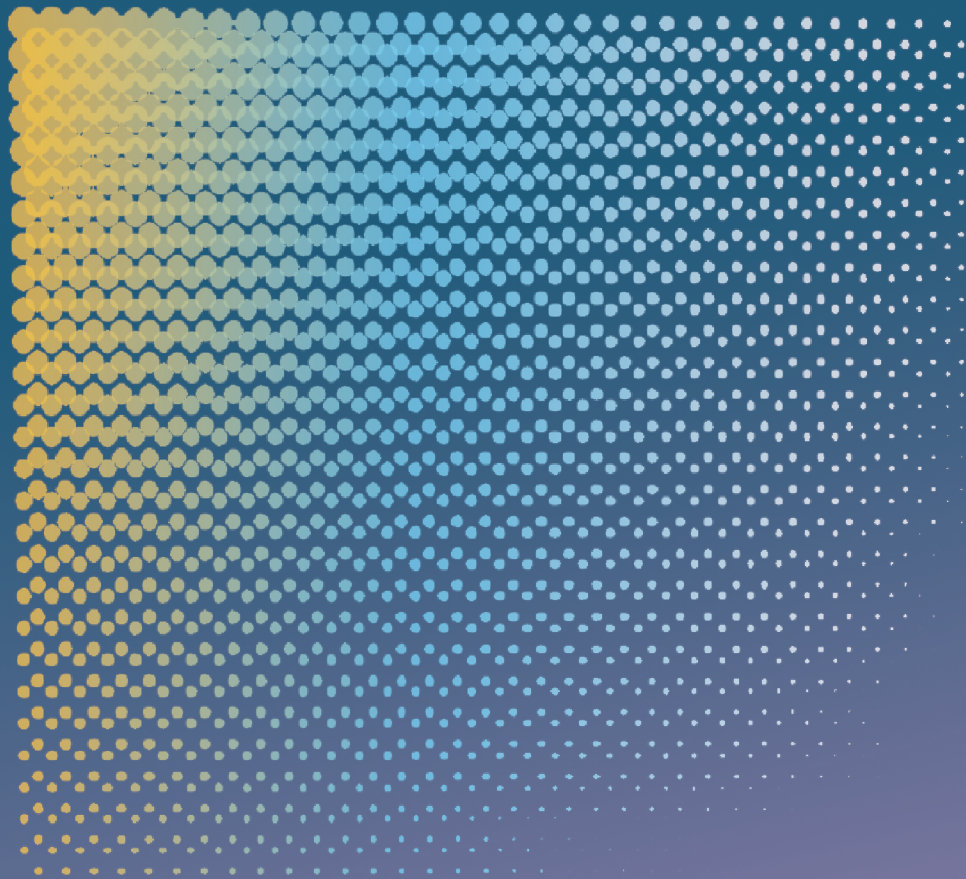
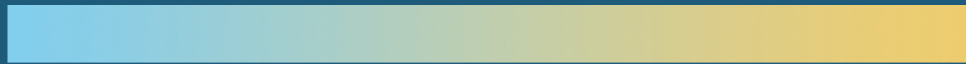


Data Ethics

in Education & the Social Sector



What Does It Mean and Why Does it Matter?

February 2021



ABOUT CENTER FOR DEMOCRACY & TECHNOLOGY

The Center for Democracy & Technology is a 501(c)(3) working to promote democratic values by shaping technology policy and architecture, with a focus on the rights of the individual. Based in Washington, D.C., and with a presence in Brussels, CDT works inclusively across sectors to find tangible solutions to today's most pressing technology policy challenges. Our team of experts includes lawyers, technologists, academics, and analysts, bringing diverse perspectives to all of our efforts. Learn more: cdt.org/

ABOUT STUDENT PRIVACY

CDT's vision for the *Student Privacy Project* is to create an educated citizenry that is essential to a thriving democracy by protecting student data while supporting its responsible use to improve educational outcomes. To achieve this vision, CDT advocates for and provides solutions-oriented resources for education practitioners and the technology providers who work with them, that center the student and balance the promises and pitfalls of education data and technology with protecting the privacy rights of students and their families.

ABOUT EQUITY IN CIVIC TECH

As governments expand their use of technology and data, it is critical that they do so in ways that affirm individual privacy, respect civil rights, foster inclusive participatory systems, promote transparent and accountable oversight, and advance just social structures within the broader community. CDT furthers these goals by providing balanced advocacy that promotes the responsible use of data and technology while protecting the privacy and civil rights of individuals. We engage with these issues from both technical and policy-minded perspectives, creating solutions-oriented policy resources and actionable technical guidance.

AUTHORED BY

Elizabeth Laird, CDT Director, Equity in Civic Technology

Hannah Quay-de la Vallee, Senior Technologist

This report was created with consulting support from Maya Lagana, Independent Consultant, & based on advisory guidance from Jake Metcalf, Ethical Resolve.

Data Ethics in Education and the Social Sector: *What Does It Mean and Why Does it Matter?*

Executive Summary

Data and technology are playing an increased role in educating and supporting students and their families. This has become even more pronounced as a result of the global pandemic, during which schools are relying on technology to support distance learning and may collect new data about students' health to provide safe in-person instruction. With this growing reliance on data and technology comes an increased responsibility to ensure not only that data is kept private and secure, but also that new technologies are used only to benefit students, not limit their educational opportunities. To achieve this goal, policymakers and practitioners need to look beyond privacy and security to how data and technology might be misused or have unintended consequences. Responsible planning and clear guardrails can help ensure technology is used to support students and their families.

A data ethics approach can support the education sector in shaping policies and practices that promote the responsible use of data and technology. Simply put, data ethics is the application of ethics concepts and resources to the specific challenges of responsible data use. However, operationalizing ethical data practices requires further understanding and scoping around two key areas:

- Supporting fair and equitable use of data and technology to maximize the potential to improve the public good and lives of individuals; and
- Managing risks that lead to negative impacts on individuals, especially vulnerable populations.

As highlighted in *Figure 1*, important topics within both of these areas should guide the education sector's use of data and technology.

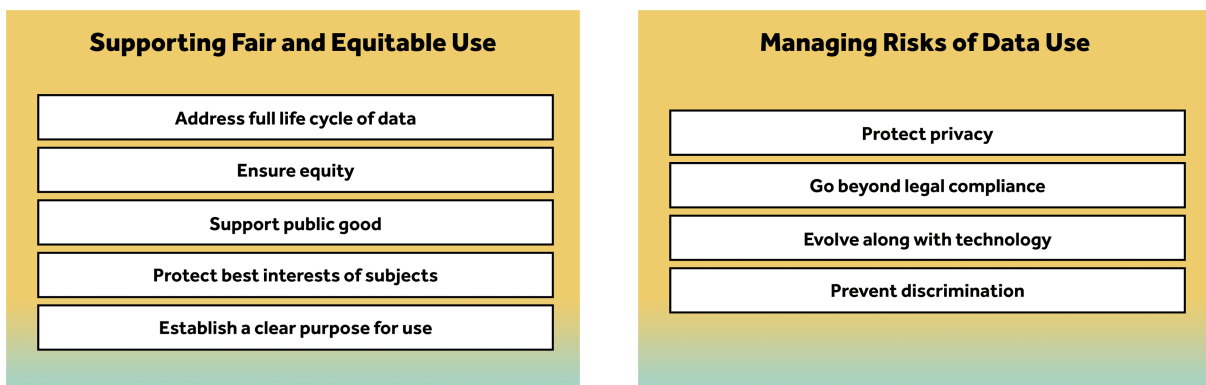


Figure 1. Data ethics definitions and related topics

Additionally, *Figure 2* provides an overview of the range of reasons why it is important to address data ethics, both to achieve proactive benefits and avoid potential risks:



Figure 2. Reasons to address data ethics

Finally, there are several key issues that must be addressed to ensure that data is used responsibly and ethically:

- **Governance and oversight:** Data governance is the management of data and technology, including people, processes, and structures that oversee their use. Governance is fundamental to supporting ethical data use and should address issues like data quality; data retention, deletion, and minimization; promoting user access and data ownership; and outlining and enforcing appropriate uses of data and technology. These structures are important in their own right, and also lay the foundation for addressing other important data ethics issues.
- **Stakeholder engagement:** Organizations should engage stakeholders like students and families, teachers, and administrators throughout the process of adopting and using data and technology. Doing so has multiple benefits, most notably increased buy-in and trust in the use of data as well as the organization more broadly. It can also result in the early detection of concerns, and allow for more inclusive and robust solutions for those concerns.
- **Equity and bias:** The use of data and technology has the potential to improve equity and limit biases, but only if the collection, analysis, and use of data is designed intentionally to meet these goals. There are many examples of how the use of data and technology reinforces the biases and inequities present in society, particularly in research uses and artificial intelligence applications.
- **Transparency:** Transparency throughout the data life cycle enables individual providers of the data to be informed participants, helping build trust in the process.
- **Capacity-building:** Because data and technology are rapidly evolving, it is a challenge for organizations, especially under-resourced organizations like schools, to have the capacity to enact and follow ethical data practices and policies. Organizations can build capacity by participating in trainings, creating guidance resources, and having dedicated staff to support team members in this work.

- **Secondary data use:** Once data has been collected, it is sometimes reused for additional purposes beyond the original intended use, potentially diverging from the scope of what the data subject previously consented to or was notified of. This can violate the trust and consent of the data subject, since it does not meet their expectation of how the data was meant to be used.
- **Privacy and security:** Privacy is the idea that people should be able to control their own information, and that the entities that are authorized to collect and use that information do so in ways that respect an individual's autonomy. Security is the practice of preventing unauthorized access to information and the systems that hold it. In education, data privacy is governed by laws at both the state and federal levels, although these laws often do not fully cover associated best practices.
- **Open data access and research:** Some data sets may be aggregated or stripped of identifiable information for purposes of sharing with researchers or the broader public, which can raise issues of privacy and security, risk of re-identification, secondary data use, transparency, and consent.
- **Consent:** Consent can play a role in supporting responsible data use, even in instances where it is not legally required, as it can assist with ensuring stakeholder buy-in and transparency. However, ethical data use cannot rely on users' consent alone. There are important challenges to informed consent, including whether the user really reads and understands the ways their data may be used, and whether the user feels they have a meaningful, non-coercive choice. Therefore, consent alone is not sufficient to protect individuals and should be complemented with other practices.

Although the field of data ethics is still evolving in education, responsible data use is of critical importance for all organizations that seek to utilize data and technology to improve outcomes for students and their families. A particular use of data or technology might be legal, private, and secure, but not in the interest of students and families. Therefore, data ethics is an important framework to support beneficial data and technology uses while minimizing potential harms. This report seeks to provide a working definition of data ethics, explain why it matters for educators and public officials, and explore key topics and practices that need to be addressed to operationalize data ethics principles.

Introduction

Although the use of data and technology in education and the social sector has increased over time, the global pandemic has led to a sharp expansion of their use.¹ School systems have had to grapple with how to use data and technology responsibly while still protecting student privacy.² For example, schools may be able to securely collect new data about students like their health symptoms or their COVID status, but there are active debates about what should be done with that information.³ Schools must create a safe environment for teachers and staff while ensuring that this sensitive information is not used to discriminate against, stigmatize, or otherwise cause harm to students.

Given this unexpected expansion, it is more urgent than ever for educators and administrators to have the skills and knowledge to use technology and data responsibly. Whether schools are providing services in-person, virtually, or through a hybrid of the two⁴, they must reckon with the increase in the use of educational technology. Recent research suggests that approximately three in four teachers and parents⁴ support an increased level of online learning after the pandemic ends, reiterating that the need for responsible data use will only grow.

Despite significant discussion of student privacy, there has been limited exploration of broader issues around responsible use of data and technology in education—a field often referred to as data ethics. In light of this gap, this report seeks to provide a working definition of data ethics, explain why it matters for educators and public officials, and explore key topics and practices that need to be addressed to operationalize data ethics principles. The main focus of this paper is high-level data ethics issues and practices in education; however, the *Deep Dive in K-12 Education* on p. 21-24 and Appendices A-C provide more detailed information about these issues in higher education, the social sector, and other emerging technologies.

Look for this icon throughout the document to identify topics covered in more detail in the Deep Dive on Data Ethics in K-12 Education (on p. 21-24 of the report).



What Is Data Ethics?

Data ethics is the application of ethics concepts and resources to the specific challenges of responsible data use. However, there is wide variation in how this term is interpreted and operationalized. Based on research of the field of data ethics and its application to education

* For CDT-published resources around student privacy during COVID-19, please visit cdt.org/student-privacy.

and the social sector, CDT has developed the following working definition of data ethics that will be used throughout this document:

A set of evolving principles on how to collect, manage, share, and use data in a secure and equitable way that avoids harm to the individual and is in support of the public good.

In addition to this working definition, operationalizing ethical data practices requires further understanding and scoping around two key areas:


- **Supporting fair and equitable use** of data and technology to maximize the potential to improve the public good and lives of individuals; and
- **Managing risks** related to data and technology use that lead to negative impacts on individuals, especially vulnerable populations.

Supporting fair and equitable use: One focus of data ethics is the fair and equitable use of data and technology to support an organization’s mission and goals. This includes furthering the aims of the organization as well as supporting individuals.⁵ In the case of education, this means using data and technology to improve outcomes for every student.

To support fair and equitable data use, practitioners and policymakers should consider five guiding principles:

- *Address the full life cycle of data:* Data ethics should be incorporated into the full life cycle of data, including collection, management, analysis, application, and ultimately deletion or archiving once the data is no longer being used.
- *Ensure equity:* Data ethics requires organizations and practitioners to be responsible for ensuring equitable use of data, but it is important to consider what equity means in this context. Equity can refer to minimizing bias, addressing power imbalances, using data to highlight existing inequities, and ensuring equitable access to data and technology.⁶
- *Support the public good:* Using data and technology responsibly includes a clear moral imperative to promote the greater good rather than self-interest. Supporting the public good can include making data publicly available to researchers for broader benefit (while bearing the privacy and security of data subjects in mind), as well as looking at how data is used once it has been collected.
- *Focus on the best interests of data subjects:* In addition to supporting the public good, an ethical approach should also ensure that the interests of the individuals providing the data are considered by those using the data. Centering the best interests of data subjects necessarily means protecting their privacy, but should also consider how individuals can benefit from data and technology use, granting individuals access to

their data, and minimizing potential harms like discrimination. Organizations should consider this aspect of data ethics as the technology-minded version of the Hippocratic oath, in that the data collector should follow a principle of “do no harm” to the data subject.⁷

- *Establish a clear purpose for data use:* Organizations should be thoughtful in how they collect, manage, and use data, rather than using data for data’s sake.⁸ They should establish a purpose-driven plan for how they will use data and technology to benefit both the public good and individuals. 

Managing risks of data and technology use: The benefits of using data to drive improvement across sectors are clear, but so are the risks to privacy and equity. High-profile privacy breaches and questionable uses of data and technology as schools moved online in spring 2020 provide recent examples of how data and technology can harm individuals rather than help them.⁹

Data ethics provides a framework to help organizations and practitioners reap the benefits while mitigating and minimizing the harms to individuals. To that end, practitioners and policymakers should address three aspects of data and technology:

- *Protect privacy:* Protecting sensitive data is a key component of data ethics. This includes privacy as well as secure storage, user access control, and data retention policies.¹⁰
- *Go beyond legal compliance:* Laws related to how data is collected, shared, and used are often insufficient to address new uses of data and new technologies, as technology often evolves more rapidly than the legal framework that governs it. A focus on data ethics can help organizations identify these deficiencies and develop norms beyond existing laws.¹¹
- *Evolve as technology and policy changes:* Data ethics principles need to evolve along with technology to address the issues that arise as data and technology shift.¹² Organizations should revisit their policies on a regular basis to keep pace with changes in technology and policy.
- *Prevent discrimination:* Data ethics seeks to prevent potential harms that could occur from data and technology use including discriminatory actions, some of which may carry legal risks as well. Emerging uses of algorithmic systems are particularly susceptible to exacerbating inequities.¹³

Why Does Data Ethics Matter?


There are a series of benefits that organizations can achieve by incorporating data ethics considerations, as well as potential risks if they ignore or fail to address the concerns these considerations are meant to target.

Benefits of incorporating data ethics:

- *Balance competing interests:* The use of data requires balancing competing interests, including a common tension around individual versus collective good. One role of data ethics is to balance these two interests appropriately so that individuals do not experience undue harm for the collective benefit, particularly if those individuals are unlikely to experience the benefit themselves. This tension commonly plays out between the benefits of research that aims to improve quality of life for many people and the privacy concerns of an individual whose data would be used in that research. While there is not a universal answer to the question, organizations can use data ethics principles as a framework with which to weigh these risks.¹⁴
- *Build and maintain equity:* The use of data and technology has the potential to improve outcomes and increase equity if used responsibly; however, it can also perpetuate existing biases if safeguards are not put into place. Incorporating data ethics principles can help organizations ensure that they are considering equity when engaging with data and technology.¹⁵
- *Adhere to a moral imperative:* While different organizations approach their roles in a variety of ways depending on the sector, mission, and other factors, there is a strong moral imperative to support responsible data use.¹⁶ This is because of the potential harms to individuals, particularly when considering sensitive personal information, as well as the outsized harm that data can cause to vulnerable groups.
- *Gain public trust:* An organization's success is at least in part dependent on its reputation within its community. Consideration and responsiveness to data ethics considerations is increasingly important to maintaining reputation and, correspondingly, public trust.¹⁷ Organizations should not just address these issues, but be transparent and proactively communicate with the community about their efforts.¹⁸
- *Ensure data-driven continuous improvement:* Data ethics principles can also help establish trust in organizations, allowing those organizations to continue collecting and learning from data from their community. Failure to ethically use data and technology can lead community members to disengage from any data collection, limiting an organization's ability to use data for good.¹⁹
- *Prevent large-scale crises:* There are many examples of data ethics issues that only get attention when a large-scale crisis such as a data breach occurs; in some ways, data breaches are a risk of failure to incorporate ethical considerations. If not prevented, data breaches and other large-scale crises are damaging to organizations in terms of both reputational risk and their ability to achieve their goals if data is no longer usable (either because it has been corrupted or because users rescind consent).²⁰ Being proactive in incorporating data ethics can help to both avoid such crises from occurring and also limit the fallout to reputational risks and individual harm if such a crisis were to occur.²¹

In addition to these themes, the increasing ubiquity of data and technology in all sectors makes it important for every organization to pay attention to issues of data ethics.

Risks of not addressing data ethics:

- *Entrench inequalities:* Mission-oriented organizations, like those in the education and social sector, need to guard against potential misuses of data and technology that work at cross-purposes of their goals to increase equity and opportunity.²²
- *Lose public trust:* Organizations need to be ahead of crises that would result in loss of public trust. In the government and nonprofit sectors, this is particularly important given that these organizations often offer important services, which they are unable to provide if the community distrusts them.²³ Loss of public trust can limit an organization's efficacy, as well as limit its ability to carry out an equity-driven mission if data is not being used in a fair way.²⁴
- *Suffer backlash against data use:* There are numerous examples, including in education, of inattention to data ethics leading to a backlash against data and technology use. The public will assume that, because an issue arose, it is not possible to use data responsibly or that the risks outweigh the benefits, or that the organization is unable or unwilling to mitigate those risks. This is a reason why education (both K-12 and higher ed) are behind other sectors in the use of data, despite the many benefits of data and technology.²⁵ 
- *Breach sensitive information:* Security breaches not only damage public trust but can have harmful and long-lasting effects on individuals, including identity theft, stigmatization, and cyber-bullying. Ethical data practices can abate the risk of security breaches, and minimize privacy harms as well as the loss of trust and sense of privacy.²⁶

What Are Key Data Ethics Issues, and What Can Leaders Do About Them?

Ethical data practices and policies span multiple important issue areas. The sections below provide overviews of the following issues, as well as actions that leaders can take to support the responsible use of data:

- Governance and Oversight
- Stakeholder Engagement
- Equity and Bias
- Transparency
- Capacity-Building
- Secondary Data Use
- Open Data Access and Research
- Privacy and Security
- Consent

Governance and Oversight

Data governance is “the overall management of data, including its availability, usability, integrity, quality, and security,”²⁷ and includes people, processes, and structures that are

responsible for data and technology. It serves as a foundation from which to address many other issues related to the responsible use of data, as it is a framework for building and managing data use policies.²⁸

Data governance commonly addresses issues around ethical data practices, such as:

- *Data minimization*: Organizations can be inclined to collect as much data as possible, regardless of whether there is an identified need for the data. The principle of “data minimization” encourages the collection only of data that has a clear purpose and use, minimizing the risks of collecting data that is not used or needed (like data breaches or using data out of context). Additionally, select and limited data sharing and privacy-protective user access may assist in data minimization by ensuring that the same data is not being collected multiple times for different purposes.²⁹
- *Data ownership*: Defining who has the ultimate control and legal rights over the data is an important decision that is best done early and documented in a formal agreement.³⁰
- *Data storage*: Governing data storage includes the rules and processes to ensure appropriate deletion, retention, and maintenance of data. It is important that these rules and processes are established on the front end, shared transparently, and designed to protect user data.³¹
- *Data sharing*: Data governance is even more essential when data is shared across multiple agencies, as the issues become more complex.³² Organizations will need to consider whether sharing is appropriate, necessary, and consistent with users’ expectations, and develop clear policies that govern the roles, responsibilities and processes for sharing. This includes requirements around privacy, storage, use, and deletion. Effective data governance approaches provide a framework for different organizations to ensure that they are using consistent practices and procedures when dealing with their shared data.³³
- *Data deletion*: Because all data comes with some amount of risk, collecting unnecessary data or keeping data after it is no longer useful creates an unnecessary risk that that it could be used out of context or exposed in a breach. Creating policies and procedures governing when and how to delete data can minimize the risks that come with amassing unnecessary data.³⁴
- *User access*: Limiting user access to only individuals who have a clear need for it can help organizations ensure privacy protection and minimize the likelihood of inappropriate data access or misuse.³⁵
- *Data quality*: A perennial issue that underlies all of these considerations is the need to ensure that data is accurate and valid before using or sharing it. Organizations have an ethical obligation to ensure the accuracy of data, and if they do not, any insights gleaned from that data or actions taken based on that data may be misguided and do more harm than good.³⁶ Organizations should consider adopting mechanisms for users to view and request the correction and deletion of information held about them.
- *Documentation*: For transparency purposes and to ensure policies are maintained for as long as data is being used, organizations should put in place appropriate documentation

around data use, including internal policies and procedures, data sharing agreements, decision-making frameworks, and well-constructed contracts, particularly as they relate to technology vendors.³⁷ Ideally, government agencies (such as school districts or state education agencies) can create systemic frameworks for these types of documents to help support consistency across organizations, rather than having individual organizations create their own unique documents.³⁸

It is important to note that many of these issues are also important when working with outside vendors and across the life cycle of data-related projects.³⁹



Actions for Leaders:

Data governance is an ongoing process that is never finished. Leaders can support effective governance of data and technology by taking the following actions:

- Set a clear vision and mission for data governance;
- Establish a data governance body and ensure that it has members that represent all relevant parties and a broad range of expertise;
- Clarify the roles and responsibilities of both the governing body and individual members, including decision-making rights;
- Create transparent processes, including for decision-making; and
- Provide resources and training to ensure the long-term sustainability of the governing body.⁴⁰

Stakeholder Engagement

Organizations will benefit from engaging stakeholders as they develop their data governance frameworks⁴¹. Data and technology initiatives benefit from diverse perspectives, which help surface potential problems and develop frameworks that work for a broad cross-section of users. Stakeholder engagement will also increase buy-in and trust in how data and technology are used, which can increase faith in an organization more broadly. Moreover, there are risks to not engaging stakeholders in decision-making about data and technology, in that organizations are more likely to encounter pushback on how data is being used. In the event of a breach or other issue, people are more likely to be understanding if they had buy-in on the front end, seeing firsthand that meaningful steps were taken to put protections in place.⁴²

Stakeholder engagement can range from informational, to advisory, to giving stakeholders decision-making authority. Depending on the topic, it may or may not be appropriate or necessary for stakeholders to have decision-making power. Some argue that certain groups (e.g., parents of K-12 students) should be involved in decision-making, while others argue that for certain issues, engagement should be more informational in nature.⁴³

Actions for Leaders:

To create trust, understand and address community concerns, and strengthen relationships with the community, leaders can engage stakeholders—especially those about whom data is being collected—by taking the following actions:

- Identify a clear purpose for community engagement;
- Engage communities early in the decision-making process;
- Proactively communicate with families, giving them information upon which to base their input;
- Build capacity among communities to help them engage;
- Prioritize inclusivity within community engagement efforts;
- Be open to implementing feedback gathered; and
- Clearly communicate why the organization is collecting and using the data or deploying a new technology tool, to ensure understanding and build buy-in.⁴⁴

Equity and Bias

The use of data and technology has the potential to improve outcomes, in that they can be used to benefit individuals (e.g. share data to coordinate services across different public agencies to provide better care) as well as improve the public good (e.g. analyze data to reveal policies and practices that are working and could be replicated). However, these benefits will only be realized if the collection, analysis, and use of data is designed intentionally to meet these goals and minimize potential bias. To this end, it is important to identify and address the ways in which using data and technology could inadvertently create, entrench, or worsen inequities or have other unintended consequences.

Certain data elements can be inherently biased, and including them in analyses will bias the outcomes of the analysis towards (or against) a particular group. In education, students of color are disciplined at a greater rate than their peers (both in terms of number of infractions as well as the severity of consequences), so using discipline data in certain analyses could result in the over- or under-identification of students of color, which could negatively affect their outcomes. Alternatively, using data from a non-representative sample (e.g. omitting certain groups or simply failing to collect sufficient data from those groups) and then applying the findings to the broader population can result in practices or policies that are not beneficial for certain populations within the broader community.⁴⁵

For example, federal funding for schools is determined, in part, by the U.S. Census; however, language barriers as well as fears of government and immigration enforcement often results in an under-counting of Latinx and immigrant families. With the most 2020 Census data collection, the inclusion of a question about citizenship status threatened to further bias the underlying data, exacerbating risks of using this data and resulting in some school districts losing hundreds

of millions of dollars and limiting their ability to serve all students, especially the highest-needs students.⁴⁶

In addition to the danger of bias in how data is collected and used, organizations should also consider other equity issues that can arise from the use of data and technology. Among those issues is equitable *access* to data, which can be influenced by the power dynamics of a given situation. The organizations that have authority over individuals may be the same organizations that control whether and with whom their data is shared. For example, in the case of education records, public school parents have legal rights to access and correct records, but federal student privacy laws do not require schools to honor those requests or delete information after a certain period of time. This imbalance means some parents might be successful in cleaning up or limiting access to their child's records, while other parents will be unsuccessful in their attempts or not even try for fear of damaging relationships with the school staff. In this dynamic, individuals may lack the power or comfort to request access to information about them (to correct or delete it), or the power to push back if this request is not honored.⁴⁷

Lastly, emerging technologies have the potential to exacerbate bias. For example, predictive analytics, particularly when machine learning is utilized, can significantly increase inequitable outcomes if bias is not accounted for in their design and evaluation.⁴⁸ For example, as previously mentioned, students of color are disproportionately disciplined at a greater rate than their peers, so early warning systems that use discipline data to predict whether a student is on track or at risk of dropping out of school will identify more students of color. In the case of Pasco County,⁴⁹ this information was being shared with law enforcement, which could further expose students of color to law enforcement and reinforce the school-to-prison pipeline. Additional examples of emerging technologies, as well as their impact on equity and bias, can be found in Appendix C.

Actions for Leaders:

To address equity and minimize bias in data and technology, leaders can take the following actions:

- Examine input data at the outset and mitigate potential biases before the system is deployed;⁵⁰
- Be conscious of biases and ensure they are not being built into algorithms or other decision-making structures;⁵¹
- Be aware of the data set or technology tool's fitness for its intended purpose and use it only for this purpose;
- Avoid over-reliance on algorithmic tools and instead use them as one input for decision-making;⁵²
- Conduct ongoing stakeholder engagement and transparency to ensure biases are being identified and addressed; and

- Conduct continuous oversight of algorithms and decision-making processes to ensure they are appropriate, and address any biases that emerge.⁵³
- More information about how to ensure equitable access to data is provided in the section on *Open Data Access and Research*.



Transparency

Data ethics seeks to ensure transparency at all stages of the data life cycle, from collection through analysis and use, to support data quality, create trust, and establish buy-in. Transparency about data and technology can be defined in a number of ways, but can generally be put into three categories: transparency about data collection, data use and storage, and data-driven decision-making.⁵⁴ Specifically regarding decision-making, transparency includes visibility into how decisions are being made based on data, including methodology, decision-making processes, and the underlying data itself.⁵⁵

Lack of transparency can create issues both for organizations collecting data and individuals whose data is being used. If individuals do not understand the decision-making processes that utilize their data, they will not have buy-in to the decisions that impact them directly.⁵⁶ Organizations should build transparency to support data ethics goals, such as: building public trust in data and its uses; ensuring accuracy of data; establishing accountability; and creating equity in access to data that can be used for research purposes.⁵⁷

Actions for Leaders:

To support meaningful transparency, leaders should take the following actions:

- Be clear on the what, who, how, and why to all stakeholders on the front end and across all stages of the data life cycle;
- Go beyond compliance and seek to build understanding, avoiding the common trap of publishing an overly complex notice about data usage that is not actually accessible for key stakeholders;⁵⁸ and
- Provide information in multiple formats and modalities to reach different audiences.



Capacity-Building

As data and technology rapidly evolve, many organizations may feel they lack the capacity and knowledge to develop and maintain appropriate data practices. However, in order for organizations to use data ethically and to its full potential, staff and other stakeholders must have a clear understanding of data, technology, and their purposes, or access to technical support so they can address potential issues that may arise.⁵⁹

When building capacity around data and technology, the following areas are especially important to consider:

- Awareness of all the data maintained by the organization and the ethical implications of holding and using that data;
- Building an organizational culture that is data-driven and places a high premium on the responsible use of data;⁶⁰
- Relevant laws and policies that should govern the use of data;
- Engagement with affected communities about the collection of their data and how it will be used;⁶¹
- Decision-makers' understanding about the data that they are using to make decisions, including the data's strengths and limitations, and the potential risks of bias;
- Awareness of bias among those who are developing algorithmic decision-making systems, as well as an understanding of how additional qualitative or anecdotal information might be relevant and when that contextual data will require human intervention in the automated system;⁶² and
- Documentation to ensure knowledge transfer in the event of staff transition.⁶³

Actions for Leaders:

Organizations can increase their capacity to engage in responsible data use by participating in trainings, creating guidance resources, and having dedicated staff to support team members in this work, incorporating professional development best practices more generally. To do this, leaders should take the following actions:

- Ensure all staff and other key stakeholders are trained to understand how data is collected and used;
- Provide customized training dependent on individuals' role and prior knowledge;
- Create documented policies and procedures, including versions that can be shared publicly to communicate the organization's approach;
- Provide regular follow-up trainings, both to refresh knowledge and discuss new developments; and
- Provide oversight and accountability to ensure that implementation is occurring and not forgotten once the training is completed.⁶⁴



Secondary Data Use

As discussed in the *Transparency* section, organizations should communicate to stakeholders what data will be collected about them and how it will be used. However, once data has been

collected, it often may be re-used for additional purposes beyond the original intended use, potentially diverging from the scope of what the data subject was notified of or consented to. Algorithmic decision-making systems often rely on data from disparate, integrated data sets to identify unanticipated patterns, which incentivizes data holders to integrate and repurpose data sets without knowing in advance how the data will be used. While repurposing data may be useful for gaining insights and improving systems, it complicates other data ethics issues like transparency, community engagement, and consent.

Secondary data use can become an issue with any data that is collected, but occurs most frequently in these scenarios:

- Data was collected for informational purposes and then is used for decision-making;
- Data was originally not going to be shared with outside agencies, but in the end the decision is made to share the data externally;
- Data was collected to support the individual but then is used for a collective purpose, such as research; or
- Data was collected, aggregated, and used for systemic decisions but is then disaggregated and used to make decisions about individuals.⁶⁵

A risk of secondary data use is losing public trust, especially if used for purposes beyond which consent was given or in a way the individual would not expect. Additionally, secondary data use can increase privacy and security risks if data is kept longer than initially intended or shared beyond its initial purpose, and data may be misused if it is shared with other parties who may have different missions, different data governance standards, or a lack of understanding of the data provided.⁶⁶

Secondary data use is especially pertinent to research. Data is often collected across fields to track and support individual outcomes (e.g. test scores or health screenings), but may also be helpful to research to support the broader sector. Often these research projects have not yet been identified at the time of the data collection, so consent can be difficult or impossible to collect. In some cases, de-identified or aggregate data could be used for research purposes and may pose less of a privacy risk, but de-identification must be done carefully by someone with proper training to minimize the risk that the data is re-identified, thus exposing the individuals to privacy loss, financial risk, or other harms.

Actions for Leaders:

To limit secondary data uses, especially those that are not communicated and for which consent was not given, leaders can take the following actions to minimize related risks:

- Establish strong data governance that evaluates secondary data use independently for their harms and benefits, which can serve to limit the

misuse of data (it is also important to acknowledge that particularly sensitive data like genetic and health data cannot be re-used absent direct and express opt-in).⁶⁷



- Delete data after it is no longer needed for its initial purpose and ensure only anonymized data is stored;
- If consent is collected, provide the full range of potential uses; and
- Proactively communicate any changes in the use of data and collect consent from original data providers.

Open Data Access and Research

Initiatives that are aimed at providing open data access and supporting research refer to increasing access to information that could be analyzed and used to make better systemic decisions. The concept of open data includes treating data as a public good, particularly for research purposes. Typically access to “open data” refers to access to *aggregated* data, or data where much of the identifiable information has been removed, although not in all cases. Open data access is intended to achieve two primary goals:

- Broaden the potential uses of data that has already been collected to maximize its potential impact; and
- Ensure that lack of resources or other barriers do not create inequitable access to data.⁶⁸

However, there are also ethical concerns related to open data access, including:

- *Privacy*: The more broadly data is shared, the more opportunities there are for subjects’ privacy to be compromised. Although removing personal information can help to minimize these concerns, advanced technology makes it difficult to eliminate this risk. This includes the increased potential for reidentification of individual information when information is released publicly.
- *Secondary data use*: By its nature, making data available to the public encourages secondary data uses. While these uses are generally for research and often only utilize aggregated data, the data is often available to the general public and can be misused.
- *Consent and transparency*: Depending on the type of data released and how broadly access to it is granted, it is often difficult to obtain consent from data subjects for all uses, or to be transparent about how it might be used.⁶⁹

- *De-identification and anonymization:*^{2**} As data sets get bigger, more data is available in the public realm, and matching software grows more advanced, guaranteeing that data will remain de-identified becomes increasingly difficult. Organizations will need to grapple with the degree to which de-identifying data, such that the information cannot be traced to individuals, is possible while still providing some level of access to the data.⁷⁰ Techniques such as differential privacy or federated learning may help to achieve these goals.
- *Data reporting:* Related to transparency, data reporting must be based on quality data and valid and reliable analyses, as well as displayed in ways that do not mislead the consumer of the data. One common issue that arises here is trying to simplify data for a range of audiences in a way that alters the conclusions or does not adequately discuss the limitations of the data.⁷¹

Actions for Leaders:

Leaders can take the following actions, based on emerging practices that attempt to balance the benefits and risks of open data access:

- Utilize clear governance practices that regulate who receives access and for what purpose, particularly for individual data records;
- Provide detailed explanations of the data, both raw and analyzed, to ensure potential users have a clear understanding of it;
- Remove the maximum amount of personally identifiable information possible while still maintaining the usefulness of the data; and
- When sharing with communities, include capacity-building activities to ensure understanding.⁷²

Privacy and Security

Privacy and security are fundamental issues to address in support of the goal of ethical and responsible data and technology use. Privacy is the idea that people should be able to control their own information, and that the entities that are authorized to collect and use that information do so in ways that respect an individual's autonomy.⁷³ Security is the practice of preventing unauthorized access to information and the systems that hold it.⁷⁴

** The terms "de-identification" and "anonymization" are sometimes used interchangeably, while in other contexts they have distinct (but related) meanings. In computer science, de-identification typically refers to the removal of personally identifiable information (PII) from a record (or the decoupling of PII from the rest of the record so that it can be re-identified later if needed), while anonymization is generally used to refer to additional, more complex methods such as shuffling or adding noise to the data, which can make recovery more difficult than simple de-identification.

Having fundamental practices and policies in place to ensure the privacy and security of sensitive data is essential to establishing trust and enabling organizations to build towards more advanced uses of data and technology.

Many organizations focus their privacy and security work on legal compliance at the federal and state level. However, these laws become outdated quickly, resulting in practices that do not go far enough in protecting privacy and have not evolved along with the types of data and technology being used. For example, the primary federal law that governs student privacy, the Family Educational Rights and Privacy Act, was passed in 1974. Other laws like the Federal Trade Commission Act permit most companies' data practices so long as they are disclosed in a company's terms of service, but users rarely read those notices, and they expect more. Therefore, organizations should consider how their privacy and security efforts need to go beyond legal compliance to truly protect individuals and their data.

Actions for Leaders:

To go beyond legal compliance and ensure the privacy and security of sensitive information, leaders should take the following actions:^{3***}

- Put in place a Chief Privacy Officer or someone in a similar role who is responsible for managing an organization's privacy responsibilities and compliance with legal requirements;
- Create detailed privacy policies that set forth how privacy and security will be addressed as the organization collects, shares, and uses data and technology;
- Tailor communications and materials to various audiences, including details for those who are interested/have subject matter expertise, as well as user-friendly information for those who do not have a background in data and technology;
- Minimize data collected and technology used;
- Enact policies and practices that govern data and technology, including minimizing how data is collected and shared, specifying when information is deleted, and prohibiting harmful uses of data even if consent is given;
- Provide regular training and support to staff on these issues; and
- Put in place processes to review changes in technology, policy, and the law.

*** CDT's resources to assist education leaders with protecting sensitive information can be found at cdt.org/civictech.

Consent

Gathering consent from data subjects raises key ethical issues. As data collection and use have become more complex and widespread, issues related to consent have also become more pronounced. Consent places the burden to protect information on the data subject and fails to prevent harmful data uses.⁷⁵ Additionally, in education, it is not always feasible to collect consent when schools are using data and technology to provide core services to students like transportation, nutrition, or even school funding. Therefore, it is important for organizations to adopt responsible data practices distinct from stakeholder consent, such as committing to practices for data minimization, data deletion, and prohibiting harmful uses of data even if consent is given.⁷⁶


That said, when user consent is collected, it should be done so thoughtfully, which entails addressing the following issues:

- *Duration of consent:* Now that data can be stored for long periods relatively cheaply, organizations must consider whether giving consent once means giving indefinite consent, or if consent should expire after a certain period of time.⁷⁷
- *Data uses:* As secondary data use has become more common, organizations must consider whether or not consent can be granted for multiple data uses, or how to go back to the individuals that provided the data to ask for consent for a new purpose, particularly if time has passed.⁷⁸
- *Lack of options:* As data use has become more prevalent, there are situations where consent is required, but if an individual were to opt out, they would not be able to access services. In these cases, consent is not a true choice.
- *Lack of expertise:* As data and data analysis become more complex, it grows harder for stakeholders, who will often lack expertise in data analysis, to understand how their data will be used. This issue can be exacerbated by poor communication during the consent-gathering process. Organizations can help counteract this concern by providing education as part of the process of requesting consent.⁷⁹

Actions for Leaders:

Consent is not sufficient to protect the rights of individuals, so leaders should take the following actions to manage its limitations:

- As mentioned in the *Privacy and Security* section, enact policies and practices that govern data and technology, including minimizing how data is collected and shared, specifying when information is deleted, and prohibiting harmful uses of data even if consent is given;

- When consent is not feasible or preferable, ensure transparency and proactively communicate to data subjects as to what data is being collected, how it is being used, and how it is being kept safe;
- Develop technical and organizational resources that facilitate “dynamic consent” by identifying potential short and long-term reuses for the research subject to consider, and ensure research design that stands up to challenges for ongoing informed consent;⁸⁰
- Rather than relying on prior consent, provide ongoing opportunities to review data and how it is being used; and
- Ensure other principles around responsible data use are in place to address the limitations of consent.⁸¹ 

**The
Deep Dive
on Data Ethics
in K-12 Education
*is on the next page.***



Deep Dive: Data Ethics in K-12 Education

While the use of data and technology in educational contexts has been growing steadily in recent years, there was an abrupt increase as COVID-19 forced most of the country into online learning. This accelerated the urgency of adopting data ethics principles in K-12 education for technology to be a trusted, safe, and effective tool to support students moving forward.

Defining data ethics, articulating why it matters, and applying best practices are important across all sectors, but there are specific considerations for the K-12 education sector.

What is Data Ethics? → *Additional Considerations for K-12 Education*

When thinking about how to define data ethics within K-12 education, considerations must include the protection of both student and family privacy and student and family rights in the use of data.⁸² The goal of data use by K-12 organizations is to improve student learning, but the related laws—including the Family Educational Rights and Privacy Act (FERPA)—are insufficient in the current context to ensure data is used responsibly, which creates the need for data ethics policy and practice.⁸³

Why Does Data Ethics Matter? → *Additional Considerations for K-12 Education*

- **EdTech Vendors:** The use of education technology (EdTech) vendors in the K-12 education space introduces particular data ethics concerns. These vendors may have greater expertise on technical issues than education organizations, but may have different priorities than education organizations themselves. Education organizations have an obligation to educate themselves so they can enter into vendor agreements in an informed manner.⁸⁴
- **View of Education as a Public Good:** While reputation and public trust is important in all sectors, the public's perception of K-12 education makes it especially important. A breach of trust in school management, including as it relates to data, can have harmful impacts. This can include further restricting the collection, sharing, and use of data that could otherwise be used to assist students and families; staff turnover and school executives losing their jobs are other potential repercussions of breaches of public trust.⁸⁵

What Are Key Data Ethics Issues, and What Can Leaders Do About Them? → *Additional Considerations for K-12 Education*

- **Transparency:** In K-12 education, transparency should be primarily aimed at achieving three objectives:
 - *Making parents and the public aware of data practices.* School systems should proactively communicate about student data being collected and how that

information is used, stored, and deleted. This is true of educational organizations (e.g., school districts) as well as third-party vendors. When transparency has not been prioritized, stakeholders have been surprised by and unhappy with how schools were sharing data with outside parties, leading to a backlash against the collection of data.⁸⁶

- o *Reporting of student outcomes.* Organizations often report data in a manner that supports the claims they are making in a given moment, but fails to be fully transparent about the true implications of the results.⁸⁷
- o *Explaining and clarifying algorithmic systems.* K-12 organizations often rely on vendors for algorithms and other analyses, but these are not transparent to the organizations and certainly not to the community.⁸⁸
- **Capacity-Building:** Teachers often do not receive training on data and technology, even though they are the ones collecting and interacting with student data. In fact, recent research indicates that almost half of teachers have not received substantive training on student privacy.⁸⁹ In addition, because K-12 education sees high rates of staff turnover and often has decentralized decision-making structures, documentation of policies and procedures is vital to avoid knowledge loss during these transitions.⁹⁰
- **Equity and Bias:** Mitigating bias in K-12 education requires an analysis of at least two specific areas:
 - o *Predictive analytics.* Defined as “using massive amounts of historical data to predict future events,”⁹¹ an example of predictive analytics that might be used in K-12 education is an early warning system that identifies students who may be at risk of academic failure or dropping out based on past history and demographic characteristics.⁹² As described in the *Equity and Bias* section, algorithmic systems have the potential to exacerbate inequities if these risks are not mitigated.
 - o *Discipline and special education data.* Certain data sets have been shown to reflect systemic inequities, with discipline and special education data reflecting an over-identification of marginalized students, including students of color.⁹³ Although sharing this data can help teachers and schools best serve students, it can also bias teachers and limit opportunity for students, through tracking or other actions taken by teachers and staff.⁹⁴
- **Governance and Oversight:** Data governance in K-12 education is required both for data that stays within the K-12 sector as well as data that is shared across systems and sectors.
 - o *Longitudinal and cross-sector.* Data governance can support ethical data use for longitudinal data systems that track students over time, as well as systems for sharing education data with other agencies. This often requires a governance body that has representatives from districts, schools, and state agencies.
 - o *Data use.* School districts should have a group that is responsible for data governance internally to ensure that student data is used responsibly.⁹⁵

- **Stakeholder Engagement:** In K-12 education, stakeholder engagement should include under-represented perspectives, like teachers, parents, and students, on issues related to student-level information as well as broader system-wide data.⁹⁶
- **Secondary Data Use:** In K-12 education, secondary data use most often arises with student information systems, where the district or school uses the data to inform its operations while the company that created the system also uses the data to create product improvements.⁹⁷ As a result, school leaders should consider all issues related to secondary data use, including but not limited to data quality, consent, and stakeholder engagement, before using data that was collected for one purpose for something else.

Data Ethics Dynamics That Are Consequential to K-12 Education

There are some dynamics that are specific to K-12 education and span multiple data ethics issues. Such dynamics that should inform data ethics policies and practices include:

- **Education as a Public Good:** K-12 education is considered a public good, so there can be concerns about privatizing K-12 education, especially regarding companies like EdTech vendors. In addition, the use of data to drive decisions is seen as analogous to the for-profit world and therefore can be an unwelcomed addition to the sector.⁹⁸
- **Children:** K-12 education is unique in that the primary providers of much of the data are minors. This raises additional questions of privacy and consent and adds an additional stakeholder (parents and guardians) to engage, which increases complexity.⁹⁹
- **Decentralized Decision-Making:** Much of the day-to-day decision-making in K-12 education happens at the school or classroom level. While this can have benefits, it can create challenges in responsible data use due to the number of users, different systems, varied training needs, etc.
- **Student Mobility:** Students, especially the most vulnerable, change schools frequently, which has led to efforts to find ways to share information with new schools to inform their education, like integrating data systems and creating data sharing agreements. However, this creates privacy and security risks, as well as the potential for bias if data is used out of context, so it must be implemented responsibly and securely.¹⁰⁰
- **Online Learning:** K-12 education has integrated online learning in recent years, including fully online schools, virtual tutoring, and online courses. These have expanded significantly since the shift to distance learning as a result of the global pandemic. Online learning raises concerns of bias and activity tracking, particularly if the system has an algorithmic decision-making component. For these systems, it is important to ensure transparency to the community and hold vendors accountable through contracts.¹⁰¹

- **Discipline:** Much of the concern around data sharing in K-12 education relates to discipline data, as it raises concerns of bias where students may be treated differently based on their discipline history. Clear data deletion and user access policies can help to address these concerns.¹⁰²
- **Resource Constraints:** K-12 organizations tend to have significant resource constraints and often do not prioritize issues of data ethics in their resource allocations, which can lead to less ethical practices, signal to stakeholders that these issues are not a priority, and lead to mistrust. It can also alter the power dynamics between K-12 organizations and EdTech vendors, as they do not have the resources to develop the expertise needed to more equitably interact with vendors.¹⁰³
- **Non-Student Data:** The focus of this document is student data, but there are other types of data in K-12 education where issues of data ethics arise, like teacher and school finance data.¹⁰⁴

Conclusion and Next Steps

As data and technology use becomes increasingly prevalent in education and other social service settings, organizations must focus on the responsible use of data. The field of data ethics is continuously evolving, which requires those entrusted to use data and technology to change as well. Decision-makers who may not consider themselves technical or data experts are nevertheless being charged with decisions that can have a significant impact on people's privacy, or that risk perpetuating inequality through poor digital design. Organizations and agencies must equip themselves to meet this challenge, recognizing the ways in which data and technology can be both used and misused, and taking proactive steps to avoid potential risks.

The following sections set forth specific considerations for decision-makers in the settings of higher education, the social sector (human services, housing and other government service areas designed to support individuals), and other emerging technologies.

For more information from CDT about core issues explored in this paper, visit cdt.org/civictech.

Appendix A: Data Ethics in Higher Education

Defining data ethics, articulating why it matters, and applying best practices are important in all sectors, but there are specific considerations for the higher education sector.

What Is Data Ethics? *Additional Considerations for Higher Education*

Data ethics considerations within higher education should be focused on promoting student learning, either through using data in support of organizational operations or directly in support of student learning.¹⁰⁵

Why Does Data Ethics Matter? *Additional Considerations for Higher Education*

- **EdTech Vendors:** Similar to K-12 education, higher education often relies on EdTech vendors, and therefore these organizations must ensure that they have the expertise to develop contracts with vendors that protect the interests of both students and the organizations.¹⁰⁶
- **Fiscal Sustainability:** Higher education organizations increasingly rely on advanced data and technology to support efforts related to recruitment and enrollment. However, they also run the risk of these efforts undermining public trust and their reputations, and thereby fiscal sustainability.¹⁰⁷

What Are Key Data Ethics Issues, and What Can Leaders Do About Them? *Additional Considerations for Higher Education*

- **Transparency:** In higher education, transparency should be primarily aimed at two issues:
 - *Predictive analytics.* Students are often unaware of algorithms and early warning systems, which are also used in K-12 education, and how these systems are making recommendations and decisions. In many cases, even the universities do not fully understand the algorithms used because they are held by the vendors, which raises obstacles to transparency or judgment.¹⁰⁸
 - *Organizational operations:* Much student data is used to manage the organization and inform enrollment decisions, but this is not made clear to students who believe data is only being used to support their academics.¹⁰⁹
- **Capacity-Building:** Similar to the K-12 education sector, there is a disconnect between those creating the data and those responsible for improving student outcomes. Capacity-building efforts in higher education need to address the relationship with vendors and the concentration of data knowledge in a small number of staff who do not directly interact with students.¹¹⁰

- **Equity and Bias:** Challenges related to supporting fair and equitable data use include the intersection between student outcomes and the financial sustainability of the organizations. For example, higher education organizations may lower their expectations for groups of students who have been traditionally underserved and steer them towards less challenging programs to increase their chances of completion (and therefore increased tuition payments), even if they would be successful on the traditionally more challenging paths. Algorithms can exacerbate these inequities by utilizing demographic or past performance information about students to make decisions about them, resulting in students being challenged and, in some cases, encouraged to drop out, which is at odds with ensuring all students receive a high-quality education.¹¹¹
- **Governance and Oversight:** Data governance in higher education should focus on at least three issues:
 - Oversight of how data is shared between operational and academic departments as well as with outside vendors;¹¹²
 - Management of how algorithmic systems and predictive analytics are used;
 - Secondary data use when information is collected for academic purpose but then used to inform university operations.¹¹³
- **Stakeholder Engagement:** Stakeholder engagement in higher education is often focused on students but should also include university staff.¹¹⁴

Data Ethics Dynamics That Are Consequential to Higher Education

Dynamics that are specific to higher education, and span multiple data ethics issues, should inform data ethics policies and practices:

- **Predictive Analytics:** Although algorithmic and early warning systems are used in K-12 education, they are more common in higher education.¹¹⁵ Higher education organizations use predictive analytics in three primary ways: selection of students for academic advising, personalized learning, and enrollment management.¹¹⁶
- **Online Learning:** Like K-12 education, higher education has seen an increased prevalence of online learning tools in recent years, which raises questions about activity tracking and potential bias as well as transparency, as often students do not fully understand how their information is being used.¹¹⁷
- **Operational Data:** Often, higher education organizations collect data for academic purposes like advising, but then also use it for operational needs such as enrollment management. This raises questions of transparency, as students are not made aware of this use of their data and have not given informed consent. This lack of transparency can also lead to bias, as it can impact the types of student schools recruit and enroll.¹¹⁸

- **Research Function:** Most large universities have well-established research functions that include an Institutional Review Board structure that is aligned with corresponding human subject research guidelines. However, they typically only apply to traditional human subject research, so similar structures and guidelines are needed for research driven by more complex data and technology.¹¹⁹
- **Financial Aid:** Higher education organizations collect a large amount of financial data on students and their families to inform financial aid packages, so this personal information should be kept secure and only used it for its intended purpose.¹²⁰

Appendix B: Data Ethics in the Social Sector

Defining data ethics, articulating why it matters, and applying best practices holds true across all sectors, but there are specific considerations for the social sector, defined here as human services, housing, and other government service areas designed to support individuals.

What Is Data Ethics? *Additional Considerations for the Social Sector*

Data ethics in the social sector should focus on ensuring that the use of data and technology is supporting the most disadvantaged groups, given that its purpose is to support individuals. Social sector public agencies should ensure that data and technology are used ethically and focus on the individual good and collective benefit. Doing so can change the historical paradigm in the social sector, in which data has sometimes been viewed as being used against the community as opposed to helping it.¹²¹

Why Does Data Ethics Matter? *Additional Considerations for the Social Sector*

- **Power Dynamics:** Although other sectors may have a power imbalance between organizations and their users, these situations are especially prevalent in the social sector where organizations are most often working with groups that are economically or otherwise disadvantaged or vulnerable. The power imbalance between these groups, who are most often the ones providing the data, and those using the data makes it especially important that data ethics issues are addressed, especially related to equity.¹²²
- **Sensitivity of Data:** Much of the data collected and used in the social sector is especially sensitive and can cause significant damage to an individual if not protected. This includes income and other financial data, criminal records, and healthcare data.¹²³
- **Data Integration:** It is becoming more common for public agencies within the social sector to integrate data across sectors. Given the complexities of cross-agency data sharing, it is especially important to address data ethics as risks can scale and spread significantly when integrating data sets.¹²⁴
- **Due Process Obligations:** Public agencies that use data and technology, including algorithm-driven decision-making, to provision services and benefits are at risk of violating constitutional or statutory due process rights. Individuals are entitled to notice, a right to challenge decisions, decision-making that is not arbitrary, and ascertainable standards in a decision regarding their government-issued benefits.¹²⁵

What Are Key Data Ethics Issues, and What Can Leaders Do About Them? *Additional Considerations for the Social Sector*

- **Transparency:** In the social sector, transparency should prioritize the following issues:

- o Sharing data analyses with the providers of data, so they can use it to make informed decisions and improve their lives
 - o Communicating efforts to integrate data across agencies, especially to the providers of the data¹²⁶
 - o Building on human subject research guidelines within the social sector to ensure individual protections and transparency¹²⁷
 - o Ensuring that algorithms that are used within the social sector are explained and communicated¹²⁸
- **Capacity-Building:** Capacity-building in the social sector should be focused on educating the community about how their data is being used and how they might be able to utilize this data to their advantage. Social sector data that is publicly available, especially data like the Census collected by the government, should be accompanied by a public education campaign.¹²⁹
 - **Equity and Bias:** In the social sector, algorithms—such as those used in predictive policing, bail reform, housing opportunities, credit scoring, and employment practices—are put in place, in part, to automate processes due to resource constraints; however, they can exacerbate inequities. Reasons for potential bias in the social sector include historic inequities of the system, changing policies that are not embedded in the algorithms, and power dynamics inherent in the sectors. These biases can lead to discrimination in all of these areas, which serves to only exacerbate systemic inequity rather than address them.¹³⁰
 - **Governance and Oversight:** Because data sharing across agencies is becoming more common in the social sector, leaders should establish a data governance body that includes representatives from all of the agencies to establish rules for data sharing and ensure the rules are followed appropriately.¹³¹
 - **Secondary Data Use:** In the social sector, secondary data use is often related to research or when official government data (i.e., Census data) is used for additional purposes and potentially beyond its original intent.¹³²

Data Ethics Dynamics That Are Consequential to the Social Sector

Dynamics that are specific to the social sector, and span multiple data ethics issues, should inform data ethics policies and practices:

- **Surveillance:** In criminal justice, much of the conversation is around the appropriate level of surveillance, both of physical and online activity, and how this data can be used.¹³³
- **Resource Constraints:** Much of the social sector operates under resource constraints and consequently seeks alternative uses of data and technology to automate processes.

In addition, the government is seeking alternatives to current costly data collection processes, including components on the Census, that would use potentially comparable data but could contain inaccuracies and be used in ways not originally intended.¹³⁴

- ***Use of Alternative Data:*** Many of the algorithms and decision-making processes in both criminal justice and financial services rely on a limited set of historically available data points; however, there are more varied types of data currently available that could be integrated into these systems to limit the potential biases and make them more representative.¹³⁵
- ***Short and Long-Term Impact:*** Although concerns of misuse exist in all sectors, using data irresponsibly in certain social sectors could have dramatic effects on an individual, including in child welfare, immigration, and criminal justice. Therefore, individuals interacting with the social sector should have greater control and agency over decisions regarding their data, and to do this, the social sector must educate individuals and ensure transparency.¹³⁶
- ***Perception of Organizations:*** Many of the organizations working with data in the social sector are viewed unfavorably by the communities in which they work, making issues such as transparency and stakeholder engagement, as well as broader issues of data ethics, especially important.¹³⁷

Appendix C: Emerging Technologies

While many of the issues articulated throughout this paper are relevant to emerging technologies, this section articulates specific emerging technologies that raise issues of data ethics and have applications in education and the social sector.

- **Predictive Analytics:** Predictive analytics utilize data science techniques and platforms to make predictions on the basis of disparate data sets, typically using an algorithmic model trained on integrated learning data sets, to chart a student or other individual's course. An example is early warning systems.¹³⁸ Some ethical concerns related to the use of such models include potential bias and secondary data use.
- **Algorithmic Tools:** In addition to predictive analytics, other uses of algorithms include virtual tutors and automated test grading and other review processes, increasing the risk of bias and other ethical issues.¹³⁹
- **Data Integration:** Data integration is discussed throughout this paper, but also important as a standalone issue is the extent to which technological advances are enabling a higher degree of integration than ever before. This raises a number of issues, including (but not limited to) access, purpose, and storage. However, it is important to note that siloed data can also cause ethical issues.¹⁴⁰
- **Virtual Reality:** Virtual reality is not a new technology but is only beginning to be used in education and the social sector for things such as remote teaching and experiential learning. Some ethical issues that arise include privacy (e.g., when student experience with the technology is shared with the vendor or other third parties) and inequitable access.¹⁴¹
- **Activity Tracking:** As the ability to track student behavior increases, both online and offline, additional ethical issues about the benefits and risks of collecting and analyzing this data arise.¹⁴²
- **Facial and Other Physical Recognition and Tracking:** As the technology to recognize individuals' physical characteristics grows more effective, there is an increased ability to track movements and behavior. One common use of this has been related to school safety, raising issues of bias (as these systems are not always equally effective across different demographic groups) and privacy.¹⁴³
- **Surveillance:** Related to facial recognition, cameras and other technology have become cost-effective and high-quality enough to allow organizations to constantly surveil students and others, raising concerns about privacy and how this data might be used.¹⁴⁴ False positive results in surveillance technologies can strain resources and increase the potential for discriminatory outcomes.

Endnotes

1. Laura Fay, “Ambitious Research Project — to Review How Every School in America Responded to COVID-19 — Aims to Deliver Its First Findings in Early July,” June 26, 2020, <https://www.the74million.org/article/ambitious-research-project-to-review-how-every-school-in-america-responded-to-covid-19-aims-to-deliver-its-first-findings-in-early-july/>.
2. Mike Montgomery, “The Pandemic Should Have Been Edtech’s Moment To Shine. So Far, It Hasn’t Been,” accessed August 7, 2020, <https://www.forbes.com/sites/mikemontgomery/2020/06/04/the-pandemic-should-have-been-edtechs-moment-to-shine-so-far-it-hasnt-been/#63b7959149f3>.
3. Ruth Reader, “A school mandated that students wear a COVID-detecting ‘BioButton.’ They fought back,” *Fast Company*, August 2020, <https://www.fastcompany.com/90537201/why-these-students-fought-back-against-their-universities-covid-19-program>.
4. Elizabeth Laird, “Protecting Students’ Privacy and Advancing Digital Equity,” (CDT, October 2020), <https://cdt.org/insights/research-report-protecting-students-privacy-and-advancing-digital-equity/>.
5. Joseph Jerome and Natasha Duarte, “Towards a Data Ethic for the Public Interest,” (CDT, 2017), <https://www.law.berkeley.edu/research/bclt/past-events/june-2017-10th-annual-privacy-law-scholars-conference-plsc/agenda-plsc-2017/>.
6. Andrew Cormack, “A Data Protection Framework for Learning Analytics,” *Journal of Learning Analytics* 3, no. 1 (2016): 91–106.
7. M Knight, “What Are Data Ethics?,” *DataVersity*, December 22, 2017, <https://www.dataversity.net/what-are-data-ethics/>.
8. Brian Bollier, “The Promise and Peril of Big Data,” (Aspen Institute, 2010).
9. Montgomery, “The Pandemic Should Have Been Edtech’s Moment To Shine. So Far, It Hasn’t Been.”
10. Sharon Slade and Paul Prinsloo, “Learning Analytics: Ethical Issues and Dilemmas,” *American Behavioral Scientist* 57, no. 10 (October 1, 2013): 1510–29, <https://doi.org/10.1177/0002764213479366>.
11. White House Big Data and Privacy Working Group, “Big Data: Seizing Opportunities, Preserving Values,” Interim Progress Report (White House Big Data and Privacy Working Group, February 2015).
12. Jacob Metcalf, Emily Keller, and danah boyd, “Perspectives on Big Data, Ethics, and Society,” (The Council for Big Data, Ethics, and Society, n.d.).
13. Hannah Quay-de la Vallee and Natasha Duarte, “Algorithmic Systems in Education,” A Series of Papers on Student Privacy (Center for Democracy and Technology, 2019).
14. Andrea Alarcon et al., “Data & Civil Rights: Education Primer,” Data & Civil Rights (Data & Society Research Institute, October 30, 2014), <http://www.datacivilrights.org/pubs/2014-1030/Education.pdf>.

15. Cormack, “A Data Protection Framework for Learning Analytics.”
16. Slade and Prinsloo, “Learning Analytics.”
17. David Kay, Naomi Korn, and Charles Oppenheim, “Legal, Risk and Ethical Aspects of Analytics in Higher Education,” *CETIS Analytics Series* 1, no. 6 (n.d.).
18. Neil Richards and Jonathan King, “Big Data Ethics,” *Wake Forest Law Review* 49 (May 2014): 393–432.
19. Daniel Castro and Travis Korte, “Parents and Educators Should Embrace, Not Fear Student Data” (Center for Data Innovation, December 3, 2013).
20. Consortium for School Networking, “Student Data Privacy Roadmap,” (n.d.), <https://trustedlearning.org/framework/>.
21. Accenture, “The Ethics of Data” (Accenture, 2016).
22. “Technological School Safety Initiatives: Considerations to Protect All Students,” accessed February 1, 2021, <https://www.brennancenter.org/sites/default/files/analysis/20190524schoolsafety.pdf>.
23. M Rye, “The 4 Pillars of Responsible Use of Nonprofit Data,” *NTen*, August 13, 2019.
24. R Bean, “A Rising Crescendo Demands Data Ethics and Data Responsibility,” *Forbes*, October 29, 2018.
25. A Gauss, “Why We Need Data Ethics to Do Good,” *Classy* (blog), (n.d.), <https://www.classy.org/blog/why-we-need-data-ethics/>.
26. Elizabeth Laird and Hannah Quay de Valle, “Protecting Privacy While Supporting Students Who Change Schools” (CDT, 2019).
27. Corey Chatis and Kathy Gosa, “Communicating the Value of Data Governance,” SLDS Issue Brief (Institute of Education Sciences, n.d.).
28. C Towers-Clark, “The Ethics of Data Governance - Data Comes with Benefits And Liabilities,” *Forbes*, January 23, 2019.
29. Joanna Grama, “Protecting Privacy and Information Security in a Federal Postsecondary Student Data System” (Institute for Higher Education Policy, 2019).
30. Brian Bollier, “The Promise and Peril of Big Data” (Aspen Institute, 2010).
31. Brian Bollier, “The Promise and Peril of Big Data” (Aspen Institute, 2010).
32. John Fantuzzo et al., “The Integrated Data System Approach: A Vehicle to More Effective and Efficient Data-Driven Solutions in Government,” *Actionable Intelligence for Social Policy* (University of Pennsylvania, 2017).
33. Chatis and Gosa, “Communicating the Value of Data Governance.”
34. White House Big Data and Privacy Working Group, “Big Data: Seizing Opportunities, Preserving Values.”
35. Omer Tene and Jules Polonetsky, “Big Data for All: Privacy and User Control in the Age of Analytics,” *Northwestern Journal of Technology and Intellectual Property* 11, no. 5 (April 2013): 240–72.
36. Brian Bollier, “The Promise and Peril of Big Data,” (Aspen Institute, 2010).
37. Joel Reidenberg et al., “Privacy and Cloud Computing in Public Schools,” *FLASH: The Fordham Law Archive of Scholarship and History*, Book 2 (2013).
38. Taddeo, “Data Philanthropy and the Design of the Infraethics for Information Societies.”

39. UCLA Data Governance Task Force, “Data Governance Task Force Final Report and Recommendations,” (UCLA, May 2016).
40. Data Quality Campaign, “Roadmap for Cross-Agency Data Governance,” Quality Implementation Roadmaps (Data Quality Campaign, January 30, 2018).
41. Elizabeth Laird, “Responsible Use of Data and Technology in Education: Community Engagement to Ensure Students and Families Are Helped, Not Hurt,” (CDT, February 2021),
<https://cdt.org/insights/responsible-use-of-data-and-technology-in-education-community-engagement-to-ensure-students-and-families-are-helped-not-hurt/>.
42. Ben Green and Lily Hu, “The Myth in the Methodology: Towards a Recontextualization of Fairness in Machine Learning,” in *Machine Learning: The Debates Workshop at the 35th International Conference on Machine Learning*, 2018.
43. National Forum on Education Statistics, “The Forum Guide to Data Ethics,” 2010.
44. National Forum on Education Statistics, “The Forum Guide to Data Ethics,” 2010.
45. Andrea Alarcon et al., “Data & Civil Rights.”
46. Corey Mitchell, “High Stakes for Schools If 2020 Census Undercounts Latino Families,” Education Week, November 22, 2019,
<https://www.edweek.org/policy-politics/high-stakes-for-schools-if-2020-census-undercounts-latino-families/2019/11>.
47. Bean, “A Rising Crescendo Demands Data Ethics and Data Responsibility.”
48. Cormack, “A Data Protection Framework for Learning Analytics.”
49. F. Chris Curran, “‘Early Warning’ Systems in Schools Can Be Dangerous in the Hands of Law Enforcement,” *The Conversation*, accessed February 1, 2021,
<http://theconversation.com/early-warning-systems-in-schools-can-be-dangerous-in-the-hands-of-law-enforcement-152701>.
50. Hannah Quay-de la Vallee and Natasha Duarte, “Algorithmic Systems in Education,” A Series of Papers on Student Privacy, (Center for Democracy and Technology, 2019).
51. Copeland, “10 Principles for Public Sector Use of Algorithmic Decision Making.”
52. Chen and Liu, “Big Data Ethics in Education: Connecting Practices and Ethical Awareness.”
53. Dillon Reisman et al., “Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability,” (AINow, April 2018).
54. Brian Bollier, “The Promise and Peril of Big Data,” (Aspen Institute, 2010).
55. David Robinson, “The Challenges of Predictions: Lessons from Criminal Justice,” *Journal of Law and Policy for the Information Society* 151 (2018).
56. Jules Polonetsky and Omar Tene, “The Ethics of Student Privacy: Building Trust for EdTech,” *International Review of Information Ethics*, 21 (July 2014).
57. Xiaojon Chen and Chen Ying Liu, “Big Data Ethics in Education: Connecting Practices and Ethical Awareness,” *Journal of Educational Technology Development and Exchange* 8, no. 2 (2015).
58. Leah Plunkett, Alicia Solow-Niederman, and Urs Gasser, “Framing the Law and Policy Picture: A Snapshot of K-12 Cloud Based Ed Tech and Student Privacy in Early 2014,” Research Publication, Student Privacy Initiative (The Berkman Center for Internet & Society at Harvard University, June 3, 2014).

59. Eddie Copeland, “10 Principles for Public Sector Use of Algorithmic Decision Making,” *Nesta* (blog), February 20, 2018, <https://www.nesta.org.uk/blog/10-principles-for-public-sector-use-of-algorithmic-decision-making/>.
60. Rye, “The 4 Pillars of Responsible Use of Nonprofit Data.”
61. John Fantuzzo et al., “The Integrated Data System Approach: A Vehicle to More Effective and Efficient Data-Driven Solutions in Government,” *Actionable Intelligence for Social Policy* (University of Pennsylvania, 2017).
62. Iris Palmer, “Choosing A Predictive Analytics Vendor,” (New America, 2018).
63. National Forum on Education Statistics, “The Forum Guide to Data Ethics.”
64. National Forum on Education Statistics, “The Forum Guide to Data Ethics.”
65. Accenture, “Building Digital Trust: The Role of Data Ethics in the Digital Age.”
66. Brad Wheeler, “Who Is Doing Our Data Laundry?,” *EduCause Review*, March 13, 2017.
67. Chatis and Gosa, “Communicating the Value of Data Governance.”
68. R Bean, “A Rising Crescendo Demands Data Ethics and Data Responsibility,” *Forbes*, October 29, 2018.
69. Accenture, “Building Digital Trust: The Role of Data Ethics in the Digital Age,” (Accenture, 2016).
70. Richards and King, “Big Data Ethics.”
71. Laura Jensen and Vanessa Roof, “The Ethical Use of Student Data and Analytics,” (The Reinvention Center, 2015).
72. Accenture, “The Ethics of Data” (Accenture, 2016).
73. Elizabeth Laird, “Chief Privacy Officers: Who They Are and Why Education Leaders Need Them,” (CDT, 2019), <https://cdt.org/insights/chief-privacy-officers-who-they-are-and-why-education-leaders-need-them/>.
74. Elizabeth Laird, “Chief Privacy Officers: Who They Are and Why Education Leaders Need Them,” (CDT, 2019), <https://cdt.org/insights/chief-privacy-officers-who-they-are-and-why-education-leaders-need-them/>.
75. Michelle Richardson et al., “Americans Deserve a Law Protecting Their Digital Privacy – Here’s Our Proposal,” (CDT, December 2018), <https://cdt.org/insights/americans-deserve-a-law-protecting-their-digital-privacy-heres-our-proposal/>.
76. Woodrow Hartzog and Evan Selinger, “Big Data in Small Hands,” *Stanford Law Review Online* 66, no. 81 (September 3, 2013).
77. Cormack, “A Data Protection Framework for Learning Analytics.”
78. Jensen and Roof, “The Ethical Use of Student Data and Analytics.”
79. Kate Crawford and Jason Schultz, “Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms,” *SSRN Scholarly Paper* (Rochester, NY: Social Science Research Network, October 1, 2013), <https://papers.ssrn.com/abstract=2325784>.

80. Isabelle Budin-Ljøsne et al., “Dynamic Consent: A Potential Solution to Some of the Challenges of Modern Biomedical Research,” *BMC Medical Ethics* 18, no. 1 (January 25, 2017): 4, <https://doi.org/10.1186/s12910-016-0162-9>.
81. Cormack, “A Data Protection Framework for Learning Analytics.”
82. National Forum on Education Statistics, “The Forum Guide to Data Ethics.”
83. Priscilla Regan and Jane Bailey, “Big Data, Privacy and Education Applications,” 2018.
84. CJ Hoofnagle, “EdTech Promise and Peril,” (TLBS, Istanbul, 2016), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3359205.
85. Charles Fadel, Wayne Holmes, and Maya Bialik, “The Social Consequences of AI in Education,” in *Artificial Intelligence In Education: Promises and Implications for Teaching and Learning* (Center for Curriculum Redesign, 2019), 136–45.
86. National Forum on Education Statistics, “The Forum Guide to Data Ethics.”
87. CCSSO, “Student Privacy and Data Security,” (CCSSO, 2016), <https://ccsso.org/resource-library/student-privacy-and-data-security-state-education-agency-discussion-framework>.
88. “Robot Teaching, Pedagogy and Policy,” in *The Oxford Handbook of Ethics of AI* (2019: Oxford University Press, n.d.).
89. Elizabeth Laird, “Research Report: Protecting Students’ Privacy and Advancing Digital Equity,” (CDT, October 2020), <https://cdt.org/insights/research-report-protecting-students-privacy-and-advancing-digital-equity/>.
90. CCSSO, “Student Privacy and Data Security.”
91. Iris Palmer and Manuela Ekowo, “The Promise and Peril of Predictive Analytics In Higher Education,” (New America, 2016).
92. Jill Barshay and Sasha Aslanian, “Colleges Are Using Big Data to Track Students in an Effort to Boost Graduation Rates, but It Comes at a Cost,” *Hechinger Report*, August 6, 2019.
93. Nora Gordon, “Disproportionality in Student Discipline: Connecting Policy to Research” (Brookings, 2018)
94. Elizabeth Laird and Hannah Quay-de la Valle, “Protecting Privacy While Supporting Students Who Change Schools,” (CDT, 2019).
95. Chatis and Gosa, “Communicating the Value of Data Governance.”
96. Plunkett, Solow-Niederman, and Gasser, “Framing the Law and Policy Picture: A Snapshot of K-12 Cloud Based Ed Tech and Student Privacy in Early 2014.”
97. National Forum on Education Statistics, “The Forum Guide to Data Ethics.”
98. A Giambone, “When Big Data Meets the Blackboard,” *The Atlantic*, June 22, 2015.
99. Priscilla Regan and Jolene Jesse, “Ethical Challenges of Edtech, Big Data and Personalized Learning: 21st Century Student Sorting and Tracking,” *Ethics and Information Technology* 21, no. 3 (2019).
100. Laird and Quay-de la Valle, “Protecting Privacy While Supporting Students Who Change Schools.”
101. Regan, Jesse, and Khwaja, “Big Data in Education: Developing Policy for Ethical Implementation in the US and Canada.”

102. Priscilla Regan, “Ethical and Administrative Policy Concerns about Use of Big Data in K-12 Education,” 2017,
https://www.ftc.gov/system/files/documents/public_comments/2017/11/00022-141722.pdf.
103. National Forum on Education Statistics, “The Forum Guide to Data Ethics.”
104. Matt Kasman and Jon Valant, “The Opportunities and Risks of K-12 Student Placement Algorithms,” A Blueprint for the Future of AI (Brookings Institution, February 28, 2019).
105. James Willis, “Ethics, Big Data and Analytics: A Model for Application,” *EduCause Review*, May 2013,
<https://er.educause.edu/articles/2013/5/ethics-big-data-and-analytics-a-model-for-application>.
106. Palmer, “Choosing A Predictive Analytics Vendor.”
107. Barshay and Aslanian, “Colleges Are Using Big Data to Track Students in an Effort to Boost Graduation Rates, but It Comes at a Cost.”
108. A Brooker, “Defining Data in Conversations with Students about the Ethical Use of Learning Analytics,” 2017,
<http://2017conference.ascilite.org/program/defining-data-in-conversations-with-students-about-the-ethical-use-of-learning-analytics/>.
109. John Campbell, Peter Deblois, and Diana Oblinger, “Academic Analytics: A New Tool for a New Era,” *Educause Review*, August 2007.
110. Sharon Slade, “Applications of Student Data in Higher Education: Issues and Ethical Considerations,” (ITHaka S+R, September 6, 2016).
111. Iris Palmer and Manuela Ekowo, “Predictive Analytics in Higher Education: Five Guiding Practices for Ethical Use,” (New America, 2017).
112. Slade and Prinsloo, “Learning Analytics.”
113. J Johnson, “The Ethics of Big Data in Higher Education,” *International Review of Information Ethics*, July 2014, <http://www.i-r-i-e.net/inhalt/021/IRIE-021-Johnson.pdf>.
114. Palmer and Ekowo, “Predictive Analytics in Higher Education: Five Guiding Practices for Ethical Use.”
115. Ronald Yanosky, “The Analytics Landscape in Higher Education, 2015,” (EduCause Center for Analysis and Research, October 2015).
116. Palmer and Ekowo, “The Promise and Peril of Predictive Analytics In Higher Education.”
117. Martin Kurzweil and Mitchell Stevens, “Setting the Table: Responsible Use of Student Data in Higher Education,” *EduCause Review*, June 2018.
118. J Johnson, “The Ethics of Big Data in Higher Education,” *International Review of Information Ethics*, July 2014, <http://www.i-r-i-e.net/inhalt/021/IRIE-021-Johnson.pdf>.
119. Slade, “Applications of Student Data in Higher Education: Issues and Ethical Considerations.”
120. Palmer and Ekowo, “The Promise and Peril of Predictive Analytics In Higher Education.”
121. Moish Kutnoski, “The Ethical Dangers and Merits of Predictive Policing,” *Journal of Community Safety and Well-Being* 2, no. 1 (March 2017): 13–17.
122. Alexa Hasse et al., “Youth and Artificial Intelligence: Where We Stand,” Research Publication (Berkman Klein Center for Internet & Society at Harvard University, May 24, 2019).
123. Paul Stiles and Boothroyd, Roger, “Ethical Use of Administrative Data for Research Purposes,” (Actionable Intelligence and Social Policy, 2015),

https://www.aisp.upenn.edu/wp-content/uploads/2015/09/0033_12_SP2_Ethical_Admin_Data_001.pdf.

124. Fantuzzo et al., “The Integrated Data System Approach: A Vehicle to More Effective and Efficient Data-Driven Solutions in Government.”
125. Lydia Brown, Michelle Richardson et al., “Challenging the Use of Algorithm-Driven Decision-Making in Benefits Determinations Affecting People with Disabilities,” (CDT, 2020), <https://cdt.org/wp-content/uploads/2020/10/2020-10-21-Challenging-the-Use-of-Algorithm-driven-Decision-making-in-Benefits-Determinations-Affecting-People-with-Disabilities.pdf>.
126. Fantuzzo et al., “The Integrated Data System Approach: A Vehicle to More Effective and Efficient Data-Driven Solutions in Government.”
127. Paul Stiles and Boothroyd, Roger, “Ethical Use of Administrative Data for Research Purposes.”
128. Upturn, “Civil Rights, Big Data, and Our Algorithmic Future” (Upturn, 2014).
129. Hermann Habermann, “Ethics, Confidentiality, and Data Dissemination,” n.d.
130. John Logan Koepke and David Robinson, “Danger Ahead: Risk Assessment and the Future of Bail Reform,” *Washington Law Review* 93 (September 2017).
131. Fantuzzo et al., “The Integrated Data System Approach: A Vehicle to More Effective and Efficient Data-Driven Solutions in Government.”
132. Upturn, “Civil Rights, Big Data, and Our Algorithmic Future.”
133. Upturn, “Civil Rights, Big Data, and Our Algorithmic Future.”
134. Siu-Ming Tam and Jae-Kwang Kim, “Big Data Ethics and Selection Bias,” *Statistical Journal of the IAOS* 34 (2018): 577–88.
135. Robinson, “The Challenges of Predictions: Lessons from Criminal Justice.”
136. Aaron Rieke, Miranda Bogen, and David Robinson, “Public Scrutiny of Automated Decisions,” (Omidyar and Upturn, 2014).
137. Moish Kutnoski, “The Ethical Dangers and Merits of Predictive Policing,” *Journal of Community Safety and Well-Being* 2, no. 1 (March 2017): 13–17.
138. Meredith Whittaker et al., “AI Now Report 2018,” (AI Now Institute at New York University, 2018).
139. Todd Feathers, “Flawed Algorithms Are Grading Millions of Students’ Essays,” *Vice News*, August 20, 2019.
140. Fantuzzo et al., “The Integrated Data System Approach: A Vehicle to More Effective and Efficient Data-Driven Solutions in Government.”
141. “Robot Teaching, Pedagogy and Policy,” in *The Oxford Handbook of Ethics of AI* (2019: Oxford University Press, n.d.).
142. Priscilla Regan, Jolene Jesse, and Elsa Khwaja, “Big Data in Education: Developing Policy for Ethical Implementation in the US and Canada,” (American Society for Public Administration Annual Conference, Seattle WA, 2016).
143. Jiayun Feng, “China to Curb Facial Recognition Technology in School,” *SupChina*, September 5, 2019.
144. Charlie Warzel, “How Much School Surveillance Is Too Much?,” *New York Times*, August 27, 2019, Sec. Opinion.