four principles to make

ADVANCED DATA ANALYTICS

work for children and families

THE ANNIE E. CASEY FOUNDATION

The Annie E. Casey Foundation is a private philanthropy that creates a brighter future for the nation's children by developing solutions to strengthen families, build paths to economic opportunity and transform struggling communities into safer and healthier places to live, work and grow. For more information, visit www.aecf.org.

introduction

ADVANCES IN DATA SCIENCE and computing power are rapidly shifting the opportunities available to citizens, changing how systems interact with almost every aspect of our lives: as students, job seekers, patients, citizens and consumers.

This document presents the Annie E. Casey Foundation's point of view on the use of advanced analytics in social programs and policy, based on an exploration of the promise and consequences of advanced analytics for youth, families and communities. It also outlines a set of principles that distinguish between the useful, acceptable and harmful applications of new tools. Developed through broad consultation with data scientists, civil rights groups, public leaders and family advocates convened by the Foundation, as well as extensive consultation with Foundation staff and leadership, these principles represent points of agreement between those who are optimistic about the good these tools can do for families and those who remain extremely concerned about their abuse:

- *Expand opportunity.* Advanced analytics should open doors to greater opportunity for children and families, not merely incrementally improving the status quo.
- *Provide transparency and evidence.* Agencies and businesses providing services fundamental to the welfare of families and communities owe the public evidence about the use and effects of their decision-making analytics.
- *Empower communities.* These data tools should empower communities to hold systems accountable.
- *Promote equitable outcomes.* New tools and models should explicitly promote, and be judged against their ability to achieve, more equitable outcomes for historically disadvantaged groups.

Taken together, these principles direct a spotlight toward areas that need a great deal more attention, investment, innovation and consultation with families and affected communities. At a time when the COVID-19 pandemic has laid bare the inequities in many public systems and prompted reimagination of more effective approaches, organizations and advocates who care about children and families must not only seize the opportunity to prevent emerging technology from causing harm but also work affirmatively to support and promote the constructive use of these tools for equity and progress.

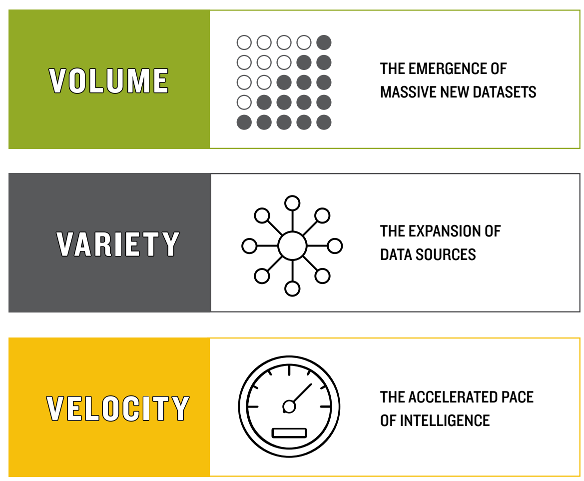
THE ESCALATING ROLE

of advanced analytics and big data

Advanced data analytics are deeply embedded in how public and private systems make decisions, already shaping the opportunities available to youth and families. Algorithms guide decisions on everything from which neighborhoods police patrol to whether loan officers extend credit and which job applicants a hiring manager selects to interview. (See Table 1.) With the volume, variety and velocity of information available on all of us (also known as “Big Data”) growing dramatically in the past decade, these analytical tools will play an increasingly central role in translating that information into decisions about families and communities in the coming decade.

BIG DATA is an umbrella term referring to the dramatic change in the *volume, variety* and *velocity* of information available to — and about — all of us.

- **VOLUME:** The digitization of all aspects of our lives and environments has created massive new datasets to mine for insight.
- **VARIETY:** In addition to administrative data intentionally collected by programs, companies can scrape social media sites and handwritten case notes to create unanticipated composite records.
- **VELOCITY:** The speed of collecting and analyzing information is constantly increasing, as processors, bandwidth and data-science techniques improve.



The new tools computer programmers and statisticians have developed to interrogate these large and varied datasets go by many names — such as machine learning and artificial intelligence — and they introduce both additional opportunities and challenges compared to more familiar statistical modeling methods, as discussed below. But whatever technique is used, the core questions raised by the use of advanced analytics for prediction, risk assessment and decision making are similar.

Table 1

APPLICATIONS OF ADVANCED ANALYTICS

- Affordable housing protection
- Criminal record expungement
- Early warning systems in K–12
- Evidence-based sentencing
- Insurance pricing
- Job awareness and employment decisions
- Learning analytics in higher education
- Policing
- Predicting risk of injury to youth in foster care
- Pretrial risk assessment
- School assignment and transportation
- Screening calls to child protective services
- Tenant screening

Many uses of advanced analytics are controversial. Civil rights advocates worry that more algorithmic and automated decision-making systems will worsen existing inequities across lines of race, age, income and gender. Organizers, academics and journalists have increasingly documented these problems in many advanced analytical tools in common use. (See the "Evidence-Based Sentencing" example in Appendix A.) Some have gone as far as to call for abolishing these tools, arguing that agencies' application of advanced analytics to the increasing kinds and quality of data available to them increases opportunities for discrimination.

Algorithms implemented by systems (public or private) will tend to automate decision-making criteria already in place, learning from data that capture how these systems have historically allocated resources or punishment and amplifying the existing incentives for frontline workers, including case managers, judges and loan officers. The problem for civil rights advocates isn't chiefly bias or naivety on the part of those building these models — it's the racism built into the policy and historical practice of many institutions. Data science can be used to combat these problems rather than worsen them, but doing so requires putting these tools in the hands of different people with explicitly different intention.

This challenge is of particular concern to organizations advocating for youth and youth adults, who are more likely to be assessed by tools that treat age as a liability rather than an asset. For example, an analysis of the most common pretrial risk assessment models — Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) Violent Recidivism Risk Score — estimated that fully 60% of its risk score was a function of a person's age rather than any criminogenic factor or case history.[1] This practice runs contrary to what researchers have learned about adolescent development in the past decade: Youth are exceptionally resilient and able to recover from past adversities but also vulnerable to the harm caused by deep engagement with child welfare and justice systems.[2]

Such structural biases in decision making are not unique to risk assessments built on advanced analytics. "Differentiation" and "discrimination" are two sides of the same coin. Agencies and businesses must make decisions about people all the time, often in situations of great uncertainty. In every one of those situations, managers already have screening rules in place — implicit or explicit, simple or complex. The lessons that civil rights organizations and system reformers have learned about implicit and structural bias, organizational incentives and the way they pollute and skew data apply with equal force to this newest generation of analytics tools and predictive algorithms.

What's more, new data and data tools have real potential to do great good. Predictive risk assessments can be fairer and more accurate than the actuarial tools already in wide use throughout government and business. Each time systems introduce new technologies into familiar decision-making processes, advocates get an additional chance to critique and improve how decisions affecting children and families are made — but only if agencies and regulators include them as critical partners.

Many of these predictive models can be thoroughly tested before being implemented. They can help more accurately identify and measure sources of disadvantage, making the mental models used by judges and social workers more explicit and transparent. As prominent social scientists and practitioners argue, "when algorithms are involved, proving discrimination will be easier — or at least it should be, and can be made to be."[3] In many cases, the existing frameworks used to make decisions about youth are so fundamentally biased that improving the accuracy of an assessment disproportionately benefits groups subject to discrimination.

It is true that some data science techniques introduce genuinely new challenges when applied to decisions about people's rights and well-being. Programmers increasingly rely on algorithms to parse and identify patterns in vast reservoirs of data and then use those patterns to make predictions about people. Often called machine learning, this approach can create spurious or dangerous correlations unless data scientists work closely with practitioners to test their conclusions. What's more, the methods involved can generate findings that are difficult to interpret. For example, it can be difficult to explain what combination of characteristics a machine learning algorithm identified as predictive of heart failure. In some contexts, this is less important — just knowing a patient is likely to be at risk can prompt a doctor to provide preventive care. But when making high-stakes decisions about whether a consumer qualifies for a bank loan or a young man is eligible for parole, being able to explain and defend that risk assessment becomes a matter of civil rights.

BASELINES FOR COMPARISONS

In judging whether a new application of advanced analytics is fair and worthwhile, the answer depends largely on how the question is framed.

Is the new tool:

- fair and accurate, or at least as close as possible given the available data?
- an improvement over the process in use now?
- better than any *other* reform that could be applied to this system?

Which combination of these elements constitutes the "right" standard for review is a matter for debate — and it is common for data scientists, public administrators and civil rights advocates to disagree.[4]

Upturn, a Washington, D.C.–based organization that advocates for more equitable use of technology, uses this illustration of data mining creating a spurious correlation: A data scientist might use machine learning techniques to analyze millions of electronic health records and observe that patients with an asthma diagnosis are, surprisingly, *less* likely to die when hospitalized with pneumonia. Does asthma protect against pneumonia? Just the opposite: What the data reflect here is that hospitals rank any asthmatic patient admitted for pneumonia as "high risk" and treat them with corresponding preventive protocols. A data scientist not familiar with hospital protocol would draw exactly the wrong conclusion about the relationship between asthma and pneumonia from the data mining alone.[5]

Advanced analytics also empower agencies and child advocates to create new, never-before-possible solutions. For example, machine learning tools are uniquely helpful for studying how youth move through complex social systems and refining evidence on which programs work for whom, under what circumstances. As Sarah Heller, a leading researcher on strategies for reducing violence among disadvantaged youth, put it: "The real world is often more complicated than one-way tests. ... If it is a combination of characteristics (like age, neighborhood and family income) that defines who is benefiting from a program, we are never going to find that with our standard ways of testing."[6] These tools can help to model large, multivariate problems — like how to simultaneously create a more equitable bus schedule for affluent and low-income communities and rebalance school bell times to create a later start for older youth — that previously could not be solved together.

Despite the potential to use advanced analytics to improve decision making for youth and families, many prominent civil rights organizations in the country have aligned in opposition to them — and a new generation of data-driven civil rights organizations such as Data for Black Lives have grown up specifically to guard against their misuse. Among the most frequently cited reasons for curbing or ending the use of algorithms in decision making are:

- distrust in the capacity, incentives and historical practices of systems that control the underlying data and decision-making processes;
- a lack of transparency, not just about how the models work but also the lack of any notification that these algorithms and "automated decision systems" are even in use; and
- fundamentally different visions for reform that focus on, for example, decriminalization rather than fairer enforcement. Relying on advanced analytics to mitigate near-term problems, reformers worry, might inhibit or foreclose longer-term structural solutions.

While calls to "abolish" this whole category of technology are not likely to work, these concerns are well founded, and organizations and advocates that care about children and families have a vital role to play. Advanced analytical tools are already embedded in business and government operations, and increasingly in state and federal policy. Whether they benefit or harm communities will depend on their design, use and oversight; the technologies themselves are neither "good" nor "bad." It's critical that communities and advocates contribute to a policy framework and set of professional standards that can protect children from negligent or harmful uses of advanced analytics and highlight where these tools can be used to transform ineffective and inequitable systems.

The social sector cannot stay on the sidelines of this debate. Public agencies, philanthropies and policymakers will have to make judgment calls about uses of advanced analytics. Situations that can bring about such judgment calls include:

- the growing number of deliberative city and state task forces on public algorithms;
- calls from citizens and activist groups, such as those organized to "Stop the Cradle to Prison Algorithm" in St. Paul Public Schools or to force police departments to end predictive policing (as happened recently in both Los Angeles and Chicago); and
- a need to modernize our existing approaches to risk assessment, most of which have well-established flaws and biases.

The challenge facing the field is how to distinguish among the useful, acceptable and harmful applications of advanced data analytics, and to help practitioners and advocates use the consequences of these tools for youth, families and communities as the benchmark. This challenge can — and should — surface fierce debate about the risks and opportunities presented by new kinds of algorithmic decision making.

But surprisingly broad agreement exists about the values that should guide how these tools are used and regulated. The implications of these values on policy and practice, explored in the recommendations below, outline a vision for fair and effective use of advanced analytics that looks very different from the investments being made today.

4 PRINCIPLES TO USE ADVANCED ANALYTICS FOR GOOD

expand
OPPORTUNITY

PRINCIPLE I

Most of the established uses of advanced analytics in education, social services and criminal justice direct the attention of counselors and social workers to potential problems facing youth and families — for example, dropouts, abuses and infractions. Interventions are frequently intended to prevent worse harm to the individual, family or community, but very often they result in punishments that disproportionately track along familiar lines of age, gender and race. An alternative and encouraging use of advanced analytics focuses not only on risk but also on identifying so-called odds beaters and new opportunities for youth. Systems also might use advanced analytics to identify untapped talent, sequence more successful pathways into the workplace and — as Los Angeles recently did — expunge old charges from criminal records.

An example of the benefits:

The Children's Data Network at the University of Southern California is helping the state's departments of education and social services move beyond documenting their poor results with high-needs students (especially youth who have been in foster care, experienced homelessness, run away or lived in migrant families) to begin investing in doing better. Their project, Exploring Resilience Among Vulnerable Students in California, explores why some students succeed, despite these negative experiences, and will identify protective factors — individual, family and school experience — that merit more investment.[7]

Good intentions, however, can cause harm if the analytics effort is not well designed.

Chicago police used advanced analytics to develop a "heat list" of individuals more than 500 times as likely as the average city resident to be involved in gun violence, as either a victim or shooter. Despite a commitment to follow up with these mostly young men with prevention services, the police "failed to provide any services or programming" and "instead increased surveillance and arrests."[8] A 2016 study by the RAND Corporation found no perceptible change in gun violence. As the author wrote, "to make a long story short: it didn't work," and the project demonstrated how initiatives aimed at personalizing support for young people who need it can shift on the ground to targeting "chronic offenders" for incarceration.[9]

Recommendation: This is a promising but underexplored area for research and investment. Government and its philanthropic partners, in particular, will need to test whether novel applications of data science (like so-called "precision analytics") can create new insights for the field and under what circumstances it is responsible to apply them. These advanced analytics are likely to be particularly helpful in understanding and tailoring opportunities for complex populations — for example, youth involved in several systems and youth aging out of foster care.

provide
TRANSPARENCY
and evidence

PRINCIPLE 2

The public has a right to know what decisions are being informed or automated by advanced analytical tools, how they have been independently validated and who is accountable for answering and addressing concerns about how they work. These public responsibilities for oversight and regulation cannot fall to private third parties lest, as one close observer put it, "government, which alone is accountable to the public, [becomes] hollowed out, dumb and dark."[10] Equally, the public right to know should not be trumped by trade secrecy claims in commercial contracts. Uses of advanced analytics must earn and maintain a social license to operate, which includes providing clear evidence and ongoing accountability for how these models affect our communities.

An example of the benefits:

Policymakers and advocates have convened task forces in cities like New York City and Pittsburgh and states like Vermont to inventory and set policy around the use of advanced analytical tools. These task forces have attracted community organizing efforts that are remarkably robust, including a public symposium in New York City's Riverside Church, hosted by the NAACP Legal Defense and Educational Fund and attended by legislators, public interest lawyers, tenants-rights organizers and block captains from around the city. Community members and advocates have pushed these task forces to go further, to be more transparent and to recommend substantive, enforceable policies for regulating the selection, implementation and use of advanced analytics. One group, the AI NOW Institute, even wrote a "shadow report" including the recommendations it believed the New York City Automated Decision System Task Force should have included.[11]

Without this kind of transparency, vetting and careful calibration of these tools with local stakeholders, costly mistakes are more likely.

Eckerd Connects built and validated its "Rapid Safety Feedback" tool in Hillsborough County, Florida, to identify children in the care of the child welfare system likely to be at greater risk for injury or death. Subsequently, Eckerd Connects marketed the tool to a number of state agencies for broader use. Child advocates were furious to learn that some states were pursuing its use without having consulted with them. Moreover, researchers worried that Eckerd Connects was slow to provide the level of access to its model that would allow independent verification of its accuracy and validity in these new settings. Illinois, the first state to implement the Rapid Safety Feedback tool, had to rapidly cancel its contract after implementation went badly awry, with the tool assigning thousands of Illinois children a 90% or greater probability of death or injury. At the same time, "predictive analytics (wasn't) predicting any of the bad cases," Illinois Department of Children and Family Services Director BJ Walker told the *Chicago Tribune*.[12]

Recommendation: The local and state task forces above are excellent laboratories for testing how to engage youth and communities about advanced analytics, and for developing initial policy frameworks for regulating their use. These early efforts deserve promotion and study, both to identify successful and scalable strategies and to build a deeper bench of expert advocates who can help communities respond to data initiatives. Public and private funders should avoid supporting private algorithms whose design and performance are shielded by trade secrecy claims, and they should discourage agencies from using them. Procurement processes should disadvantage or disqualify competitors who cannot meet reasonable requirements for auditing and transparency. To create viable alternatives, the philanthropic sector and government should fund and promote efforts by agencies and their research partners to develop, evaluate and adapt transparent and effective models. One well-known effort is the Allegheny Family Screening Tool first used in Allegheny County, Pennsylvania, to screen calls to child protective services. Its development and evaluation have taken place in public view and, accordingly, been the subject of vigorous debate.[13]

EMPOWER
communities

PRINCIPLE 3

Most advanced analytical tools, whether created by business or government, treat children and families as clients, patients and consumers. They aim to assess and treat people, rather than to understand and reform the systems acting upon those people. But the results systems achieve when intervening in the lives of children and families are not only — or even primarily — a function of those families' specific assets or challenges. Results are heavily influenced by local policy decisions and the actions of employers, judges and social workers, which deserve equal scrutiny from these same tools. Focusing analysis and risk assessment on individual profiles, rather than structural barriers to opportunity, is a choice that we should reverse.

The field can build on emerging good practices.

New York City has used advanced analytics to identify landlords who discriminate against potential tenants because of their source of income, in violation of the city's Human Rights Law. The Mayor's Office of Data Analytics "pinpointed areas with available housing, high-performing schools, and little crime, but suspiciously low use of housing vouchers," writes Chris Bousquet for Data Smart Cities at the Harvard Kennedy School of Government's Ash Center.[14] Using this model to target 24 neighborhoods for increased investigation by the Fair Housing Justice Center, New York City lodged 120 income discrimination complaints against landlords from the Council on Human Rights and levied a record-high $100,000 civil penalty. Agencies are not the only actors that can do this: Communities have a right to expect that data collected by public agencies are shared with the public, with appropriate privacy protections, so that citizens can conduct their own analyses.

Too often, this public accountability is deliberately inhibited by claims of trade secrecy and the unwillingness of public agencies to share their data or work.

Advocates and researchers have documented the many fundamental problems with using existing police data to predict crime.[15] Dozens of U.S. police departments continue to defend these tools and to use predictive algorithms to target individuals and neighborhoods for surveillance. Police departments have resisted, however, using these same predictive technologies to examine how citizen complaints and data on officer conduct might anticipate and prevent police use-of-force incidents. At a time when national reform efforts are focused on police use of force and the fitness of individual officers, analytics are being used unilaterally to assess and surveil citizens, especially youth of color, while shielding officers from oversight.

Recommendation: Greater investment is needed to amplify the voices of youth and their communities in debates about the use of advanced analytics — not only to critique existing algorithms but also to access the underlying system data to build alternative models and analyses. Agencies should be encouraged to examine how the characteristics of their own staff and policies can bias their decision making when building advanced analytical tools — for example, analyzing how the age, race or tenure of judges affects their risk assessments. If these tools are valuable to government and business decision makers, they are also valuable in the hands of communities. That sense of reciprocity means that organizations, and government agencies especially, must make the data they collect available for scrutiny and public use.

promote **EQUITABLE** *outcomes*

PRINCIPLE 4

The rationale for introducing complex new technologies into decisions about youth, families and communities is not to make incremental improvements to the status quo. These new investments are only worthwhile if they aim to correct the well-documented bias embedded in the design and use of existing models used by business and government. To a much greater degree than previous risk assessments and decision-making tools, the performance of these advanced analytical models can be tested and shared before they are implemented. We must do so and reject tools that worsen the disproportionately negative treatment of people of color.

An example of the benefits:

In 2013, 10,300 Chicago kids younger than age 6 had elevated blood lead levels. Less than half of children are tested each year, and the harm to them is done before this risk is recognized. There are 200,000 older buildings in Chicago likely to have lead paint, disproportionately housing Black and brown families.[16] To reduce the generational damage that lead has done to these communities, the University of Chicago used two decades of data on childhood lead tests, lead inspections and housing data to model the risk for specific neighborhoods and buildings.[17] The city is proactively targeting these areas for public service advertisements, outreach and inspections. Similar preventive work is underway in cities like Cleveland and Philadelphia.

An example of the benefits:

In St. Paul, Minnesota, a diverse group of youth and community advocates fought for a year to stop a school-county effort to screen each of the county's youth for risk of juvenile delinquency based on ZIP codes, income, truancy data, race and other indicators. Citing the wide and longstanding disparity in the school district's disciplining of students of color, advocates successfully argued that without a commitment to reduce systems' disproportionate harm to communities of color, the analytics project should be terminated.[18]

Recommendation: The expectation that advanced analytical tools should be introduced only when they reduce the opportunity deficit for disadvantaged groups would have far-reaching consequences. It will take organizing and advocacy to establish that expectation and new policy development to begin to institutionalize and enforce it. This strategy has an increasingly firm theoretical and academic foundation but few practical tools and examples. Philanthropy and government can help communities test and improve upon those that do exist — for example, the Algorithmic Equity Toolkit from Washington State's ACLU[19] — and promote the use of equity audits from pioneering firms like O'Neil Risk Consulting & Algorithmic Auditing (ORCAA).[20]

conclusion

Advocates for children and families have an opportunity to shape the rapidly evolving field of advanced analytics and create frameworks that make these tools work for, not against, young people.

Schools, nonprofit service providers and public agencies collect an immense amount of information that goes largely unused, except for billing and administration. These data can be used instead to get teachers on-track indicators for their students, to help social workers identify children that urgently need help (and not intrude into the family lives of those who do not) and to speed the expungement of criminal records for nonviolent offenders. In the hands of communities and regulators, these data can — and should — identify, measure and provide tools to correct the bias embedded in decisions by police, courts, banks and insurers, which fall disproportionately on the shoulders of Americans who are low-income and of color. Advanced analytics are a tool, just like data, that organizations dedicated to improving outcomes for youth and young adults can use to build a shared understanding of their conditions and needs, to target systemic barriers and inequities and to achieve better results.

The field, however, should urgently listen to the warnings of civil rights groups and community organizers who have been monitoring the early expansion of these tools and have documented how algorithms can contribute to laundering prejudice into "scientific" risk and remove key decisions affecting families from public debate. If this new generation of data analytics is created by the same organizations with the same incentives to serve the same vested interests as the last generation of tools, the results for children and families will not improve and decisions about credit, incarceration and employment will disadvantage families even further.

Philanthropy, the social sector and policymakers have an opportunity to safeguard data, technology and advanced analytics for the public good. Both proponents and critics of these tools, which are already deeply embedded in decisions affecting families, broadly agree on the principles that should guide their development and regulation. Decision makers at every level can take action today to orient these tools toward equity, transparency, empowerment and opportunity. There are already organizers, data activists and social entrepreneurs blazing a path for the field to follow, many of them youth and young adults who are keenly aware of the stakes of this debate for their communities.

Affordable Housing Protection

- To preserve rent-controlled units in its high-priced real estate market, New York City created a model to predict families in units likely to come under illegal pressure to vacate by landlords and to identify that harassment in something near real time. The model determined that complaints about dust and asbestos, illegal work and 311 calls related to construction and air quality were highly correlated with the loss of rent-regulated units in the following year, and the city began prioritizing inspections associated with those complaints. The targeted inspection resulted in an $8 million suit against a prominent developer.[21]

Criminal Record Expungement

- In 2019, Code for America partnered with the District Attorney in Los Angeles to erase 62,000 felony convictions for marijuana possession with an algorithm that can scan thousands of criminal records each second to identify those eligible for expungement.[22]

Evidence-Based Sentencing

- Courts have sought to apply risk assessment instruments to sentencing for more than half a century to aid judges and to improve the consistency and fairness of decisions. The digitization of records within criminal justice and availability of software to parse that data have contributed to a shift toward so-called "evidence-based sentencing" aimed at assessing how likely a person is to reoffend and assigning punitive or rehabilitative services accordingly. Civil rights organizations and researchers have argued that these tools violate defendants' rights to due process unless their fairness and accuracy cannot be challenged in court. Journalistic work like the Pro Publica series "Machine Bias" have documented significant concerns about the quality of the underlying algorithms in commercially available risk assessments — and, perhaps more significantly, the degree to which defendants, their counsel and the public are denied access to these models' code and training data.[23] In the 2016 case *State v. Loomis,* however, the Supreme Court of Wisconsin ruled use of these sentencing tools legal, on the theory that they provide only one of many inputs into a broader set of factors considered by judges.[24] The Supreme Court declined to hear an appeal of this decision.

Higher Education

- Georgia State University (GSU) has transformed itself from a "night school for white businessmen"[25] in the 1960s to one of the country's leading institutions awarding bachelor's degrees to African Americans. Predictive analytics have been an important part of that success. The university's GPS Advising platform qualifies students who meet certain criteria for additional support — such as summer session, financial literacy coaching, micrograants — and is monitored constantly by advisors. These advisors can create "early alert" holds on students' accounts during the first six weeks of the semester to require one-on-one meetings with a coach. GSU's graduation rate has increased from 32% (2013) to 54% (2017) as enrollment of both Pell Grant-receiving students and students of color has increased. More generally, learning analytics is already a $500 million market serviced by at least 30 for-profit companies, including EAB, Civitas Learning and Hobsons (GSU's vendor).

Insurance Pricing

- Commercial insurers are increasingly looking to external sources of data available about customers to more accurately assess their actuarial risk. This data mining narrows risk pools and drives down costs for some but has "a significant potential negative impact on the availability and affordability of life insurance for protected classes of consumers." Those classes included "race, color, creed, national origin . . . and sexual orientation."[26]

Job Awareness

- Rather than being conducted through classified ads or job centers, matches between workers and employment opportunities increasingly are being mediated by platforms like Zip Recruiter, LinkedIn and Facebook. The machine learning in use by these platforms makes independent assumptions about which positions might be attractive to each of its users.[27] In a recent test of Facebook, ads for a house cleaner and construction worker were shown almost exclusively to women and men, respectively, even though the ad was purchased with no expressed gender preference.[28] This kind of invisible algorithmic bias does not require intent.

- Alternatively, tools like Textio are beginning to leverage artificial intelligence to help employers use language to attract more diverse candidate pools and deliberately mitigate bias in job advertisements that historically discourage applications from women and people of color.[29]

Policing

- Police use neighborhood- and person-based approaches to predicting crime, both of which have well-documented problems[30] related to bad data, bias and little-to-no evidence of efficacy.[31] The so-called chronic offender bulletins produced by Palantir for Los Angeles's Operation LASER were so relentless that some residents reported being stopped by police as many as 11 to 30 times *per week*. LASER was abandoned in April 2019.[32] Lack of police transparency about these new tools with citizens and even city legislators[33] has prompted major controversy in several cities.

- Conversely, Charlotte has used some of these same predictive techniques to forecast the likelihood of officer misconduct,[34] and advocates in Phoenix, Arizona, have tried to do the same over the objections of the Fraternal Order of Police.[35]

Pretrial Risk Assessment

- In 2017, New Jersey replaced its cash bail system with a predictive model using just nine factors to generate a risk score that advises a judge whether to release or jail someone arrested, pending trial. As a result, thousands of New Jersey residents with low incomes were released from jail, as crime following the first six months of the risk assessment's implementation declined.[36] California subsequently passed a law, SB 10, which replaced the state's cash bail system with one based on risk assessment.[37] Late in 2018, President Donald Trump signed into law the First Step Act, which directed the Justice Department to develop its own risk assessment tool for the Federal Bureau of Prisons, known as PATTERN (Prisoner Assessment Tool Targeting Estimated Risk and Need).[38] The civil rights community, led by the Leadership Conference on Civil and Human Rights, opposed each of these moves. In early 2020, the Pretrial Justice Institute reversed its support for pretrial risk assessment as an element of its "smart" pretrial justice framework, writing, "We made a mistake — we did not have the right people at the table when we were designing our roadmap to decarceration."[39]

School Assignment And Transportation

- Algorithms also increasingly play a role in school assignments for students and how they get to school. New York City high schools use a "deferred acceptance algorithm" to match students and schools based on their mutual preference.[40] Meant to provide a fairer shot of placement in the city's selective high schools for students of color who are low-income, the screening system has expanded to fully one in five middle and high schools in New York City and has been criticized for replacing "tracking within schools with tracking *by* school."[41]

- In 2014, Boston used a new geographically driven algorithm to create "choice baskets" for families to preserve the city's school choice system, while reducing some of the strain associated with long busing times between neighborhoods. The "home-based assignment system," however, could not overcome Boston's very unequal social geography; there simply weren't enough good schools accessible to high-poverty neighborhoods.[42] In 2017, Boston Public Schools went further, using MIT's expertise and machine learning to redesign their bus routes, provide more equitable travel times across different neighborhoods[43] and change school bell times to start the school day later in line with recommendations from adolescent psychology. The solution provided by MIT was technically impressive — the district judged that it would save $15 million and deliver on promises of more equitable travel times and better school schedules — but it was rejected by parents who had not been consulted as a part of the algorithm's development and would have been disadvantaged by the changes.

Tenant Screening

- New digital advertising models create "target audiences" for apartment listings that can deliberately or inadvertently exclude certain groups, including protected classes (discrimination by age, gender, race or income). For example, Facebook has been sued by the Department of Housing and Urban Development (HUD) for housing discrimination because its tool encourages landlords to create "Lookalike Groups" that Facebook uses to make automated decisions about which of its users can view the advertised property for rent.[44]

Endnotes

1 Stevenson, M. T., & Slobogin, C. (2018, August 2). Algorithmic risk assessments and the double-edged sword of youth. *Washington University Law Review, 96*(3), 681–706. Retrieved September 24, 2020, from https://openscholarship.wustl.edu/law_lawreview/vol96/iss3/6

2 National Academies of Sciences, Engineering, and Medicine. (2019). *The promise of adolescence: realizing opportunity for all youth.* Washington, DC: The National Academies Press. https://doi.org/10.17226/25388

3 Kleinberg, J., Ludwig J., Mullainathan, S., & Sunstein, C. R. (2018). Discrimination in the age of algorithms. *Journal of Legal Analysis, 10.* Retrieved from www.nber.org/papers/w25548

4 Partnership on AI. (2019, April 29). *Report on algorithmic risk assessment tools in the U.S. criminal justice system.* Retrieved from www.partnershiponai.org/report-on-machine-learning-in-risk-assessment-tools-in-the-u-s-criminal-justice-system

5 Rieke, A., Bogen, M., & Robinson, D. G. (2018). *Public scrutiny of automated decisions: early lessons and emerging methods.* Washington, DC: Upturn and Omidyar Network. Retrieved from https://omidyar.com/wp-content/uploads/2020/09/Public-Scrutiny-of-Automated-Decisions.pdf

6 Doleac, J. (Host). (2020, March 17). *Episode 25: Sara Heller* (Audio podcast). Probable Causation. Retrieved from www.probablecausation.com/podcasts/episode-25-sara-heller

7 Children's Data Network. (2020). *Exploring resilience among vulnerable students in California.* Retrieved from www.datanetwork.org/research/exploring-resilience-among-vulnerable-students-in-california

8 Saunders, J. (2016, October 11). Pitfalls of predictive policing. *The RAND Blog.* Retrieved from www.rand.org/blog/2016/10/pitfalls-of-predictive-policing.html

9 Saunders, J. (2016, October 11).

10 Brauneis, R., & Goodman, E. P. (2018). Algorithmic transparency for the smart city. *Yale J.L. & Tech, 20*(1), 103–176. Retrieved from https://digitalcommons.law.yale.edu/yjolt/vol20/iss1/3. And, Goodman, E. P. (2019, February). The challenge of equitable algorithmic change. *The Regulatory Review.* Retrieved from www.theregreview.org/wp-content/uploads/2019/02/Goodman-The-Challenge-of-Equitable-Algorithmic-Change.pdf

11 Richardson, R. (Ed.) (2019, December 4). *Confronting black boxes: A shadow report of the New York City Automated Decision System Task Force.* New York, NY: AI NOW Institute, New York University. Retrieved from https://ainowinstitute.org/ads-shadowreport-2019.pdf

12 Jackson, D., & Marx, G. (2017, December 6). Data mining program designed to predict child abuse proves unreliable, DCFS says. *Chicago Tribune.* Retrieved September 25, 2020, from www.chicagotribune.com/investigations/ct-dcfs-eckerd-met-20171206-story.html

13 Allegheny County. (n.d.). *The Allegheny family screening tool: Predictive risk modeling in child welfare in Allegheny County.* Retrieved from https://alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx

14 Bousquet, C. (2018, February 28). How New York is protecting affordable apartments with analytics. *Data-Smart City Solutions.* Retrieved from https://datasmart.ash.harvard.edu/news/article/how-new-york-protecting-affordable-apartments-analytics

15 Richardson, R., Schultz, J. M., & Crawford, K. (2019, February 13). Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *NYUL Rev. Online, 94,* 15. Retrieved from www.nyulawreview.org/wp-content/uploads/2019/04/NYULawReview-94-Richardson_etal-FIN.pdf. And, Robinson, D. G. (2017). The challenges of prediction: Lessons from criminal justice. *ISJLP, 14*(2), 151–186. Retrieved from https://kb.osu.edu/bitstream/handle/1811/87012/ISJLP_V14N2_151.pdf?sequence=1

16 Potash, E., Brew, J., Loewi, A., Majumdar, S., Reece, A., Walsh, J.,…Ghani, R. (2015). *Predictive modeling for public health: Preventing childhood lead poisoning.* In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 2039–2047). Retrieved from www.researchgate.net/profile/Subhabrata_Majumdar/publication/281528176_Predictive_Modeling_for_Public_Health_Preventing_Childhood_Lead_Poisoning/links/5a7ca4f8a6fdccc013f323a8/Predictive-Modeling-for-Public-Health-Preventing-Childhood-Lead-Poisoning.pdf. And, Data Across Sectors for Health. (n.d.). *Childhood lead paint hazard data sharing, Chicago, IL* (DASH project profile). Retrieved from http://dashconnect.org/wp-content/uploads/2017/03/Chicago.pdf

17 Data Across Sectors for Health. (n.d.).

18 Melo, F. (2019, January 28). St. Paul, Ramsey County to end youth data-sharing agreement after withering criticism. *Pioneer Press.* Retrieved September 25, 2020, from www.twincities.com/2019/01/28/st-paul-ramsey-county-to-end-youth-data-sharing-agreement-after-withering-criticism

19 American Civil Liberties Union of Washington. (n.d.). *Algorithmic equity toolkit.* Retrieved from www.aclu-wa.org/AEKit

20 O'Neil Risk Consulting & Algorithmic Auditing. (n.d.) *What we do.* Retrieved from https://orcaarisk.com

21 Bousquet, C. (2018, February 28).

22 Tchekmedyian, A., & Miller, L. (2020, February 13). L.A. County District Attorney Jackie Lacey announces dismissal of 66,000 marijuana convictions. *Los Angeles Times.* Retrieved September 25, 2020, from www.latimes.com/california/story/2020-02-13/la-district-attorney-clear-marijuana-convictions

23 Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine bias. *ProPublica.* Retrieved from www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

24 Kehl, D., Guo, P., & Kessler, S. (2017). *Algorithms in the criminal justice system: Assessing the use of risk assessments in sentencing.* Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School. http://nrs.harvard.edu/urn-3:HUL.InstRepos:33746041

25 Fausset, R. (2018, May 15). Georgia State, leading U.S. in Black graduates, is engine of social mobility. *The New York Times.* Retrieved September 25, 2020, from www.nytimes.com/2018/05/15/us/georgia-state-african-americans.html

26 Ralph, O. (2020, February 17). Big data opens up vast new possibilities for insurers. *The Financial Times.* Retrieved September 25, 2020, from www.ft.com/content/d7000d78-519c-11ea-90ad-25e377c0ee1f

27 For extensive discussion of this and other sources of bias within hiring and employment, see Bogen, M., & Rieke, A. (2018, December). Help wanted: An examination of hiring algorithms, equity and bias. *Upturn.* Retrieved from www.upturn.org/reports/2018/hiring-algorithms

28 Bogen, M., & Rieke, A. (2018, December).

29 Collier, D., & Zhang, C. (2016, Fall). *Can we reduce bias in the recruiting process and broaden/diversify pools of candidates by using different types of words/styles in job descriptions?* Ithaca, NY: Cornell University. Retrieved September 25, 2020, from http://digitalcommons.ilr.cornell.edu/student/140

30 Richardson, R., Schultz, J. M., & Crawford, K. (2019, February 13).

31 Saunders, J. (2016, October 11).

32 Puente, M. (2019, April 12). LAPD ends another data-driven crime program touted to target violent offenders. *Los Angeles Times.* Retrieved September 25, 2020, from www.latimes.com/local/lanow/la-me-laser-lapd-crime-data-program-20190412-story.html

33 Winston, A. (2018, February 27). Palantir has secretly been using New Orleans to test its predictive policing technology.

The Verge. Retrieved September 25, 2020, from www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd

34 Arthur, R. (2016, March 9). We now have algorithms to predict police misconduct. Will police departments use them? *FiveThirtyEight.* Retrieved September 25, 2020, from https://fivethirtyeight.com/features/we-now-have-algorithms-to-predict-police-misconduct

35 Price, J. (2019, August 29). Phoenix police want to predict officer misconduct. Why 'purged' records may make that impossible. *The Arizona Republic.* Retrieved September 25, 2020, from www.azcentral.com/story/news/local/arizona-investigations/2019/08/29/phoenix-police-want-predict-officer-misconduct-why-purged-records-may-make-impossible/2130810001

36 Golob, C. A., Redcross, C., & Valentine, E. J. (2019, November). *Evaluation of pretrial justice system reforms that use the public safety assessment.* New York, NY: MDRC. www.mdrc.org/publication/evaluation-pretrial-justice-system-reforms-use-public-safety-assessment-0. And, Dabruzzo, D. (2019, November 14). New Jersey set out to reform its cash bail system. Now, the results are in. *ArnoldVentures.org.* Retrieved September 25, 2020, from www.arnoldventures.org/stories/new-jersey-set-out-to-reform-its-cash-bail-system-now-the-results-are-in

37 Judicial Council of California. (n.d.). *SB 10: Pretrial release and detention.* Retrieved September 25, 2020, from www.courts.ca.gov/pretrial.htm

38 Federal Bureau of Prisons. (n.d.). *An overview of the First Step Act.* Retrieved September 25, 2020, from www.bop.gov/inmates/fsa/overview.jsp

39 Pretrial Justice Institute. (2020, February 7). *Updated position on pretrial risk assessment tools.* Retrieved from www.pretrial.org/wp-content/uploads/Risk-Statement-PJI-2020.pdf

40 Roth, A. (2015, July 2). Why New York City's high school admissions process only works most of the time. *Chalkbeat.* Retrieved from https://ny.chalkbeat.org/2015/7/2/21096226/why-new-york-city-s-high-school-admissions-process-only-works-most-of-the-time

41 Hu, W., & Harris, E. A. (2018, June 18). A shadow system feeds segregation in New York City schools. *The New York Times.* Retrieved September 25, 2020, from www.nytimes.com/2018/06/17/nyregion/public-schools-screening-admission.html

42 O'Brien, D. T., Hill, N. E., Contreras, M., Phillips, N. E., & Sidoni, G. (2018, July). *An evaluation of equity in the Boston Public Schools' home-based assignment policy.* Boston, MA: Boston Area Research Initiative. Retrieved from https://news.northeastern.edu/wp-content/uploads/2018/07/BPSHBAP.pdf

43 Goodman, E. P. (2019, February 12). Smart algorithmic change requires a collaborative political process. *The Regulatory Review.* Retrieved from www.theregreview.org/2019/02/12/goodman-smart-algorithmic-change-requires-collaborative-political-process

44 Jan, T., & Dwoskin, E. (2019, March 28). HUD is reviewing Twitter's and Google's ad practices as part of housing discrimination probe. *The Washington Post.* Retrieved September 25, 2020, from www.washingtonpost.com/business/2019/03/28/hud-charges-facebook-with-housing-discrimination

Design & Production: Shagas Design / Photography: © Jason Miczek

701 ST. PAUL STREET BALTIMORE, MD 21202 410.547.6600 WWW.AECF.ORG