



"A lot of the times, the failings are not in AI. They're human failings and we're not willing to address the fact that there isn't a lot of diversity in the teams building the systems in the first place."

– Vivienne Ming

"You want people in AI who have compassion, who are thinking about social issues."

– Dr. Timnit Gebru

RESPONSIBLE AI AND TECH JUSTICE: A GUIDE FOR K-12 EDUCATION

A guide designed for K-12 educators and students to support the critical interrogation of artificial intelligence and its implications on individuals, communities, and the world.

Responsible AI and Tech Justice Guide © 2023 by Shana V. White and Allison Scott is licensed under Attribution-NoDerivatives 4.0 International. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nd/4.0/>

Responsible AI & Tech Justice: A Guide for K-12 Education

1.18.2024

CONTRIBUTING AUTHORS AND ADVISORS

KAPOR FOUNDATION

Shana V. White
Dr. Allison Scott
Dr. Sonia Koshy

SENIOR ADVISORY BOARD

Dr. Saifya Noble
Dr. Emily Bender
Dr. Brandeis Marshall
Dr. Chris Gillard
Dr. Ali Alkhatib
Dr. Kamau Bobb
Dr. Jane Margolis
Aerica Shimizu Banks
Adrienne Williams
Tawana Petty

ADVISORY COMMITTEE MEMBERS

Dr. Sepehr Vakil
Dr. Jean Ryoo
Dr. Aman Yadav
Dr. Marie Heath
Dr. Dan Krutka
Christy Crawford
Diane Levitt
Irene Lee
Rebecca Luebker
Sarah Dunton

ACKNOWLEDGEMENTS

The authors of this guide would like to extend sincere gratitude, thanks, and appreciation for the decades of work by leading scholars on artificial intelligence, bias in technology, racism in computing, education equity, social justice and technology, and technology ethics and justice. We appreciate the time and dedication of the Senior Advisory Board members and Advisory Committee Members who generously contributed their expertise to this guide. This project is sponsored by the Kapor Foundation, with generous support from co-chairs, Mitch Kapor and Dr. Freada Kapor Klein.

Background

“I believe AI is going to change the world more than anything in the history of humanity. More than electricity.”

– [Kai-Fu Lee, 2021](#)

The recent technological advances in artificial intelligence (AI) have captivated the attention of students, teachers, investors, entrepreneurs, tech leaders, activists, and policymakers alike, with the global artificial intelligence market projected to grow from [\\$515B in 2023 to \\$2,025B by 2030](#), over [45% of the US population](#) using generative AI tools, and at least [25 states, Puerto Rico and the District of Columbia](#) introducing legislation related to AI in 2023. Yet, despite the potential for AI to identify breakthroughs in [disease prevention and treatment](#); improve efficiency in communication through auto-complete, virtual assistants, and chatbots; reduce [business costs](#) and improve efficiency and [worker productivity](#); and enable greater personalization to [improve educational outcomes](#) for students with disabilities and multilingual learners, advancements in AI are not without concern.

“Among the great challenges posed to democracy today is the use of technology, data, and automated systems in ways that threaten the rights of the American public...”

– [OSTP AI Bill of Rights](#)

Algorithms and AI tools have contributed to [increased polarization](#), the [proliferation of mis/disinformation](#), increased online safety and privacy concerns, and [mental health challenges](#) among teens and young adults. The massive amount of [energy needed to power AI](#) technologies can impact carbon emissions at a time where climate change is a [leading global concern](#). Researchers have documented biases and [systemic inaccuracy in facial recognition](#) software that [misidentify Black individuals](#) at much higher rates, and have led to cases of [wrongful arrest](#). Algorithms used to determine health care and insurance decisions have [demonstrated significant bias](#) against Black patients; mortgage-approval application algorithms [disproportionately denied loans](#) to applicants of color; algorithms widely-used in the hiring process have demonstrated [bias against women](#) and [individuals with disabilities](#); and algorithms used to predict recidivism were much more likely to [assess Black defendants as high risk](#), impacting sentencing and parole decisions. These biases in algorithmic tools, their underlying datasets, and their utilization have profound implications on the lives of individuals from the most marginalized communities, thus replicating and exacerbating societal inequality. Further, the lack of diversity within the [AI workforce](#) and among AI company leadership, [boards](#), and investors impact the types of technology tools created, the ethical, moral, and social decisions made in the development and deployment process, and opportunities for economic success within one of the most lucrative industries. It is critical to understand that the AI technologies are themselves not neutral, and are a reflection of the [intentional and unintentional biases of their creators](#).

“If we’re not careful, AI will perpetuate bias in our world. Computers learn how to be racist, sexist, and prejudiced in a similar way that a child does. The computers learn from their creators — us.”

– Aylin Caliskan
Computer Scientist

At a time when there is so much promise juxtaposed with growing fear of the powerful breakthroughs in AI and its increasing impact on our everyday lives, the comprehensive understanding and critical interrogation of AI technologies must no longer be limited to a select few. We must seize upon this moment to improve equity and justice in AI by preparing a diverse and robust AI workforce; providing reskilling opportunities to effectively incorporate AI into existing jobs and roles; conducting algorithmic audits and identifying biases in algorithms; investing in solutions aiming to utilize AI for positive social impact; and implementing regulation and accountability measures to mitigate risk and harm. To achieve this, we must begin by making fundamental changes in K-12 education to ensure all students and teachers are equipped with the knowledge, skills, and resources to become critical consumers and ethical producers of the next generation of technologies.

“We need people designing technologies for society to have training and an education on the histories of marginalized people, at a minimum, and we need them working alongside people with rigorous training and preparation from the social sciences and humanities.”

– Dr. Safiya Umoja Noble,
*Algorithms of Oppression:
How Search Engines Reinforce Racism, 2018*



The Need for A Justice-Centered Approach to AI Education

“The global injustices experienced by the most marginalized within our societies are replicated and re-embedded into the most popular technologies. Unless we examine the ethics of these technologies, what data is collected and used, and who it is shared with, we will kick the can of racial injustice down the road into global societies that will only become “better,” and “smarter” at reproducing racialized bias, racism, caste, and ethnic discrimination”

– The [AIxRacial Justice Toolkit 2022](#)

As the AI landscape continues to rapidly evolve, tech companies, advocacy groups, legislators, and individual citizens continue to grapple with how to effectively minimize risk, mitigate harm, and harness this powerful technology for the good of society. Most recently, we have seen policy efforts to regulate this nascent space in the form of international legislation on AI from the [European Union](#) and a Presidential Executive Order on [safe, secure, and trustworthy AI](#) in the United States. In the entertainment industry, we have seen class-action lawsuits from [artists](#), [writers](#), and [musicians](#) over copyright infringement, while civil rights groups are [sounding the alarm](#) about AI's amplification of bias and discrimination. Meanwhile, there is an all-out arms race by [tech companies](#) and [investors](#) seeking to capitalize on this emerging technology.

In the K-12 education space, we have seen significant attention, discussion, and debate about how AI might [transform the future of teaching and learning](#), improve educational outcomes through differentiation and accessibility, and/or exacerbate challenges including bias, inaccuracy of information, and data privacy concerns. New [frameworks](#), guides, and [toolkits](#) have been developed to help educators, legislators, and education technology creators understand the opportunities and risks that are presented by the evolution of artificial intelligence and adopt policies and practices to enable responsible, ethical, and inclusive development and utilization of AI tools in education. Yet, a key element is glaringly missing from these conversations about the future of AI in education. While it is critical to delineate guidance for how teachers, students, and schools use AI tools in the advancement of education, it is equally important to prioritize how students and educators interrogate ethics, equity, and justice in the creation, deployment, and utilization of AI technologies as a core component of a robust K-12 education.

The examination of ethical and equitable concerns about AI technologies has historically been a component of computing education, incorporated into [teaching](#) and [learning standards](#) and the learning goals of [Advanced Placement CS](#) courses. More recently, these topics have been incorporated into [national guidelines](#) for K-12 AI education. However, the critical examination of the technologies that are ubiquitous in students' and educators' lives, impact social and democratic structures, and have disproportionate impacts on marginalized communities must not be a peripheral part of computing education. Instead, we argue that the critical interrogation of AI's development and impact must be a core component of K-12 computing education--and of education more broadly-- and we must intentionally center racial and social justice in the examination of these technologies.

Responsible AI & Tech Justice: A Guide for K-12 Education

Articulating a new vision for K-12 computer science education, the [Justice-Centered Computing Education Framework](#) outlines an approach to addressing structural educational inequality, critically interrogating the computing ecosystem, and intentionally developing the resources, programs, and policies to achieve full access and meaningful outcomes for students of all identities and abilities. Within this framework, there is an explicit focus on prioritizing the interrogation of ethics, equity, and justice in the creation, deployment, and utilization of AI technologies as a component of a robust K-12 education such that all students and educators are equipped with the knowledge, resources, and opportunities to critically interrogate AI technologies and to challenge and disrupt their harms.

As a key component of Justice-Centered Computing Education, this guide articulates a vision for **Responsible AI and Tech Justice in K-12 Education** as:

A robust and comprehensive course of study that utilizes an explicit racial and social justice lens to equip all students with the knowledge and resources to critically interrogate the ethical and equitable development, deployment, and impacts of AI, while simultaneously challenging, disrupting, and remedying the harms that these technologies can produce within individuals' lives, communities, and society at large.

To translate this vision into action, six core components serve as a guide for educators, parents, policymakers, and advocates seeking to design learning experiences for students across educational settings, where both the critical interrogation of technologies and the creation of more ethical and equitable solutions are prioritized. The core components cover large concepts and are complemented by sets of sample interrogation questions, which provide suggested ideas to cover, while allowing space for creativity, innovation, and evolution in content as the space continues to rapidly evolve. A table of sample resources, ranging from activities and lessons to articles and books aligned with each of the core components is outlined in Appendix 1. It is our goal that these six core components will be used by a range of K-12 educators across disciplines, including teachers, instructional coaches, administrators, and curriculum providers to directly inform their instruction, pedagogy, policies, and curricula resources. We aim to ensure that these concepts become deeply ingrained within computer science classrooms and pathways, across other content areas and disciplines, and within outside of school time programming.



The Six Core Components:

1. **Examine the AI technology creation ecosystem** from who designs and develops products and how they are developed, to who invests in their creation and benefits from their adoption.
2. **Interrogate the complex relationship between technology and human beings**, including human-computer interaction and topics of values, ethics, privacy, and safety.
3. **Explore the impacts and implications of AI technologies on society**, including positive benefits, negative consequences, and the perpetuation of exclusion, marginalization, and inequality.
4. **Interrogate personal usage of AI technologies** to become critical consumers of products and address misuse, exploitation, and safety concerns.
5. Build a **critical lens in the collection, usage, analysis, interpretation, and reporting of data**.
6. **Minimize, mitigate, and eliminate harm** and injustice caused by AI technologies through both the responsible and ethical creation process and individual and collective right to refusal.

Core components & Interrogation questions

Core Component

Sample Interrogation Questions

1. **Examine the AI technology creation ecosystem.**

- Who is involved in the ideation stage, research, and design phase of AI technology creation?
- Why did the individual/group produce this piece of technology or AI tool?
- What demographic trends exist among AI technology company boards, leadership, and technical workforce?
- What are the backgrounds, cultures, and values of AI company boards and leadership teams?
- How do the identities/backgrounds of technology producers impact algorithmic thinking?
- Who invests in AI technology tools and who benefits (financially) from their creation?
- What is Big Tech (Google, Microsoft, Amazon, etc.), who are the key players in AI (OpenAI), and what power does Big Tech hold in the past, current, and future of AI technologies?
- What are techno-optimists, what do they believe, and what do their critics believe?
- How are AI products currently regulated?
- Why is there so little regulation and accountability of AI technologies?
- How are algorithms used across social media platforms?
- How are AI technologies used for surveillance, policing, and international conflicts?

2. **Interrogate the complex relationship between technology and human beings.**

- What is HCI (human-computer interaction) & what are human-centered approaches to AI technology design and development?
- What are the similarities and differences between autonomous intelligent systems and humans?
- How do algorithms influence human behavior?
- Who trains algorithms, how are they trained, and what are they trained to optimize for?
- What does consent look like between humans and technology?
- Where are there ethical dilemmas in technology creation and design and how can ethics be prioritized prior to deployment?
- Can technologies be dehumanizing?
- How can AI technologies' advancement impact society?
- Can technology and AI solve human rights issues in society and be leveraged for good?
- How are AI technologies both potentially improving and challenging to human lives?
- What are the potential positive and negative consequences of using AI technologies for personal convenience?
- What are the potential impacts, benefits, and risks of technology being used to communicate with other human beings?
- How do AI-enabled tools capture personal data?
- How do individual biases get enhanced or replicated within AI systems?

Core Component

3. *Explore the impacts and implications of AI technologies on society.*

4. *Interrogate personal usage of AI technologies to become critical consumers.*

Sample Interrogation Questions

- How does the rise of AI and other emerging technologies contribute to furthering the 'digital divide' for marginalized communities?
 - What is algorithmic bias, its causes, and its implications?
 - What is facial recognition, how is it used, and who utilizes this technology?
 - How do institutional and structural injustices get enhanced or replicated by technology and AI tools?
 - How does social media affect a person's social, emotional, and mental health?
 - How do AI technologies impact healthcare decisions and treatments?
 - How do AI technologies impact politics, elections, and governments in the US and the world?
 - How are algorithms driving polarization and division online and offline?
 - How do AI technologies contribute to climate change?
 - How do AI technologies impact economic opportunities and outcomes as well as replicate inequality?
 - What communities have disproportionately been impacted by surveillance?
 - How are AI tools used in military conflict?
 - How can AI technologies improve education and healthcare outcomes?
-
- What is your digital footprint?
 - What is digital literacy, AI literacy, and algorithmic literacy?
 - What personal data can certain AI tools, technologies, and products access when used?
 - What data are you sharing and with whom?
 - What rights do you have to your own data?
 - How do AI tools and technologies keep a person's private information secure?
 - What are the terms of service for tools and products and how can consumers be more informed about their contents?
 - How do algorithms impact the content in your social media feeds?
 - How does social media content vary based on identity markers, demographics, and other preferences?
 - How can I harness AI technologies to find quality, factual information and what are the possible limiting factors of these AI tools?
 - What are some threats to personal safety posed by AI tools or technologies?
 - What is misinformation, disinformation, and propaganda in the digital information ecosystem?
 - What rights do people have to protect themselves from surveillance and facial recognition?
 - What protections are in place for individuals to refuse certain technology and AI tools?
 - What tech regulations and protections are available and potentially needed for consumers of AI technologies?

Core Component

5. ***Build a critical lens in the collection, usage, analysis, interpretation, and reporting of data.***

6. ***Minimize, eliminate, and mitigate harms and injustices caused by AI technologies.***

Sample Interrogation Questions

- What data sources are used for AI technologies and can these sets have biases?
 - What is the difference between training data, validation data, and testing data in datasets used for AI technologies?
 - How is the data for this AI technology obtained?
 - What is data scraping or web scraping and can it be responsibly done?
 - What are ethical challenges with data used in LLMs?
 - How can biases in datasets cause harm to individuals or groups?
 - How can data be manipulated to cause harm to individuals or groups?
 - What are responsible and ethical approaches to data collection and usage?
 - What is data capitalism?
 - What is data sovereignty?
 - What is data privacy?
 - What data protections do individuals have?
 - What are models of data ownership?
 - What are community approaches to collecting and owning data?
-
- What is digital justice?
 - What are strategies for designing for justice?
 - What are approaches to engaging diverse individuals, communities, and perspectives in the AI design process?
 - What are some guiding principles for ethical AI design?
 - What are the principles of responsible, trustworthy, and socially responsible AI technologies?
 - What are algorithmic audits and how can they be used to reduce potential harms from AI technologies?
 - How can individuals and communities refuse/opt out of using harmful technologies?
 - How can individuals and communities address the harms of AI technologies?
 - What are ways to protect marginalized communities from job displacement due to AI?
 - Can AI technology be responsibly used to undo environmental damage and address climate change?
 - Can AI technology be responsibly used to potentially undo this environmental damage?
 - What tech regulations are in place (or should be) to protect citizens and reduce harm?
 - What is transparency and why is it required for ethical technology creation?

Future Directions

The **Responsible AI and Tech Justice Guide** for K-12 Education is intended to articulate a new approach to teaching and learning in the era of rapid AI development and deployment, which prioritizes the interrogation of ethics, equity, and justice in the creation, deployment, and utilization of AI technologies and inspires the design and adoption of more equitable and just products and solutions. While this guide provides a conceptual overview of these topics, much more is needed to support the translation and adoption of these concepts. Along with aligned partners, we will support efforts to develop curricular resources, professional learning opportunities, conduct research on interventions, and implement policy change.

This guide represents a first step towards incorporating principles of ethics, equity, and justice in the critical analysis of artificial intelligence technologies within K-12 education and builds upon decades of critical scholarship and activism across disciplines ranging from computer science and machine learning, to history, political science, and sociology. As technological breakthroughs continue, as new solutions are uncovered, as new approaches are developed, and as new harms are revealed, we intend for this guide, its core components, interrogation questions, and resources, to also evolve. We invite feedback and collaboration from partners across disciplines to refine the guide, its core components, and its resources.



Appendix 1

Responsible AI and Tech Justice: Resources

1. *Examine the AI technology creation ecosystem.*

Articles & Case Studies

- [Why AI's Diversity Crisis Matters--And How to Tackle It](#) [Nature]
- [AI Investment Forecast to Approach \\$200B Globally by 2025](#) [Goldman Sachs]
- [AI Will Save the World](#) [A16Z]
- [AI Poses 'Risk of Extinction,' Industry Leaders Warn](#) [NY Times]
- [Why Timnit Gebru wants AI giants to think small](#) [Fast Company]
- [Big Tech wants AI Regulation. The rest of Silicon Valley is skeptical](#) [Washington Post]
- [What does the E.U.'s sweeping rules for Big Tech mean for your life online](#) [Washington Post]
- [Tech Elite's AI Ideologies Have Racist Foundations, Say AI Ethicists](#) [People of Color in Tech]
- [These Women Tried to Warn Us About AI](#) [Rolling Stone]
- [These are the people who could actually pause AI if they wanted to](#) [The Guardian]
- [A Battlefield AI Company Says It's One of the Good Guys](#) [WIRED]
- [The Microsoft Police State: Mass Surveillance, Facial Recognition, and the Azure Cloud](#) [The Intercept]
- [The UN hired an AI Company to Untangle the Israel-Palestinian Crisis](#) [WIRED]
- [AI companies have all kinds of arguments against paying for copyrighted content](#) [The Verge]

Lessons

- [Whose interests do "teaching machines" serve?](#)
- [Intro to Ethical Matrices](#)

Additional Resources

- Coded Bias Documentary, directed by Shalini Kantayya
- [Blood in the Machine: The Origins of the Rebellion Against Big Tech](#) by Brian Merchant
- [Algorithms of Oppression: How Search Engines Reinforce Racism](#) by Safiya Noble
- [The Atlas of AI: Power, Politics, and Planetary Cost of Artificial Intelligence](#) by Kate Crawford

2. *Interrogate the complex relationship between technology and human beings.*

Articles & Case Studies

- [Seven HCI Grand Challenges](#) [Stephanidis et al., 2019]
- [My North Star for the Future of AI](#) [The Atlantic]
- [Fei-Fei Li Started an AI Revolution By Seeing Like an Algorithm](#) [WIRED]
- [How does a computer discriminate?](#) [NPR]
- [Humans are Biased. Generative AI is Even Worse](#) [Bloomberg]
- [Generative AI's Impact on Jobs and Workflows](#) [McKinsey]
- [Is Biden doing enough to protect workers from AI?](#) [Fast Company]
- [Talking about a 'schism' is ahistorical](#) [Bender, 2023]
- [Cleaning Up ChatGPT Takes Heavy Toll on Human Workers](#) [Wall Street Journal]
- [AI Is a Lot of Work](#) [The Verge]
- [This New Autonomous Drone for Cops Can Track You in the Dark](#) [WIRED]
- [Can AI-Driven Voice Analysis Help Identify Mental Disorders?](#) [NY Times]
- [Anthony Levandowski Reboots Church of Artificial Intelligence](#) [Bloomberg]
- [If you have a face, you have a place in the conversation about AI](#) [NPR]
- [OpenAI 'was working on an advanced model so powerful it alarmed staff'](#) [The Guardian]
- [Dopamine, Smartphones, and You: A Battle for Your Time](#) [Harvard GSAS]

(cont'd)

Lessons

- [Who is responsible for discriminatory design?](#)
- [Ethical Reflection Modules for CS](#)

Additional Resources

- [Unmasking AI: My Mission to Protect What is Human in a World of Machines](#) by Joy Buolamwini
- [Artificial UnIntelligence: How Computers Misunderstand the World](#) by Meredith Broussard
- [Black Software](#) by Charlton D. McIlwain
- [Captivating Technology: Race, Carceral Technoscience, and Liberatory Imagination in Every Life](#) by Ruha Benjamin

3. Explore the impacts and implications of AI technologies

Articles & Case Studies

- [FCC adopts rules to eliminate 'digital discrimination' for communities with poor internet access.](#) [AP News]
- [Social Media is Driving Teen Mental Health Crisis, Surgeon General Warns](#) [NBC News]
- [Critics Furious Microsoft is Training AI by Sucking up Water during Drought](#) [Futurism]
- [AI's Climate Impact Goes Beyond its Emissions](#) [Scientific American]
- [America Already Has an AI Underclass](#) [The Atlantic]
- [The Secret Bias Hidden in Mortgage Approval Algorithms](#) [The Markup]
- [Eight Months Pregnant and Arrested After False Facial Recognition Match](#) [NY Times]
- [We don't know how Israel's military is using AI in Gaza, but we should](#) [LA Times]
- [Tech giants are trying to stop deepfakes from influencing the 2024 election. Good luck](#) [Fast Company]
- [Tech Disrupted Hollywood. AI Almost Destroyed It](#) [WIRED]
- [How AI reduces the world to stereotypes](#) [Rest of World]
- [The AI-Powered, Totally Autonomous Future of War Is Here](#) [WIRED]
- [UnitedHealth pushed employees to follow an algorithm to cut off Medicare patients' rehab care](#) [Stat News]
- [A Black woman invented Home Security. Why did it go so wrong?](#) [WIRED]
- [How Robots Can Assist Students with Disabilities](#) [NY Times]
- [Meet Ashley, the world's first AI-powered political campaign caller](#) [Reuters]
- [How the Federal Government Can Rein in AI in Law Enforcement](#) [NY Times]

Lessons

- [Introduction to Algorithmic Bias](#)
- [Is this a Weapon of Math Destruction?](#)
- [Technology Reset Simulation](#)
- [Drinking Water is Everywhere, Right?](#)
- [Apps for Impact](#)
- [Our Living Planet](#)

Additional Resources

- [Race and Digital Media: An Introduction](#) by Lori Kido Lopez
- [Distributed Blackness: African American Cybercultures](#) by André Brock
- [The Black Image in the White Mind: Media and Race in America](#) by Robert M. Entman and Andrew Rojecki
- [Technology and the Logic of American Racism: A Cultural History of The Body As Evidence](#) by Sarah E. Chinn

Additional Resources (cont'd)

- [Hidden in White Sight: How AI Empowers and Deepens Systemic Racism](#) by Calvin D. Lawrence
- [Racism and Racial Surveillance](#) edited by Sheila Khan, Nazir Ahmed Can, and Helena Machado
- [Seen & Unseen: Technology, Social Media, and the Fight for Racial Justice](#) by Marc Lamont Hill and Todd Brewster
- [AI and the Future of Education: Teaching in the Age of Artificial Intelligence](#) by Priten Shah
- [When the Hood Comes Off: Racism and Resistance in the Digital Age](#) by Rob Eschmann
- [Weapons of Math Destruction](#) by Cathy O’Neil

4. **Interrogate personal usage of AI technologies to become critical consumers.**

Articles & Case Studies

- [TikTok loves GenZ’s true confessions. Colleges and employers, not so much.](#) [Washington Post]
- [Children are targets for ID theft. Here’s what parents need to know](#) [Washington Post]
- [First-Gen Social Media Users Have Nowhere to Go](#) [WIRED]
- [How to avoid falling for misinformation, AI images on social media](#) [Washington Post]
- [TSA now wants to scan your face at security. Here are your rights.](#) [Washington Post]
- [Adobe is selling fake AI images of the war in Israel-Gaza](#) [Crikey]
- [Trying to Debug AI’s Biases](#) [NY Times]
- [Why you shouldn’t tell ChatGPT your secrets](#) [Washington Post]
- [How to skim a privacy policy to spot red flags](#) [Washington Post]
- [This AI makes you look like a masterpiece—while teaching you about its own bias](#) [Vox]
- [Technologies of Control: We Have to Defend Our Right to Refusal](#) [LSE]
- [Apple and Google team up to fight AirTag stalking](#) [Washington Post]

Lessons

- [Mapping the Media Education Terrain](#)
- [Are “smart” technologies worth the cost?](#)
- [Propaganda Then and Now: What has changed?](#)
- [Does new media provide more “real news”?](#)
- [Introduction to Algorithms as Opinions](#)

Additional Resources

- [Viral Justice: How We Grow The World We Want](#) by Ruha Benjamin
- [None But Ourselves Can Free Our Minds: Critical Computational Literacy as a Pedagogy of Resistance](#) by Clifford Lee
- [PowerOn! Graphic Novel](#) by Jane Margolis and Jean Ryoo
- [AI Literacy, Explained](#) by Alyson Klein
- [Affordances and challenges of artificial intelligence in K12 education: a systematic review](#) by Helen Crompton
- [Digital Defense Playbook](#), by Our Data Bodies

5. **Build a critical lens in the collection, usage, analysis, interpretation, & reporting of data.**

Articles & Case Studies

- [Indigenous Data Sovereignty Governance](#) [University of Arizona]
- [These companies will pay for your data. Is it a good deal?](#) [Washington Post]
- [Police love Google’s surveillance data. Here’s how to protect yourself.](#) [Washington Post]
- [Google promised to delete sensitive data. It logged my abortion clinic visit.](#) [Washington Post]

Articles & Case Studies (cont'd)

- [Data Work at its Layers of \(In\)Visibility](#) [Social Science Research Council]
- [Data Privacy and Ownership to Remain Key Concerns in Web Scraping Industry Next Year](#) [Forbes]
- [Thirsty Data Centers Are Making Hot Summers Even Scarier](#) [Bloomberg]
- [Health Care Data Is a Researcher's Gold Mine](#) [WIRED]
- [23andMe User Data Stolen in Target Attack on Ashkenazi Jews](#) [WIRED]
- [Generative AI Is Making Companies Even More Thirsty for Your Data](#) [WIRED]
- [Harnessing powerful AI while mitigating risks: It's about data!](#) [Fordham Institute]

Lessons

- [Datasets and Ethics Exercise](#)
- [Data Privacy Drag and Drop Exercise](#)
- [Introduction to Classification and Supervised Machine Learning](#)

Additional Resources

- [Decolonizing Data Toolkit](#)
- [Data Portraits: Visualizing Black America](#) by W.E.B DuBois
- [Discriminating Data: Correlations, Neighborhoods, and the New Politics of Recognition](#) by Wendy Hui Kyong Chun
- [The Black Technical Object: One Machine Learning and the Aspiration of Black Being](#) by Ramon Amaro
- [Artificial Whiteness](#) by Yarden Katz
- [Redlining Culture: A Data History of racial inequality and postwar fiction](#) by Richard Jean So
- [Data and Social Good: Using Data Science to Improve Lives, Fight Injustice, and Support Democracy](#) by Mike Barlow
- [Ida B. Wells Just Data Lab](#)

6. **Minimize, mitigate, and eliminate harm and injustice caused by AI technologies**

Articles & Case Studies

- [Designing AI with Justice](#) [Public Books]
- [The First Amendment Should Protect Us for Facial Recognition Technologies--Not the Other Way Around](#) [Tech Policy Press]
- [How facial recognition software in criminal investigations can harm communities of color](#) [Detroit Today]
- [AI Hurts Consumers and Workers – and Isn't Intelligent](#) [Tech Policy Press]
- [She Built an App to Block Harassment on Twitter. Elon Musk Killed it](#) [Time]
- [Everyone Is a Luddite Now](#) [WIRED]
- [In Class, Some Colleges Overlook Technology's Dark Side](#) [Inside Higher Ed]
- [Seeking Algorithmic Justice In Policing AI](#) [Science Friday]
- [Wrongful Arrest Exposes Racial Biases in Facial Recognition Technology](#) [CBS News]
- [Sam Altman's Ouster and the Age-Old Tension Between Doing Good and Making Money](#) [Time]
- [Europe agrees to landmark AI regulation deal](#) [Reuters]
- [Does AI Lead Police to Ignore Contradictory Evidence?](#) [The New Yorker]
- [Clear wants to scan your face at airports. Privacy experts are worried](#) [Washington Post]
- [NY Times Sues Microsoft and OpenAI, Alleging Copyright Infringement](#) [Wall Street Journal]
- [How AI Will Turbocharge Misinformation--And What We Can Do About It](#) [Axios]

(cont'd)

Lessons

- [How Can We Fight Algorithmic Bias?](#)
- [Why Does Ethics in CS matter?](#)
- [Bringing Climate Justice Home](#)
- [Should We Be More Like the Luddites](#)
- [Speculative Fiction Activity: Alter Ego](#)

Additional Resources

- [Race After Technology](#) by Ruha Benjamin
- [Design Justice](#) by Sasha Constanza-Chock
- [Statement from listed authors of Stochastic Parrots on the “AI pause” letter](#) by Timnit Gebru, Emily M. Bender, Angelina McMillian-Major and Margaret Mitchell
- [Your Computer Is on Fire](#) by Thomas S. Mullaney, Benjamin Peters, and Mar Hicks
- [The Atlas of AI: Power, Politics, and Planetary Cost of Artificial Intelligence](#) by Kate Crawford
- [Because Technology Discriminates](#) by Logan D.A. Williams
- [Cyber Racism: White Supremacy Online and the New Attack on Civil Rights](#) by Jessie Daniels
- [Change from the Outside: Towards Credible Third Party Audits of AI Systems](#) by Inioluwa Deborah Raji, Sasha Costanza-Chock, and Joy Buolamwini
- [Detroit Digital Justice Coalition](#)
- [Data Capitalism and Algorithmic Racism](#) by Milner and Traub



Appendix 2

Key Terms

Accountability

In AI, accountability relates to the expectation that designers, developers, and deployers will comply with standards and legislation to ensure the proper functioning of AIs during their lifecycle ([Fjeld et al. 2020](#)).

Algorithm

A procedure used for solving a problem or performing a computation. Algorithms act as an exact list of instructions that conduct specified actions step by step in either hardware or software-based routines ([Tech Target](#)).

Algorithmic Audit

A method of repeatedly querying an algorithm and observing its output in order to draw conclusions about the algorithm's opaque inner workings and possible external impact ([Metaxa, et al, 2021](#)).

Algorithm Bias

Algorithm bias describes systematic and repeatable errors in a computer system that create unfair outcomes, such as privileging one arbitrary group of users over others. It also occurs when an algorithm produces results that are systemically prejudiced due to erroneous assumptions in the machine learning process ([Florida State Libraries](#)).

Algorithmic Discrimination

Occurs when automated systems contribute to unjustified/different treatment or impacts disfavoring people based on their race, color, ethnicity, sex (including pregnancy, childbirth, and related medical conditions, gender identity, intersex status, and sexual orientation), religion, age, national origin, disability, veteran status, genetic information, or any other classification protected by law. Depending on the specific circumstances, such algorithmic discrimination may violate legal protections ([White House OTSP AI Bill of Rights](#)).

Artificial Intelligence

While there are many definitions of 'artificial intelligence,' this guide uses the definition from the White House Executive Order on Safe, Secure, and Trustworthy Development and Use of AI: the term 'artificial intelligence' means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments ([White House, EOSS-TAI](#)). A broader discussion on AI can be found in the National Institute of Standards and Technology Special Publication 1270, [Towards a Standard for Identifying and Managing Bias in Artificial Intelligence](#).

Automated Decision System

A computational process, including one derived from machine learning, statistics, or other data processing or artificial intelligence techniques, that makes a decision or facilitates human decision making, that impacts consumers ([Algorithmic Accountability Act](#)).

Consent

To agree to do or allow something; to give permission for something to happen or be done ([Merriam-Webster](#)).

Critical Theory

Critical theory is broadly defined as any approach to humanities and social philosophy that focuses on society and culture to attempt to reveal, critique, and challenge power structures ([Brittanica](#)).

Data

In computing, data is information that has been translated into a form that is efficient for movement or processing. Relative to today's computers and transmission media, data is information converted into binary digital form. It is acceptable for data to be used as a singular subject or a plural subject. Raw data is a term used to describe data in its most basic digital format ([Tech Target](#)).

Data Capitalism

An economic model built on the extraction and commodification of data and the use of big data and algorithms as tools to concentrate and consolidate power in ways that dramatically increase inequality along lines of race, class, gender, and disability ([Milner & Traub, 2021](#)).

Data Discrimination/Bias

Bias that occurs when predefined data types or data sources are intentionally or unintentionally treated differently than others ([Florida State Libraries](#)).

Deep Learning

Deep learning is a type of machine learning that can process a wider range of data resources (images, for instance), requires less human intervention, and can often produce more accurate results than traditional machine learning. Deep learning uses neural networks to ingest data and process it through multiple iterations that learn increasingly complex features of the data ([McKinsey](#)).

Design Justice

Design justice goes beyond fairness. It entails thinking about the matrix of domination—about intersecting systems of oppression—and what it means to design sociotechnical systems that can transform or overturn these systems, rather than constantly reproducing them in technology, in design, and in machine learning ([Costanza-Chock](#)).

Digital Literacy

The ability to use information and communication technologies to find, evaluate, create, and communicate information, requiring both cognitive and technical skills ([ALA](#)).

Ethics

Ethics are an external social system relating to a specific group that defines right and wrong behaviors. Ethics are codified into a set of rules or a system and adopted by people in that field. Ethical principles remain consistent across industries and institutions, as they offer strict behavior guidelines ([Center for Public Trust](#)).

Generative Artificial Intelligence

Refers to deep-learning models that can generate high-quality text, images, and other content based on the data they were trained on; generative models can take raw data and “learn” to generate statistically probable outputs when prompted ([IBM](#)).

Harm

To damage or injure physically or mentally; to cause physical or mental damage ([Merriam-Webster](#)).

Human-computer interaction (HCI)

A multidisciplinary field of study focusing on the design of computer technology and, in particular, the interaction between humans (the users) and computers ([Interaction Design Foundation](#)).

Humanize

To represent (something) as human: to attribute human qualities to (something); to address or portray (someone) in a way that emphasizes that person's humanity or individuality ([Merriam-Webster](#)).

Identity

The collective set of characteristics both visible and invisible which a person is known by or can be group by; identity characteristics include race, gender, socio-economic status, ethnicity, languages spoken as well as others ([Edutopia](#)).

Information Ecosystem

Complex, adaptive systems that include information infrastructure, tools, media, producers, consumers, curators, and sharers. Information ecosystems are complex organizations through which information flows ([Unhcr.org Information Mapping](#)).

LLM

Large language models (LLMs) are deep learning algorithms that can recognize, summarize, translate, predict, and generate content using very large datasets ([Nvidia](#)).

Machine Learning

Machine learning is a form of artificial intelligence based on algorithms that are trained on data. These algorithms can detect patterns and learn how to make predictions and recommendations by processing data and experiences, rather than by receiving explicit programming instruction ([McKinsey](#)).

Marginalized

A person or group that is treated insignificantly, pushed to the margins of society, and rendered powerless ([CultureAlly](#)).

Racial Justice

The systematic fair treatment of people of all races, resulting in equitable opportunities and outcomes for all. Racial justice — or racial equity — goes beyond “anti-racism.” It is not just the absence of discrimination and inequities, but also the presence of deliberate systems and supports to achieve and sustain racial equity through proactive and preventative measures ([NEA](#)).

Social Responsibility of AI

Refers to a human value-driven process where values such as Fairness, Transparency, Accountability, Reliability and Safety, Privacy and Security, and Inclusiveness are the principles; designing Socially Responsible AI Algorithms is the means; and addressing the social expectations of generating shared value – enhancing both AI’s ability and benefits to society – is the main objective ([Cheng et al., 2021](#)).

Tech Justice

“The right of people to decide, choose, and use technologies that assist them in leading the kind of life they value without compromising the ability of others and future generations to do the same”. This definition emphasizes technological inequalities within and between societies as well as the effects that our use of technology has on future generations ([Simon Trace](#)).

Trustworthy AI

Characteristics of trustworthy AI systems include: valid and reliable; safe, secure and resilient; accountable and transparent; explainable and interpretable; privacy-enhanced; and fair with harmful bias managed ([NIST, AI Risk Management Framework, 2023](#)).

Appendix 3

Existing Frameworks & Guidance for AI and AI in Education

K-12 Education

Resource

▶ [**AI Guidance for Schools Toolkit \(Teach AI\)**](#)

Summary

The AI Guidance for Schools Toolkit is designed to help local, state, and national education systems worldwide develop guidance on the responsible use of AI, ensure compliance with relevant policies, and build the capacity of all stakeholders to understand AI and use AI effectively. Contains 6 principles for AI in education:

1. Knowledge: Promote AI literacy
2. Integrity: Advance academic integrity
3. Evaluation: Regularly assess the impacts of AI
4. Compliance: Reaffirm adherence to existing policies
5. Balance: Realize the benefits of AI and address the risks
6. Agency: Maintain human decision-making when using AI

▶ [**K-12 AI Guidelines \(AI4K12\)**](#)

AI4K12 is developing national guidelines for AI education for K-12, as well as instructional resources and a community of practitioners. The guidelines contain 5 big ideas:

1. Perception
2. Representation & Reasoning
3. Learning
4. Natural Interaction
5. Societal Impact

▶ [**AI in Education Toolkit for Racial Equity \(The EdTech Equity Project\)**](#)

The AI in education toolkit for racial equity aims to mitigate bias in the design and development of edtech products and tools. The toolkit contains 8 phases of design within which critical questions related to racial equity are explored:

1. [Ideation](#)
2. [Logical Assumptions](#)
3. [Content](#)
4. [Datasets](#)
5. [Training Algorithms](#)
6. [UX Design](#)
7. [Testing & Evaluation](#)
8. [Implementation](#)

Resource

- ▶ [**Artificial Intelligence and the Future of Teaching and Learning \(U.S. Department of Education, Office of EdTech\)**](#)

Summary

The Office of Ed Tech outlined a set of recommendations to engage teachers, educational leaders, policymakers, researchers, and educational technology innovators, and providers as they work toward policy action related to Artificial Intelligence (AI) in education.

1. Emphasize Humans-in-the-Loop
2. Align AI Models to a Shared Vision for Education
3. Design AI Using Modern Learning Principles
4. Prioritize Strengthening Trust
5. Inform and Involve Educators
6. Focus R&D on Addressing Context and Enhancing Trust and Safety
7. Develop Education-specific Guidelines and Guardrails

- ▶ [**The Ethical Framework for AI in Education \(The Institute for Ethical AI in Education\)**](#)

The Framework is grounded in a shared vision of ethical AI in education and will help to enable all learners to benefit optimally from AI in education, whilst also being protected against the risks this technology presents. These principles include impacting:

1. Achieving Educational Goals
2. Forms of assessment
3. Administration & Workload
4. Equity
5. Autonomy
6. Privacy
7. Transparency & Accountability
8. Informed Participation
9. Ethical Design

- ▶ [**K-12 CS Content Standards \(CSTA\)**](#)

The CSTA K–12 Computer Science Standards delineate a core set of learning objectives designed to provide the foundation for a complete computer science curriculum and its implementation at the K–12 level. The five core concepts of the standards are:

1. Computing Systems
2. Networks and the Internet
3. Data and Analysis
4. Algorithms and Programming
5. Impacts of Computing

- ▶ [**K-12 CS Teacher Standards \(CSTA\)**](#)

The Standards for CS Teachers establish robust standards to help teachers deliver effective and equitable CS instruction in support of rigorous CS education for all K–12 students. There are five standards:

1. CS knowledge and skills
2. Equity and Inclusion
3. Professional Growth and Identity
4. Instructional Design
5. Classroom Practice

Resource

- ▶ [AP CS Principles \(College Board\)](#)

Summary

AP Computer Science Principles introduces students to the breadth of the field of computer science, learning to design and evaluate solutions, solve problems through the development of algorithms and programs, explain how computing innovations and computing systems, including the Internet, work, explore their potential impacts, and contribute to a computing culture that is collaborative and ethical. AP CS P contains 5 Big Ideas:

1. Creative Development
2. Data
3. Algorithms & Programming
4. Computing Systems & Networks
5. Impacts of Computing

- ▶ [AP CS A \(College Board\)](#)

AP Computer Science A introduces students to computer science through programming, with topics including the design of solutions to problems, the use of data structures to organize large sets of data, the development and implementation of algorithms to process data and discover new information, the analysis of potential solutions, and the ethical and social implications of computing systems. AP CS A contains 4 Big Ideas:

1. Modularity
2. Variables
3. Control
4. Impact of Computing

AI & Tech Policy

Resource

- ▶ [Executive Order on the Safe, Secure, Trustworthy Development and Use of Artificial Intelligence \(White House\)](#)

Summary

The Executive Order establishes new standards for AI safety and security, protects Americans' privacy, advances equity and civil rights, stands up for consumers and workers, promotes innovation and competition, advances American leadership around the world. The EO directs the following actions:

1. New Standards for AI Safety and Security
2. Protecting Americans' Privacy
3. Advancing Equity and Civil Rights
4. Standing up for Consumers, Patients, and Students
5. Supporting Workers
6. Promoting Innovation and Competition
7. Advancing American Leadership Abroad
8. Ensuring Responsible and Effective Government Use of AI

Resource

▶ [**AI Bill of Rights \(White House Office of Science and Technology Policy\)**](#)

Summary

The AI Bill of Rights outlines five principles that should guide the design, use, and deployment of automated systems that are aligned with democratic values and protect civil rights, civil liberties, and privacy. This framework applies to (1) automated systems that (2) have the potential to meaningfully impact the American public's rights, opportunities, or access to critical resources or services.

1. Safe and Effective Systems
2. Algorithmic Discrimination Protections
3. Data Privacy
4. Notice and Explanation
5. Human Alternative, Consideration and Fallback

▶ [**Advancing Racial Equity through Tech Policy \(AI Now\)**](#)

This resource argues for centering racial equity in the technology policy debate and outlines a five-point policy agenda for doing so:

1. Center racial equity in tech regulatory efforts
2. Promote democratic governance of technology
3. Build an equitable tech labor market
4. Ensure equitable access to goods, services, and information
5. Eliminate disparities in tech ownership and entrepreneurship

▶ [**Confronting Tech Power \(AI Now\)**](#)

This report aims to provide strategic guidance on the many levers we can use to shape the future trajectory of AI - and the tech industry behind it - to ensure that it is the public, not industry, that this technology serves – in the following areas:

1. Contain tech firm's data advantage
2. Build support for competition reforms as a key lever to reduce concentration in tech
3. Regulate ChatGPT and other large-scale models
4. Displace audits as the primary policy response to harmful AI
5. Future proof against the quiet expansion of biometric surveillance into new domains
6. Enact strong curbs on worker surveillance
7. Prevent "international preemption" by digital trade agreements which can be used to weaken national regulation on algorithmic accountability and competition policy

▶ [**RAI Toolkit \(Responsible AI-DoD\)**](#)

The Responsible Artificial Intelligence (RAI) Toolkit provides a centralized process that identifies, tracks, and improves the alignment of AI projects to RAI best practices and the DoD AI Ethical Principles while capitalizing on opportunities for innovation. The RAI Toolkit provides an intuitive flow guiding the user through tailorable and modular assessments, tools, and artifacts throughout the AI product lifecycle. The process enables traceability and assurance of responsible AI practice, development, and use.