks of standardized communications outreach and recruiting, six companies ultimately agreed to participate in our project, sitting for one to three hour-long interviews about their privacy practices and communications. After interviewing the companies, our team worked together to distill high-level commonalities across interviews, from which we gained the following key insights:

- Beyond complying with federal and state-level requirements, EdTech companies do not prioritize student data protections, as compared to customer acquisition and product development in their first five years.
- Due to factors such as limited resources and little demand from customers, EdTech companies do not establish formal strategies around public-facing communications about student data privacy for external stakeholders.
- Most EdTech companies use an open source, standardized privacy policy as a foundation for informing users about student data practices, which is customized as they scale up. Common practices include borrowing and/or adapting sections from competitors' privacy policies, as well as adding in sections based on customer demand and changes in federal or state-level requirements.

- Concerns about complying with privacy regulation and guidance do not seem to inhibit innovation at EdTech companies.

Although we garnered these insights from our data, the team believes that EdTech companies would also benefit from other recommendations. While most of the following recommendations were formed in the context of our small sample size of companies, they can inform and apply to emerging companies that come into the EdTech space.

- In a rapidly evolving industry landscape, the process of improving privacy practices and communications in EdTech companies should be dynamic, as opposed to one-time or periodic.
- As EdTech companies scale up, they should consider encouraging a shared responsibility for staying vigilant about changing technical standards across team members, instead of assigning this responsibility to one or two staff members. This will help increase engagement and awareness for data security practices across the company, as well as help instill a culture of respect for sensitive student data in staff.
- Instead of taking a piecemeal approach to developing privacy practices over time, young companies should consider building front-end processes and standards that guarantee evolutionary flexibility downstream.
- EdTech companies should resist collecting or storing unnecessary student data and establish strong internal controls to preclude doing so.
- EdTech companies should consider using strong and proactive public-facing communications about student data privacy as a product differentiator to stand out among competitors. Nimble adoption and understanding of new guidelines, technologies, and best practices can give initial adopters an important competitive advantage in the market.

Further, these cases offer important lessons for others within the EdTech ecosystem:

- For investors:
    - A theme emerged among these interviews about a perceived lack of meaningful interest about student data privacy from investors. Our team views this as a missed opportunity for everyone involved, and recommend that investors consider systematically assisting companies in ensuring strong privacy practices.
- For school districts:
    - School districts are the greatest force for our companies to change their privacy behaviors. Our team believes that more research needs to be done assessing school district capabilities across the U.S. for examining the technology that comes into their schools.

Our team submits that implementing these recommendations will benefit EdTech companies, their customers, and district and school-level stakeholders in this space. By employing a more proactive strategy to student data protection from day one, EdTech companies can position their vigilance as a key differentiator for their product, capture a broader market share, and share in the responsibility for protecting sensitive student data.

2. Introduction

2.1. Project Objective and Report Overview

The scope of our Systems Synthesis Project encompassed student data privacy practices of education technology (EdTech) companies that provide products, apps, or services for use by stakeholders in K-12 education. Our team interviewed six companies using a qualitative case study methodology that focused on public-facing communications on data use, privacy, and security practices. This group of companies includes a range of business models, from institutional to freemium (both for the institution and the consumer) as well as differing approaches to achieving user and revenue growth and articulating their value proposition. We synthesized the qualitative data from these case studies, distilling commonalities and key themes, and identified actionable recommendations to EdTech companies. Ultimately, our project's objective is to help nascent EdTech startups avoid communications missteps, establish trust with current and prospective customers, and build a dialogue about privacy practices between vendors and schools, districts, and other customers. This report presents the background and strategic context of student data privacy in the United States, discusses our research and data collection processes, presents our case studies, synthesizes our findings, and notes study limitations.

2.2. Advisory Board

Throughout our project's lifecycle, our team benefitted from the expertise of the following list of professionals and advocates from government, industry, nonprofits, and universities:

- ❖ **Rachel Anderson,** *Associate Director of Federal Policy and Advocacy at Data Quality Campaign (DQC)*
- ❖ **Mark Luetzelschwab,** *President of Eduphoric*
- ❖ **Jim Siegl,** *Technical Architect, Fairfax County Public Schools*
- ❖ **Amelia Vance,** *Policy Counsel, Future Privacy Forum (FPF)*
- ❖ **Elana Zaide,** *Associate Research Scholar at Princeton University's Center for Information Technology*

2.3. Project Methodology and Timeline

Our team began work on this project in Fall 2016 when the Systems Synthesis Development Committee at Carnegie Mellon University Heinz College developed a broad problem statement for our team. This problem statement gave us a starting point as guidance for reviewing relevant, evidence-based research and refining our project objective. We include the original problem statement prompt as Appendix 1 of this report.

To prepare, our team undertook a rapid literature review of available resources to familiarize ourselves with student data policy and the EdTech industry landscape.

We derived our research from a variety of source materials, including industry pledges and commitments driven by advocacy groups, state legislative activity (particularly concentrated in 2014-2015), and scholarly articles from EdTech industry trade journals.

In our review, we quickly identified a disconnect in communication between EdTech companies and their stakeholders (districts, parents, teachers, and students) regarding the collection, maintenance, and use sensitive student data. Exacerbating factors involve unevenness in data literacy among stakeholders, an outdated legislative framework, and series of headline-making scandals exposing EdTech companies misusing student data. In general, we found that many educators and parents are increasingly mistrustful how EdTech companies amass and use data to track children, as well as commercialize students' personal information.

Figure 1 denotes a high-level timeline that our team followed for this project's thirteen-week life cycle, during which we presented three times to our advisory board. It identifies key meetings, research, and dissemination activities, as well presentations to our advisory board members and policy community.

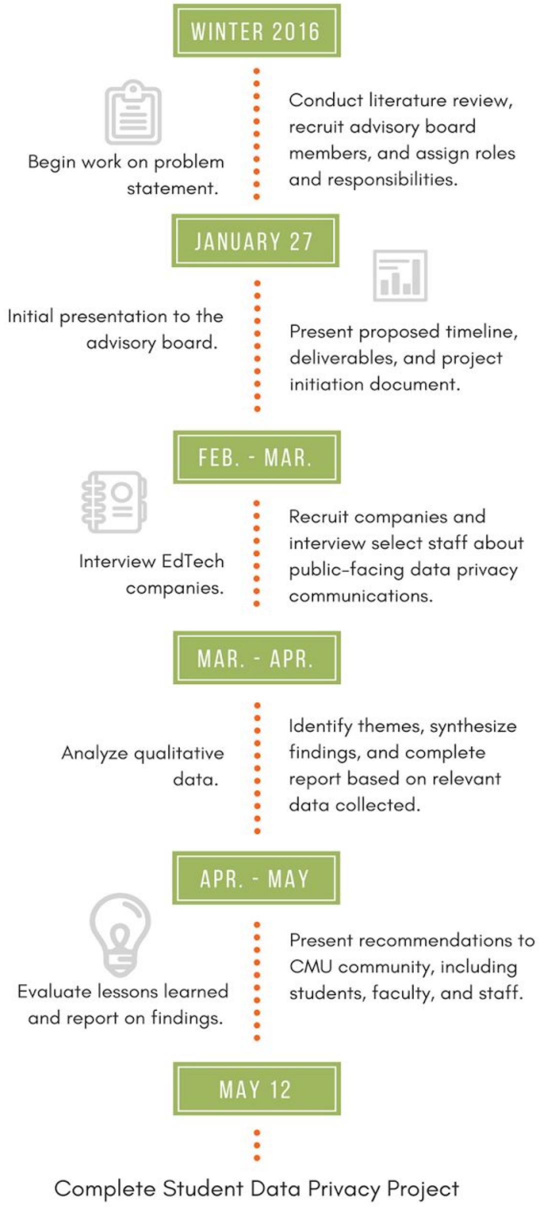

WINTER 2016
Begin work on problem statement.
Conduct literature review, recruit advisory board members, and assign roles and responsibilities.
JANUARY 27
Initial presentation to the advisory board.
Present proposed timeline, deliverables, and project initiation document.
FEB. - MAR.
Interview EdTech companies.
Recruit companies and interview select staff about public-facing data privacy communications.
MAR. - APR.
Analyze qualitative data.
Identify themes, synthesize findings, and complete report based on relevant data collected.
APR. - MAY
Evaluate lessons learned and report on findings.
Present recommendations to CMU community, including students, faculty, and staff.
MAY 12
Complete Student Data Privacy Project

Figure 1. Project Timeline

of

a

of

as

3. Strategic Context

3.1. History of Student Data Privacy

Before the mid-1970s, there was little regulation in the United States around schools sharing a child's information and records with third parties, with or without parental consent. Educational institutions could theoretically share records like enrollment forms, report cards, and disciplinary write-ups with the government, police, or vendors while also blocking parents from accessing those records.[1] At the state and federal level, education departments typically assigned a limited staff to generate basic statistical reports on basic data points, such as student enrollment,

[1] John Jennings. "1974: A Brief History of Student Data Privacy." Advancing K-12 EdTech Blog. https://www.skyward.com/discover/blog/skyward-blogs/skyward-executive-blog/february-2016/navigating-ferpa-and-protecting-student-data

teacher characteristics, or education expenditures.[2] Teachers, administrators, and parents were able to effectively protect student records, as any data storage on servers was owned by school districts and stayed within a defined physical jurisdiction. As student record management has moved from paper-based systems to leveraging digital solutions, administrators at the district and school level have started collecting student data through a combination of the two methods. Big data capabilities have also emerged as a promising new method for collecting, analyzing, and communicating information about schools, teachers, and students within and across education management systems.[3] However, questions about the security, accessibility, and sharing of student data continue to concern IT leaders, with 64 percent of a recent leadership survey stating that it is a more important issue in 2016 than the previous year.[4]

3.2. Current Strategic Context

Fast forward to 2017, however, and big data capabilities in the education field have vastly expanded, with new digital learning tools utilizing algorithmic systems on thousands of comprehensive and varied indicators on their platforms.[5] Along with government agencies and nonprofit organizations, EdTech companies (e.g., Google Apps for Education) and consulting enterprises (e.g., Pearson and Mckinsey) have seized on business opportunities in analytics and associated technologies.[6] Through collection activities and commercial transactions on servers located outside of the district's physical jurisdiction, educators, researchers, and even parents have unwittingly released sensitive student data, such as students' *de facto* personally identifiable information (PII). With the advent of Software as a Service (Saas) based tools, schools are rapidly shifting to new technologies, such as cloud computing, to find off-premise data solutions that require significantly less up-front costs but grant EdTech vendors far more access to student data.[7] An evolving "surveillance culture" has emerged, in which policymakers must carefully balance the sensitive intersection of education data and student privacy.[8] The fact that most information belongs to or relates to children only heightens the issues and sensitivities. Parents, advocacy groups, and other stakeholders in the education sphere share concerns over storing private student data on the internet, as third-party companies participate in data sharing processes, such as data mining, and sensitive student

[2] Center for Digital Education. (2013) "Big Data and Analytics in K-12 Education: the time is right." *Houghton Mifflin Harcourt Center for Digital Education.* Retrieved from http://www.hmhco.com/~/media/sites/home/Teachers/Files/HMH-CDE_Issue%20Brief_DataAnalytics.pdf

[3] Bradley Shear. (2015). "Ed Tech Must Embrace Stronger Student Privacy Data Laws." *T H E Journal.* Retrieved from https://thejournal.com/Articles/2015/05/28/Ed-Tech-Must-Embrace-Stronger-Student-Privacy-Laws.aspx

[4] CoSN K-12 IT Leadership Survey. (2016). *Consortium for School Networking.* http://www.cosn.org/sites/default/files/CoSN%20K-12%20IT%20Leadership%20Survey%20Report%202016.pdf.

[5] Obama White House Archives. (2016). "Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights." Retrieved from https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf

[6] Ibid.

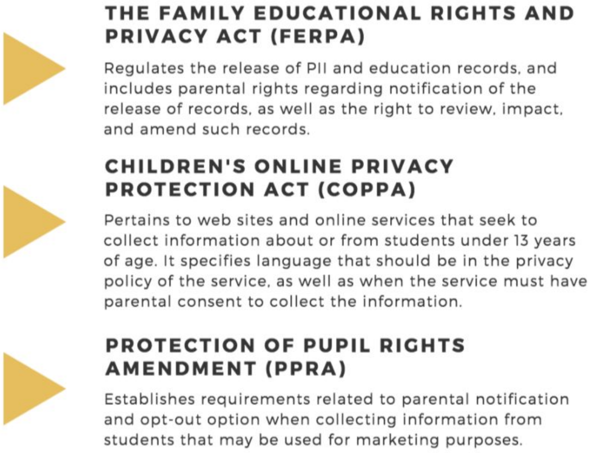
[7] The Berkman Center for Internet & Society. (2013). "Working Roadmap: Student Privacy in the Cloud Computing Ecosystem." *The Berkman Center for Internet & Society at Harvard University.* Retrieved from https://cyber.harvard.edu/sites/cyber.law.harvard.edu/files/Berkman_Center_Student_Privacy_Working%20_Roadmap-June_2013.pdf

[8] Faith Boninger and Alex Molnar. (2016). "Learning to Be Watched: Surveillance Culture at School. *National Education Policy Center at the University of Colorado, Boulder.* Retrieved from http://nepc.colorado.edu/files/publications/RB%20Boninger-Molnar%20Trends.pdf

information, such as disciplinary records and disability information can fall into the wrong hands.

3.3. Legislative Landscape

In part, student data privacy has evolved into such a hot-button policy issue due to inadequacies in the legislative landscape. Simply put, there is no timely and comprehensive federal student privacy law. State legislators, school administrators, EdTech companies, and third-party vendors have struggled to interpret the Department of Education's role, which has led to varied and state-specific interpretations of the governance framework.[9] The Family Educational Rights and Privacy Act (FERPA) is the federal law that defines how and when educational institutions can disclose PII about students from their education records.[10] Signed into law by President Ford in 1974, FERPA provides outdated and scant guidance for protecting student privacy in 2017. For example, the legislation vaguely stipulates that only those with a "legitimate educational interest" should ever access a student's records, which school districts are left to interpret and enact as they see fit.[11] A complementary federal statute that governs schools on how they track, collect, and use student data is the Protection of Pupil Rights Amendment (PPRA). Like FERPA, it was signed into law in 1974, originally as a right of parent access to federally funded experimental materials, and was amended in 1978 to add parental consent for any educational institution collecting sensitive information from it students.[12] In 2002, Congress significantly expanded PPRA to limit schools' ability to both collect certain sensitive information from students, as well as disclose it for commercial purposes.[13]



THE FAMILY EDUCATIONAL RIGHTS AND PRIVACY ACT (FERPA)

Regulates the release of PII and education records, and includes parental rights regarding notification of the release of records, as well as the right to review, impact, and amend such records.

CHILDREN'S ONLINE PRIVACY PROTECTION ACT (COPPA)

Pertains to web sites and online services that seek to collect information about or from students under 13 years of age. It specifies language that should be in the privacy policy of the service, as well as when the service must have parental consent to collect the information.

PROTECTION OF PUPIL RIGHTS AMENDMENT (PPRA)

Establishes requirements related to parental notification and opt-out option when collecting information from students that may be used for marketing purposes.

Figure 2. Guide to existing legislation

Another important piece of legislation surrounding this issue is the Children's Online Privacy Protection Act (COPPA), which applies to children under the age of 13 and requires EdTech companies to obtain parental consent before obtaining PII from children for commercial purposes.[14] In 2012, the Federal Trade Commission (FTC) helped bolster several elements of COPPA, including provisions for regulating new tracking technology.

[9] David Raths. (2016). "The Patchwork of State Student Privacy Laws." *T H E Journal*. Retrieved from https://thejournal.com/articles/2016/10/13/the-patchwork-of-state-student-privacy-laws.aspx

[10] Boninger and Molnar, 15.

[11] Ibid.

[12]i Lynn M. Daggett. (2008) "Student Privacy and the Protection of Pupil RIghts Act as Amended by No Child Left Behind." *UC Davis Journal of Juvenile Law and Policy*. https://jjlp.law.ucdavis.edu/archives/vol-12-no-1/daggett.pdf

[13] Ibid, 56.

[14] "Talking about the Facts of Education Data with School Board Members." *National School Boards Education Center for Public Education and the Data Quality Campaign.*

[15] One significant criticism of COPPA is its limited scope for protecting children under the age of 13, which can leave minors aged 13 to 17 with insufficient legal protection in their online activity.[16]

Congress has attempted several times to overhaul these legislation and address student data privacy, most recently in 2015 when lawmakers introduced the Student Digital Privacy and Parental Rights Act that was ultimately not signed into law. In 2016, federal policymakers shifted toward enacting the newly reauthorized No Child Left Behind Act, known as the Every Student Succeeds Act.[17] In lieu of Congress introducing new legislation, the federal government has responded to state requests for clarification on FERPA guidance, e.g., in 2008 and 2011, when it released regulations defining the role of the state in using student data while maintaining privacy protections around PII.[18] These regulations likewise established specific provisions, including founding the Privacy Technical Assistance Center at the Department of Education, the creation of the Chief Privacy Officer position, and the implementation of penalties for certain kinds of privacy breaches.[19]

3.4. State Legislative Activity and Challenges for EdTech Companies

As a result of federal privacy laws lagging behind the technology, state legislatures have passed a number of bills intended to regulate student data, particularly gaining momentum in legislative activity in 2014 and 2015. In 2014, states introduced and passed privacy bills focusing on state-level data collection and governing the federal government's responsibility, access, and oversight in this process.[20] In 2015, legislative activity shifted toward defining the data use and privacy activities of online service providers and vendors, as well as passing laws on supporting district's resource needs and staffing for student data privacy As a recent report from the Data Quality Campaign describes, "student data privacy bills adopt two main approaches: protecting privacy by limiting data use (a prohibitive approach) and protecting privacy by implementing data governance (a governance approach)."[21] In California, the Student Online Private Information Protection Act (SOPIPA), introduced in 2014 and enacted in 2016, is widely regarded as an effective, current, and all-encompassing bill for protecting K-12 student

[15] Ibid.

[16] Caitlin R. Costello, Dale E. McNiel, Renee L. Binder. (2016). "Adolescents and Social Media: Privacy, Brain Development and the Law. Journal of the American Academy of Psychiatry and the Law Online. 44 (3), 313-321. Retrieved from jaapl.org/content/44/3/313

[17] Data Quality Campaign Reports. (2016). "Student Data Privacy Legislation: A Summary of 2016 State Legislation." http://2pido73em67o3eytaq1cp8au.wpengine.netdna-cdn.com/wp-content/uploads/2016/09/DQC-Legislative-summary-09302016.pdf

[18] (n.d). "Talking about the Facts of Education Data with School Board Members." *National School Boards Education Center for Public Education and the Data Quality Campaign.* Retrieved from http://www.centerforpubliceducation.org/Main-Menu/Policies/Data-Privacy-Fact-Sheets-PDF.pdf

[19] Ibid, 16.

[20] Data Quality Campaign Reports. (2015). "Student Data Privacy Legislation: What Happened in 2015, and What Is Next." http://2pido73em67o3eytaq1cp8au.wpengine.netdna-cdn.com/wp-content/uploads/2016/03/DQC-Student-Data-Laws-2015-Sept23.pdf

[21] Data Quality Campaign Reports. (2016). "Student Data Privacy Legislation: A Summary of 2016 State Legislation."

data.[22] Several states are seeking to replicate their own legislation on SOPIPA for regulating collection of K-12 student on websites, applications, and other technologies.[23]

EdTech companies have also proactively taken several steps to advocate for stronger student privacy protections, such as signing industry pledges and offering transparent privacy policies that protect children's personal information. While more robust privacy laws are expected in the near future, EdTech companies can, in the meantime, do more to address public concern about student data breaches or misuse, such as selling information for marketing purposes. In the absence of legal repercussions for such misuse, EdTech companies have provided robust privacy policies and adhered to best practices regarding parental consent and notification.

4. Literature Review Synthesis

4.1. Approach

Of the 135 articles our team read in the literature review, the majority focused on aspects of data privacy that were not directly relevant to the scope of our project. Overall, most of the results of our literature search focused on implications of FERPA, COPPA, PPRA, as well as other regulations. Much of the peer-reviewed work involved analysis and use of educational data to improve student performance through evidence-based interventions. Through our rating system described in appendix 5 of this report, we narrowed the results to 13 articles that provided the most useful background and concerns about successful data privacy practices and communicating data privacy to stakeholders in the K-12 education environment. This narrow number of applicable articles to our research is evidence of the significant amount of work that remains to be done in this field.

4.2. Results

The selected articles fall into two categories: (1) those which discuss the best practices for protecting student data and maintaining data privacy, and (2) those which discuss successful communications practices on data privacy. Articles focusing on maintaining data privacy consistently suggest ensuring all employees who handle data be trained on use and privacy and that all organizations who use student data in any capacity are deeply knowledgeable about the applicable laws and regulations.

This literature review raised several salient concerns. First, there is no clearly established division of responsibility between schools and EdTech companies in regards to maintaining student data privacy. For example, one article discussed the ethics of the terms of service agreement that educators are forced to agree to when using some EdTech products and the limited bargaining power for privacy and security concerns that educators have with free apps or apps that do not directly contract with districts.[24] Another discussed the shift from school IT organizations to protect students to teachers, especially as teachers increasingly adopt digital devices, platforms, and apps in the

[22] Bradley Shear, "Ed Tech Must Embrace Stronger Student Privacy Data Laws."
[23] Ibid.
[24] Debbie Abilock, Rigele Abilock. "I Agree but Do I Know?: Privacy and Student Data". (2016) American Library Association

classroom.[25] A number of articles published focus on the roles of school districts, school administrators, and teachers, but fewer cover the responsibilities and best practices of EdTech companies in the education sphere. This suggested to us a lack of synthesized awareness in the academic community to the practices of the EdTech sphere writ large.

Notably, and perhaps connected to the relatively minor coverage of EdTech company practices, several articles suggest a distrust of EdTech companies by researchers and school leaders, centered on the potential tension between ethical treatment of student data and the need for revenue generation to sustain the company. According to articles covering controversy around technology companies with privacy, researchers, parents, trade groups, and school administrators are primarily centered on the appropriate collection of data, sale of student data to third party groups, particularly for marketing purposes, and the potential for data breaches revealing sensitive student data.

The difficulty of communicating privacy practices is a known issue. One article specifically cited the closing of inBloom as the failure to communicate their data security and privacy protections in a way that convinced stakeholders of their commitment and effectiveness.[26] Similar issues cropped up for Google when a lawsuit was filed alleging that the company had violated federal and state laws by scanning and indexing emails of apps for education users.[27] There is a need to balancing transparency without overwhelming stakeholders--particularly parents and teachers--and there is a lack of guidance on how to do this successfully.

5. Methodologies

5.1. Literature Review

Working with a reference librarian at the Library of Congress, our team conducted a comprehensive literature review of periodicals, journals, and other written sources on educational technology and privacy written in the last five years. This involved pulling over 700 possible abstracts from multiple databases, including *Education Resources Information Center (ERIC)*, *Education Full Text* database, the *ABI/Inform Collection*, and *ProQuest Education*.

From this first pull, we further reviewed the list and ended with a narrowed list of 135 sources, an index of which can be found in appendix 5. The research team read and reviewed these 135 sources and an assigned team member graded each source using a pre-established quality scoring tool. The team generated the tool in an effort to identify the sources were most relevant to our research. Figure 3 gives an overview of our literature review process.

[25] Dian Schaffhauser. “The hunt for data privacy in the classroom”. (2016). THE Journal Magazine.
[26] Ben Kamisar. “InBloom Sputters Amid Concerns About Privacy of Student Data”. (2014) Education Week.
[27] Benjamin Herold. “Google Under Fire for Data Mining Student Email Addresses”. (2014) Education Week.

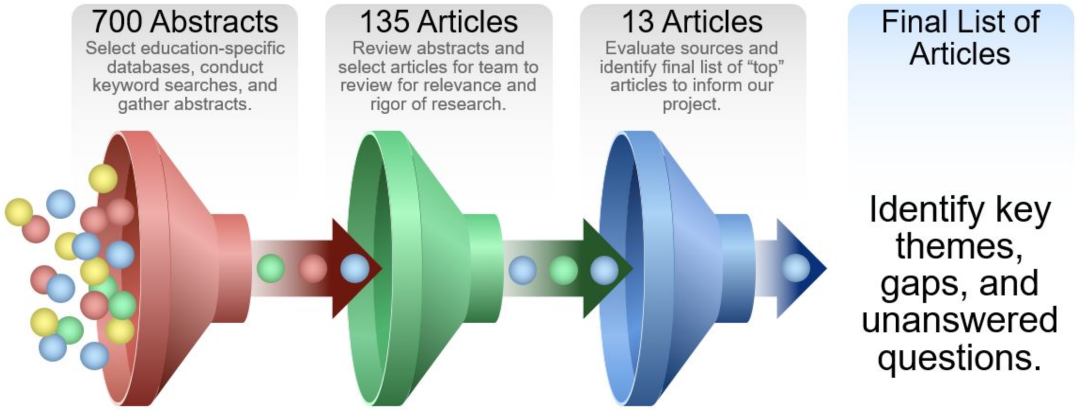

700 Abstracts
Select education-specific databases, conduct keyword searches, and gather abstracts.
135 Articles
Review abstracts and select articles for team to review for relevance and rigor of research.
13 Articles
Evaluate sources and identify final list of "top" articles to inform our project.
Final List of Articles
Identify key themes, gaps, and unanswered questions.

Figure 3. Literature Review Process

5.2. Case Selection

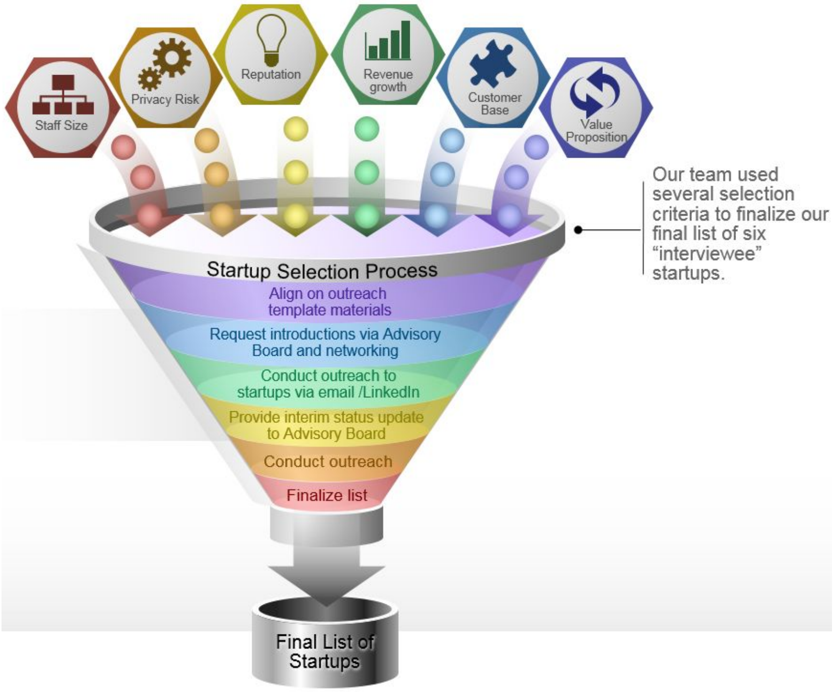
After reading and discussing the gathered literature and conducting a series of informational interviews, we decided on a qualitative case study research method. These cases were intended to serve as the bulk of our primary research and reported findings.

To identify potential case companies, our team first defined our area of research as educational technology startups in the K-12 space that were old enough to have a meaningful customer base but still small and in the process of addressing privacy issues (generally less than 30 staff). We gathered a large list of 450 possible companies from multiple startup databases and removed any companies that did not fit our criteria. This process resulted in 120 potential companies that were individually analyzed by the research team along a range of criteria (see next section). The 120 companies were put into several tiers depending on how well they fit with our qualitative research goals, with the top tier comprised of the 18 startups that aligned with our research targets and goals.

5.3. Research Methodology

Once we had a target list of startups, we began reaching out to them through direct emails, social media (such as LinkedIn), and personal connections when possible. We explicitly offered to share our findings with the companies contacted as a means of incentivizing participation. After establishing first contact, we had brief (15-30 minute) introductory calls with various employees from the startup to explain our research goals and discuss their participation. Many startups did not respond at all to our inquiries, while a few decided not to participate once they spoke with us.

We conducted one-hour interviews with one to three staff members at each company that agreed to participate, along with examining and documenting any public records of the company. These interviews were conducted primarily over the phone with two members of our six-person research team. Extensive notes on the interviews were taken and results were discussed internally amongst the research. Figure 3 below shows our team's selection criteria and roadmap for finalizing the list of EdTech startups we interviewed.


Staff Size
Privacy Risk
Reputation
Revenue growth
Customer Base
Value Proposition
Startup Selection Process
Align on outreach template materials
Request introductions via Advisory Board and networking
Conduct outreach to startups via email /LinkedIn
Provide interim status update to Advisory Board
Conduct outreach
Finalize list
Final List of Startups

Our team used several selection criteria to finalize our final list of six “interviewee” startups.

Figure 4. Process for Selecting EdTech Startups

6. Cases

In general, these six companies were from all over the United States, with one international company who had a presence in the US market. They generally had no more than thirty staff. The ages of the companies varied but generally did not exceed eight years. Some of the products offered by these companies were used in a classroom, while others were used by the administration to somehow manage the school. They all had at least some clients, though some had been fully deployed for less than a year while others had been in business for multiple years.

To ensure the absolute protection of the companies who agreed to participate in our research, we have excluded write-ups of the individual case companies. This will protect their participation in our research and allowed these companies to speak candidly with us. The more robust, aggregated results appear in our findings and recommendations.

7. Recommendations & Conclusions

7.1. Conclusions

There is much about which to be optimistic. Educational technology is revamping schools and classrooms, changing the tools teachers have and the way students engage with learning. As a research team, we had the opportunity to explore just how dramatically and rapidly the landscape is changing, and we are all hopeful for the kinds of innovations it will yield. There are much larger issues of privacy and technology that exist well outside of the classroom concerning how we use technology and who gets access to our information. No special solution to these problems exist for the realm of education.

Broader debates about civil liberties will continue to have an impact on what kinds of tools end up in our classrooms. In the meantime though, work remains to be done on how privacy and technology intersect with our students. The companies we spoke with were all happy to follow guidelines, meet technical standards, and do what was asked of them to protect their customers' data. In that vein, below are our recommendations for educational technology companies that may just be starting out, investors, and further researchers on what *can* be done in the current technology environment.

7.2. EdTech Startups

1. **View privacy practices as evolving and constantly improving:** Policymakers have implemented a variety of new legislation and regulations regarding privacy within the last five years; it is therefore important that company practices reflect these changes and acknowledge that their privacy practices may have to change as laws do. Additionally, security and privacy technologies and best practices are evolving, which will benefit companies who are able to nimbly and effectively adapt.

2. **Keep track of changing technical standards:** Some of our companies were able to regularly keep up with changes in the industry to their (and everyone's) benefit. Sometimes this was done through an educational center, trade group, or other group of like-minded interests. Investors who focused particularly on the education technology market were good sources of changing standards and might prove a useful source for new companies. Certain CTOs we spoke with had deep experience in the industry and already had methods of tracking both policy and technical privacy standards as they came to light. However it's managed, data security should be understood as a constant requirement.

3. **Set good standards:** The majority of our startups made the major privacy decisions as they initially formulated their product. Most privacy decisions after the initial planning stages were limited and largely the result of outside feedback. This suggests that new companies--or companies launching new products--should think most critically and intensely about privacy during the initial phase. (For example, one of our case studies found part of its interest to schools was its ability to limit the collection of sensitive data relative to competitors.) This is not just a technical issue, but a policy issue of what kinds of information a company captures, who can manipulate that information, and where that information lives.

4. **Do not collect or store unnecessary data:** This might not be as obvious as it seems. Several of the case companies we spoke with expressed the risks involved in collecting more data, particularly PII, than absolutely necessary for their product to function. A best practice is to be strategic and limited in data collection to minimize risk to the company and consumer. This involves drilling down to every field on a form, login for a product, or download for an app. Companies expose themselves to liability with each additional collected piece of information; new startups or product lines should gather the bare essentials to make a product work. Additionally, setting clear standards around data storage

and deletion early on will create confidence in clients and easy standards to follow for staff.

5. **Use privacy to differentiate:** Consumers generally are more aware of issues with their data security and the field of education is no different. One possible strategy for new EdTech companies is to sell not only their product but also their ability to keep limit data collection and increase security relative to their peers. Leading with good privacy standards as a marketable trait would be good for the underlying product and could help new companies attract more business relative to their peers.

7.3. Investors

Every company we spoke with discussed a general lack of pressure or interest from their capital investors in data privacy. Companies described investors as simply “checking-the-box” when it came to student data privacy. After an investor had made a decision to invest in one of our companies, the investor would do its basic due diligence and ensure that the startup did not run seriously afoul of any federal or state privacy standards, as well as being up-to-date on basic technical privacy standards. Because our companies are distributed across the United States and inhabit different parts of the EdTech space, this finding is likely particularly robust.

This lack of interest in privacy on the part of investors is a wasted opportunity for everyone involved. As our research found, new EdTech companies are particularly likely to need external support to navigate complicated and changing student privacy standards. Investors are uniquely exposed to these risks. The increased marginal cost to investors to offer technical and strategic advice to their investments is low relative to the benefit. EdTech investors can feel increased confidence that their investment is protected, while new startups can get much needed help in sufficiently prioritizing student data privacy.

7.4. More Research: School Districts

The moment when newer EdTech companies are most inclined to make significant technical or policy changes to their products to protect student data is when they are selling their product to a school or school district. The companies we spoke with generally described themselves as happy to meet whatever standards potential paying clients requested. And no other point in the development chain demanded or resulted in as much change at our case studies as when tried to get a new client.

There has been a flurry of lawmaking in the last decade at the state level; less discussed is how those state level laws will be met and enforced. It is an open question whether or not schools and school districts are uniformly well-equipped to be the arbiter of student data privacy standards for the whole of society. We suspect they are not. Given how widely resources can vary between schools within the same city, let alone the same state, our concern is that many schools lack the staff and other resources needed to effectively evaluate new technologies. New EdTech companies, unwittingly or not, could avoid the scrutiny that our case studies seemed genuinely happy to have.

However, these conclusions are beyond the scope of this effort. We encourage further work to be done into how successful and capable schools and school districts are at acting as the gatekeepers for their students' sensitive data.

8. Sources of Error and Limitations

8.1. Overview

Whenever possible, we tried to limit possible sources of error and ensure our cases were not an overly biased sample. However, there are several possible errors on the results we did produce and a variety of limitations on the reach of our conclusions. These are detailed below.

8.2. Sources of Error

While initially reaching out to companies, not every company responded or agreed to participate that we had initially chosen. One company said it was busy with a "product launch" while another cited an upcoming fundraising round. While most companies were ranked high on our scoring system (see appendix 6), it's possible that companies with serious difficulties or problems related to their privacy communications or standards simply opted themselves out of our review. This would skew the case studies we did toward companies that had not experienced problems, had better technological privacy solutions in place, and were generally more comfortable sharing this information with researchers. This might make our final conclusions seem more optimistic than they otherwise would be.

To ensure the privacy of the companies we interviewed and due to resources constraints, we did not record and transcribe our interviews. This fact prevented us from using well-known qualitative coding techniques to limit interview bias and interpretation in final results. To compensate for this we took generous notes during each interview, talked regularly across research teams to facilitate a common understanding, and collaborated closely on writing our final results.

Our interview questions were necessarily iterative and changed from company to company and interview to interview. While we collectively agreed on a mandatory set of basic questions--each two-person research team had the same original template to follow--follow-up questions, clarifications during the interview, and the natural course of conversation likely changed the tone and emphasis from interview to interview. Because our goal was not to generate final, authoritative results but instead get a general understanding of a wide swath of the educational technology field, this methodological choice made the most sense. However, it possibly introduced subtle but meaningful bias into the interviews and conclusions.

8.3. Limitations

Our use of a case interview methodology necessarily limited our findings. Because of the constraints on time and budget, generating, conducting, and analyzing a survey was not feasible. While we consciously wrote our conclusions to be suggestive rather than authoritative, we still may have over-generalized the lessons from our few cases to the broader field. We do not consider this research exhaustive, and our primary aim was discovery, hoping to cover ground that had not been well analyzed previously.

Our cases focused on a reasonably narrow developmental range--startups that were old enough to have some customers but not so old as to be very large in terms of staff size. This was a strategic choice of our research; we assumed companies any newer would not have enough to share and companies much larger would be generally less willing to speak with us in a candid way. Our hope in selecting this developmental range was to access companies that had already iterated some on their privacy communications and standards but likely were still contending with how to best move forward. This choice limits the external validity of our findings.

Our primary data--the case interviews with each company--depended on interviewees to be forward and honest about their companies and their issues with privacy practices. While we supplemented these interviews with public research, many of these companies were small enough or new enough to have little in terms of a public record. Ultimately, the extent which companies were willing share their information and struggles limited the extent of our research. Our team did not have the technical ability to do primary research into a company's inner workings, nor did we believe any companies would volunteer this type of access.

Appendix 1. Original Problem Statement

Prompt: Commercialism in the Schools & Student Data Collection

High-level Issue Definition: Big Data capabilities are revolutionizing marketing, which has become a billion dollar revenue stream for tech companies and marketing firms. Meanwhile, elementary and high schools in the U.S. are facing serious funding and accessibility to technology challenges in teaching their students. An evolving “surveillance culture” is developing in which schools are faced with a balancing act of assessing technology and allowing student data to be collected and commercially used, including de facto personally identifiable information (PII). Tech companies are opposing additional legislation to protect student data and many school districts, parents, other stakeholders are unaware of the extent their children’s data is being used for marketing purposes to the children themselves. Should, and can, there be additional protections for young students? Can enforceable legislation be designed to protect students, and what information is needed? Should technology incentives that furnish devices for classroom and student use be regulated to protect students from marketing efforts?

Appendix 2. Proposed Project Timeline

Date	Activity	Description
Friday, January 27	Meeting	● Initial Advisory Board Meeting ● Presentation of Project Initiation Document
January - March	Research	● Collection and analysis of relevant data from select EdTech companies
Friday, March 10	Meeting	● Second Advisory Board Meeting ● Presentation of the project's interim status
March - April	Project Finalization	● Completion of report, findings, and recommendations on relevant data collected from select EdTech companies
Friday, April 28	Meeting	● Presentation of the project's final report, findings, and recommendations to the Advisory Board
April - May	Dissemination	● Submission of project findings to various conferences and exploration of other methods to distribute key findings to a wider audience
Friday, May 12	Presentation	● Presentation of the project's findings to Carnegie Mellon University community, including students, faculty, and staff

Appendix 3. One-pager Sent to Prospective Companies

About Heinz College and the Master of Public Policy and Management Program (MSPPM)

The H. J. Heinz III College (Heinz College) at Carnegie Mellon University in Pittsburgh, Pennsylvania is a graduate college that consists of one of the nation's top-ranked public policy programs and information systems schools. The Master of Public Policy and Management Program (MSPPM) offers curriculum tracks in three locations, including Washington, DC (MSPPM-DC).

In the MSPPM-DC track, students take classes on campus in Pittsburgh in their first year and spend the second year in Washington, DC, working four days per week in full-time apprenticeships while completing degree requirements during weekday evenings. This integration of skill building with hands-on application, real-world experience, and networking creates a unique graduate school experience.

About the Systems Synthesis Project

In lieu of a master's thesis, MSPPM-DC students participate in a semester-long group project in the second year called the Systems Synthesis Project, where groups start with an unstructured problem, refine the scope, and then analyze, develop, and implement a solution. The Systems Synthesis Project delivers a capstone experience to students, which integrates core skills from coursework and builds hands-on problem solving, project management, and teamwork skills.

Our Systems Synthesis Project Proposal

- **Title**: *Building Effective Communications around Data Privacy Issues in the K-12 EdTech Industry*
- **Issue**: In the past two decades, big data capabilities have emerged as a promising method for collecting, analyzing, and communicating information about schools, teachers, and students within and across K-12 education management systems. Using online products and services, stakeholders (i.e., educators, parents, and teachers) must contend with a new information landscape, navigating complex questions about sensitive data. As thousands of online applications enter K-12 classrooms, there is an increasing lack of understanding between the education technology companies and the parents and students who wish to understand and protect themselves. In part, student data privacy has evolved into such a hot-button policy due to activity in state legislatures and a recent string of headlines involving high-profile companies and data privacy issues. Companies within the EdTech space have started facing criticism for a perceived lack of accessibility, transparency, and accountability in communicating their data use disclosure, privacy, and security policies to key stakeholders, i.e., schools, parents, and end users.
- **Approach:** We are using a qualitative case methodology to examine a group of 6-10 EdTech companies that meet the following criteria: (1) are 0-8 years old, (2) use click-through agreements for product delivery, and (3) offer educational services in K-12 classrooms. For each company, we will conduct a thorough analysis of public-facing communications strategy around data privacy, including interviews with staff and discussion of privacy practices.
- **Key Deliverable:** We will produce a 25-page report containing our findings, analysis, and guidance on communications best practices to help emerging EdTech companies successfully navigate messaging around data privacy. We will generate a one-page version of this report that contains our key takeaways, as well as a three-page executive summary version.
- **Timeline**: January - May 2017
- **Faculty Advisor:** Andy Richman

Appendix 4. Interview Template

Introductory blurb

We would like to learn about your privacy policies and some of the practices you employ in implementing them. We are preparing a report in which, among other things, we will be briefing our advisory board, members of the Carnegie Mellon community, and various public policy decision makers, and NGOs working on privacy issues. We intend to make policy recommendations in our report that will include many of the issues our interview will address today so thank you again for the time you are giving our project today. It is very helpful.

Introduction/professional background (5-10 minutes)

1. What led you to found [your company] and attracted you about the education technology industry in general?
2. What does privacy mean to you personally?
    a. What steps do you take to protect your own privacy?

Privacy Practices (20 minutes)

1. Could you walk us through how your company established its privacy policies?
    a. How often do you revisit your privacy practices and why do you do so?

POSSIBLE FOLLOW UPs:

- Did you reference resources or outsource this task to someone else either to draft or review?
- Did you seek professional or legal advice when creating your formal policy?
- What specific changes have you made in your privacy practices since you were founded?
- With your investors, did you discuss privacy issues and to what extent? How important was it to them to have a comprehensive privacy standard?

Company-specific Privacy Practice Communications (20 minutes)

1. When it comes to your company, what is your outlook for announcing data collection and/or sharing practices, as well as any other changes to your customers?
2. Do you hire staff who handle data collection communications with customers?
3. Have you made plans if personal information were inadvertently disclosed?
4. How often do users reach out to you about privacy concerns and what are common issues?

POSSIBLE FOLLOW UPs:

- Generally speaking, what is the customer-preferred method for reaching out?
- What are the most common privacy concerns?
- Have privacy concerns prevented a customer from signing up for your service?

- How did you consider the accessibility of your privacy practices when making it, in terms of ease of obtaining and ease of understanding?
- If you signed a pledge to safeguard student privacy, is this information known to your users?
- What resources do you provide for users who have questions about your privacy policies?
- If there is a change in your privacy policy, do you alert your users? How so?

Company Specific questions

(10 minutes)

Varies from case to case

Appendix 5. Literature Review Resources

Title	Last Name	First Name	Year	Publication
I agree, but do I know? Privacy and Student Data	Abilock	Rigele	2016	Knowledge Quest
The Privacy Problem: Although School Librarians Seldom Discuss It, Students' Privacy Rights Are under Attack	Adams	Helen R.	2011	School Library Journal
Guide to Data Privacy Issued for Parents	Armitage	Audrey	2015	Education Week
The Challenge of Data Privacy	Bames	Khaliah	2015	Education Leadership
Obama Tries New Tack to Collect Student Data	Baskin	Paul	2010	Chronicle of Higher Education
How Little Data Breaches Cause Problems	Bathon	Justin	2013	T H E Journal
Managing Student Identities in the Digital Era	Bjerede	Marie	2015	T H E Journal
Realizing Educational Rights: Advancing School Reform Through Courts and Communities	Blokhuis	J.C.	2014	Educational Theory
About This Report	Bushweller	Kevin	2014	Education Week
Microsoft Puts Privacy on Branding Agenda	Cavanagh	Sean	2014	Education Week
Ed-Tech Vendors Attend 'Boot Camp' for Data-Privacy Advice	Cavanagh	Sean	2015	Education Week
Google Commits to Pledge on Student-Data Privacy	Cavanagh	Sean	2015	Education Week
Privacy Protection	Cook	Glenn	2015	American School Board Journal
Privacy Concerns: The Effects of the Latest FERPA	Cossler	Christine	2010	School Business Affairs
Big Data Wants Your Students	Czuprynski	Christine	2015	American School Board Journal
Ed Law: Technology weaves a tangled privacy web	Darden	Edwin	2015	Phi Delta Kappan
5 (Good) Ways to Talk About Data	Datnow	Amanda	2015	Educational Leadership
Indiana Data Network Draws Opposition	Davis	Michelle	2014	Education Week

Title	Last Name	First Name	Year	Publication
New York Dumps inBloom; Now Must Find an Alternative	Davis	Michelle	2014	Education Week
TECHNOLOGY TRANSIENCE AND LEARNER DATA: Shifting Notions of Privacy in Online Learning	Dennen	Vanessa	2015	Quarterly Review of Distance Education
Partnership Books Data Privacy	Doran	Leo	2016	Education Week
Teaching Privacy in the Twenty-First Century	Edbrooke	Odette	2012	Social Education
Making the Right Commitment to Student-Data Privacy	Evans	Cameron	2015	Education Week
Student Privacy in the Age of Big Data	Ewbank	Ann	2016	Knowledge Quest
Replicating the Relationship between Teachers' Data Use and Student Achievement: The Urban Data Study and the Data Dashboard Usage Study	Faria	Ann-Marie	2014	Society for Research on Educational Effectiveness
#ED TECH	Farrace	Bob	2016	Principal Leadership
What it means when we talk student data	Fitzgerald	Bill	2015	Tech & Learning
Data-Privacy Training Lags Behind in K-12.	Flanigan	Robin	2015	Education Week
Managing School Social Work Records	Garrett	Kendra J.	2012	Children & Schools
Assessment Group Approves Privacy Rules for Student Data	Gewertz	Catherine	2013	Education Week
Privacy Group Pushes Parents To Opt Out of FERPA	Gilmore	Sara	2015	Education Week
Past, Present, and Future of Assessment in Schools: A Thematic Narrative Analysis	Green	Stephanie	2015	The Qualitative Report
Student Access to Technology and Its Impact on Achievement	Griffin	Jill	2013	Department of Educational Leadership, Southern Illinois University Edwardsville
Parents need access to education data - and need to know it's secure	Guidera	Aimee	2015	Phi Delta Kappan

Title	Last Name	First Name	Year	Publication
Schools Urged to Put a Higher Priority On 'De-Identifying' Student Data.	Harold	Benjamin	2015	Education Week
New Principles Aim to Guide Use, Safety of Student Data	Herald	Benjamin	2015	Education Week
Prominent Ed-Tech Players' Data-Privacy Policies Attract Scrutiny	Herald	Benjamin	2014	Education Week
Educators Hope Congress Provides Clarity, Support on Privacy Issues	Herald	Benjamin	2015	Education Week
President Obama Propose Student-Data-Privacy Upgrade	Herold	Benjamin	2015	Education Week
"Student Privacy Survey"	Herold	Benjamin	2014	Education Week
U.S. Outlines Data-Privacy Guidelines	Herold	Benjamin	2014	Education Week
Google Under Fire For Data Analysis of Student Emails	Herold	Benjamin	2014	Education Week
Ed-Tech Industry Weighs Impact of New Data-Privacy Laws	Herold	Benjamin	2014	Education Week
Major FERPA Overhaul Under Consideration in U.S. House	Herold	Benjamin	2015	Education Week
Draft bill on Student Privacy Raises Questions	Herold	Benjamin	2015	Education Week
Google Mines Student Data Outside Education Apps	Herold	Benjamin	2016	Education Week
House Bill Aims to Increase Federal Role in Safeguarding Student-Data Privacy	Herold	Benjamin	2015	Education Week
Legislative-Advocacy Group's Model Bill Tackles Privacy of Student Data	Herold	Benjamin	2013	Education Week
Big Data' Research Effort Faces Student-Privacy Questions	Herold	Benjamin	2014	Education Week
Ferpa Primer: The Basics and Beyond	Hlavac	George	2015	NACE Journal
Navigating Student Data Privacy Laws	Holcomb	Carolyn	2015	Risk Management

Title	Last Name	First Name	Year	Publication
Integrating Data Mining in Program Evaluation of K-12 Online Education	Hung	Jui-Long	2012	Journal of Educational Technology & Society
InBloom Sputters As Data Privacy Hits the Spotlight	Kamisar	Ben	2014	Education Week
Putting the Schools in Charge	Katzman	John	2012	The Education Digest
Data Messages	Kaufman	Daniel	2016	American School Board Journal
Craving Capacity	Kennedy	Mike	2016	American School & University
The Transformation of Schools' Social Networks during a Data-Based Decision Making Reform	Keuning	Trynke	2016	Teachers College Record
Aspirational Data Practices for Student Privacy	Kreuger	Keith	2015	School Administrator
Backtalk	Kreuger	Keith	2015	Phi Delta Kappan
New technology "clouds" student data privacy	Krueger	Keith	2015	Phi Delta Kappan
10 Steps That Protect the Privacy of Student Data	Krueger	Keith	2014	T H E Journal
10 Terms You Must Include in Contracts with Online Service Providers	Krueger	Keith	2014	T H E Journal
Let's Get Physical: K-12 Students Using Wearable Devices to Obtain and Learn About Data from Physical Activities	Lee	Victor	2015	TechTrends
Ethical and appropriate data use requires data literacy	Mandinach	Ellen	2015	Phi Delta Kappan
How Data Are Used and Misused in Schools: Perceptions from Teachers and Principals	Militello	Matthew	2013	Education Sciences
Bankruptcy Case Shines Spotlight on Data Privacy	Moiner	Michele	2014	Education Week
On the Block: Student Data and Privacy in the Digital Age--The Seventeenth Annual Report on Schoolhouse Commercializing Trends, 2013-2014	Molnar	Alex	2015	National Education Policy Center, School of Education at University of Colorado
In Data Area, Groups Push More Clarity	Molnar	Michele	2014	Education Week

Title	Last Name	First Name	Year	Publication
Safeguard Use of Student Data, White House Report Urges	Molnar	Michele	2014	Education Week
FTC Acts to Protect Student Data in Ed-Tech Firm's Bankruptcy Case	Molnar	Michele	2014	Education Week
Florida Lawmakers Approve Student-Data-Privacy Bill	Molnar	Michele	2014	Education Week
FERPA and Homelessness: A Technical Assistance Tool for NAEHCY Members	N/A	N/A	2010	National Association for the Education of Homeless Children and Youth (NAEHCY)
U.S. Department of Education's Proposed FERPA Regulations: Overview and Initial Analysis	N/A	N/A	2011	Data Quality Campaign
Call to Action: Clarify Applications of FERPA to State Longitudinal Data Systems	N/A	N/A	2011	Data Quality Campaign
Indiana Department of Education Notice of Procedural Safeguards	N/A	N/A	2010	Indiana Department of Education
Statistical Methods for Protecting Personally Identifiable Information in Aggregate Reporting.	N/A	N/A	2010	National Center for Education Statistics
Forum Guide to Data Ethics. NFES 2010-801	N/A	N/A	2010	National Forum of Education Statistics
Keeping Student Data Private	N/A	N/A	2014	T H E Journal
Manifold path to Happiness	N/A	N/A	2012	Times Education Supplement
Nonprofit Starts Campaign To Push Education as Priority	N/A	N/A	2015	Education Week
Parents Back Data Collection Depending on Its Use	N/A	N/A	2015	Education Week
Schools Gather more Data--But for What purpose?	N/A	N/A	2015	Educational Leadership
Study: No Expectation of Privacy in Massachusetts Schools	N/A	N/A	2016	Information Management
Schools Putting Kids' SSNs at Risk	N/A	N/A	2010	Information Management

Title	Last Name	First Name	Year	Publication
Surveying Encryption Practices of Technology Used Within Schools	N/A	N/A	2016	Common Sense Media
Keeping Student Data Private	N/A	N/A	2014	T H E Journal
Education Data Privacy: Seven Lessons Learned	N/A	N/A	2016	Tech & Learning
Education apps raise privacy concerns	N/A	N/A	2015	American School Board Journal
Guiding Questions for Data Analysis, by Reports	N/A	N/A	2015	Wake County Public School System (WCPSS), Data and Accountability Department
Forum Guide to the Teacher-Student Data Link: A Technical Implementation Resource. NFES 2013-802	N/A	N/A	2013	National Forum of Education Statistics
When Data Drives School Culture	N/A	N/A	2011	Leadership
All About The Data	N/A	N/A	2014	Tech & Learning
Big Data and Administrators: The Struggle to Store, Manage, and Analyze Data Safely	N/A	N/A	2015	T H E Journal
Law and American Education: A Case Brief Approach. Third Edition	Palestini	Robert	2012	Rowman and Littlefield Education
To Protect and Serve	Panter	Suzanna L.	2015	Knowledge Quest
Digital Youth in Brick and Mortar Schools; Examining the Complex Interplay of Students, Technology, Education, and Change	Peck	Craig	2015	Teachers College Record
States Working to Fill Privacy Gaps	Pik	George	2015	Information Today
Framing the Law and Policy Picture A Snapshot of K-12 Cloud-Based Ed Tech & Student Privacy in Early 2014	Plunkett	Leah	2014	Berkman Center for Internet & Society
Assessment in the Kindergarten Classroom: An Empirical Study of Teachers' Assessment Approaches	Pyle	Angela	2013	Early Childhood Education Journal

Title	Last Name	First Name	Year	Publication
Supporting assessment autonomy: How one small school articulated the infrastructure needed to own and use student data	Quartz	Karen Hunter	2014	Journal of Educational Change
Negotiating Access to Health Information to Promote Students' Health	Radis	Molly E.	2016	Journal of School Nursing
Beyond Teach and Hope: Moving from Data to Action	Regan	Kelley	2013	Teacher Education Quarterly
The Editor's Note: Mindful Privacy	Richardson	Joan	2015	Phi Delta Kappan
Promising Principles for Safeguarding Student Information	Robinson	Sharon	2016	Education Digest
Limitations and Analysis of FERPA Policies within Florida Public School Districts	Rogers	Sheryl D.	2012	ProQuest LLC
Amassing Student Data and Dissipating Privacy Rights	Rotenberg	Marc	2013	Educause Review
The Hunt for Data Privacy in the Classroom	Schaffhauser	Dian	2016	T H E Journal
Factors Influencing the Functioning of Data Teams	Schildkamp	Kim	2015	Teachers College Record
Pledge Reflects Ed-Tech Leaders' Concern for Student-Data Privacy.	Schneiderman	Mark	2015	Education Week
Online Student Collaboration and FERPA Considerations	Schrameyer	Alexander	2016	TechTrends
Ed Tech Must Embrace Stronger Student Privacy Laws	Shear	Bradley	2015	T H E Journal
Using data for decision-making: perspectives from 16 principals in Michigan, USA	Shen	Jianping	2010	International Review of Education
Proposed Data-Privacy Rules Seen as Timely for States	Sparks	Sarah	2011	Education Week
Guidance Offered on Guarding Student Privacy in School Data;	Sparks	Sarah	2010	Education Week
The Legal Implications of Surveillance Cameras	Steketee	Amy M.	2012	District Administration
Creating Anytime, Anywhere Learning for All Students: Key Elements of a Comprehensive Digital Infrastructure	Thigpen	Kamilla	2014	Alliance for Excellent Education

Title	Last Name	First Name	Year	Publication
Student data privacy is cloudy today, clearer tomorrow	Trainor	Sonja	2015	Phi Delta Kappan
Putting Data Into Practice: Lessons from New York City. Education Sector Reports	Tucker	Bill	2010	Education Sector
Legal Implications of Using Digital Technology in Public Schools: Effects on Privacy	Tudor	Joanna	2015	Journal of Law and Education
State Lawmakers Ramp Up Attention on Student-Data Privacy	Ujifusa	Andrew	2014	Education Week
States Are Expanding Access to K-12 Data, Group Says	Ujifusa	Andrew	2014	Education Week
Hearing Weighs Student-Data Privacy Concerns	Ujifusa	Andrew	2016	Education Week
Star light, star bright	Vail	Kathleen	2015	American School Board Journal
Revisiting M.R.M. in Search of Legitimate Reasons for Limiting Student Privacy Rights	Walker	Kathryn	2015	Education & Law Journal
Big Opportunities and Big Concerns of Big Data in Education	Wang	Yinying	2016	TechTrends
Privacy Concerns Often Ignored by Reformers	White	Jenni	2014	Education Week
Cyber Savvy: Embracing Digital Safety and Civility	Willard	Nancy	2012	Corwin; A SAGE Publications Company
Aligning the Effective Use of Student Data with Student Privacy and Security Laws	Winnick	Steve	2011	Data Quality Campaign
Determinants of online privacy concern and its influence on privacy protection behaviors among young adolescents	Youn	Seounmi	2009	Journal of Consumer Affairs
A 21st-Century Model for Teaching Digital Leadership	Young	Donna	2014	Educational Horizons

Appendix 6. Methodologies

Literature Review Methodology

In conducting the literature review, our team sought to gain an understanding of the material already available on this topic as well as formulate a novel research angle for our project's focus. We worked with our systems faculty advisor to narrow our scope to K-12 education in the United States, pertaining to the collection of student data.

The team worked with a librarian at the Library of Congress (LOC) to select discipline-specific databases, such as the *Education Resources Information Center* (ERIC), *Education Full Tex*t database, the *ABI/Inform Collection*, and *ProQuest Education*. The librarian advised the team on using Boolean operators in our search queries to optimize our search results, including the following,

- "student data privacy";
- "student record access";
- "student data" AND "privacy";
- "student rights" AND "privacy";
- (K-12* OR educat*) AND (privacy OR access);
- "student data" AND "education technology";
- "data" AND "education apps".

Team members filtered thousands of research citations generated by these queries by location, topic, and time frame in order to isolate sources that reflected the key themes or issues relevant to our team's research interests. This filtering process narrowed the results to about 700 abstracts. After reading each of the abstracts, the team further eliminated those that lacked relevancy.

In total, we selected 135 abstracts, listed in the appendix, for more detailed review, assigned them among team members, and carefully reviewed the articles based on factors like content, analytical rigor, credentials of authors, intended audience, and evidence-based methods. In our review, our team also tracked relevant references to other studies, articles, and compelling sources to locate others. We created a A-B-C-D quality scoring tool and included a description column for team members to provide rationale for scoring for each article. Our quality scoring tool assigned grades to articles based on factors such as relevance, audience, peer review, strength and rigor of research methods, and inclusion of anecdotal vs. evidence-based content.

Ultimately, our team used this information to produce a thorough review of the top-rated available research documents, which includes a detailed description of the list of studies that our team rated in the 'A' range, using our quality scoring tool. As mentioned in the preceding analysis, the body of literature is light on the education technology component.

Case Selection Methodology

In order to uncover the privacy communication practices of companies, we pursued a qualitative case methodology by examining a group of six EdTech companies. For each company, or case, we conducted a parallel analysis, including interviews with staff, primary analysis of data privacy practices, and examination of how the company communicates their privacy practices to key stakeholders.

Our case study methodology is broken into five parts: (1) creating a database of EdTech startup companies, (2) selecting the cases, (3) recruiting the companies, (4) executing research, and (5) analyzing the results from the case studies.

Creating a database of EdTech startup companies: To start the process, we needed to create a database of relatively new EdTech companies so that we would have a list of organizations that we could systematically research and contact. Our process for creating the database is detailed below.

1. The team downloaded lists of startup companies from relevant databases. The lists were then merged into an excel sheet and duplicate companies were removed, resulting in a file with 450 companies.
2. Not all of 450 of the companies fell into the EdTech category and thus needed to be filtered manually one by one. For our purposes, to filter only EdTech companies from the original 450 companies, we looked at the following criteria:
    - Is the company an EdTech company?
    - Do they operate in the United States?
    - Do they work with the K-12 market?

   Only companies that met the all three criteria remained on our list for a more in depth review in order to determine whether or not they would make good case study candidates. This resulted in a list of 120 EdTech companies that were based in the United States with a K-12 audience.
3. The team members each researched a portion of these 120 companies in a systematic way, filling in the information for the following pieces on each company:
    - Website and Privacy Policy: The website and privacy policy URLs
    - Year founded: The year the company was founded
    - Employees: The number of employees in the company
    - Funding: The amount of funding given to the company to date
    - Category: The EdTech segment they belonged to (games, adaptive learning, classroom management, etc.)
    - Information Risk: The amount of PII collected by the company and how they used this information
    - Entrance into schools: How these companies enter into schools (through teachers, students, administrators, etc.)

   With these fields of information entered for each company, the team effectively created a brief snapshot of each organization and their characteristics.

Creating a purposive sample: Our research focused on several different types of companies and their place in the EdTech industry. Our standards for selecting the cases were as follows:

1. We wanted a good representation of the diverse EdTech startup landscape so we chose to pursue companies that varied from each other. No two companies were the same and differed in characteristics such as Category, Information Risk, Entrance in schools, etc.
2. From our database of 120 EdTech companies, we selected 18 as potential case study prospects. Although we only needed six, the team contacted all 18 companies because we expected that some would be unresponsive or unwilling to participate in our study.

Recruiting the companies: The success of our research was contingent on the participation of EdTech companies and the reliability of the information we received from them. The recruitment strategies we used are listed below:

1. To incentivize participation, our team committed to share copies of our final report, policy brief, general research deliverables, and other insights with the companies we examined. We anticipated that recently formed companies would appreciate the opportunity to learn more about how the industry discusses privacy issues with potential customers.
2. To protect our subjects, we anonymized the companies in our report and did not refer to them by company name. Although we took this step to protect their identities, we believe that some EdTech companies are unique to the point that they may still be identifiable.

Executing Research: Once our cases agreed to participate, we analyzed each case in an analogous way, detailed below.

1. Every case had two team members devoted to it, and team pairings differed from case to case (i.e., for six cases, each team member worked on two cases and had a different partner for each case). This ensured that ideas among the group cross-pollinated and that each case analyzed benefited from multiple perspectives.
2. Techniques of research and inquiry were broadly identical to ensure that the team maintained research fidelity across cases. The team met regularly to discuss progress and to confirm that we maintained consistency in our case methodology. However, there were instances in which each research team asked questions that applied to only that particular company and not others. Since we were explicitly relying our collective ability to pull multiple qualitative threads together, we did not want to severely limit any individual team’s ability to fully analyze their case company.
3. For each company, or case, we used structured interviews with staff, primary analysis of data privacy practices (as much as possible), examination of how the technology is used in the classroom, and other kinds of evaluation. Through our interviews, we verified what kind of data storage and management practices exist at the companies we examined. When our technical knowledge fell short, we sought advice from experts in the community.

Analyzing the results: Upon completion of the individual company interviews, we generated a full final report to discuss our findings, as described below.

1. These findings are separated into two parts.
    a. The first part includes a reviews and analysis of each company case. The write-up of individual company cases was written by each paired research team, and it examined how the company deals with privacy concerns, exploring technical, legal, and policy dimensions of the issue as it relates to each case company.
    b. The second part of the final report synthesized the trends we all saw in all of our cases, the literature we have reviewed, insight from our advisory board, interviews we have done throughout the project, and our own reflections. This second part was a collaborative effort across the entire research team. We also discussed what we *do not* know and could not learn through this process, the pitfalls of our own methodology, and the questions that remain unanswered.