HARNESSING AI INNOVATION FOR STRUGGLING FAMILIES

Jacqueline G. Schafer†

Abstract

State child welfare systems impact one of the oldest fundamental liberty interests recognized by the U.S. Supreme Court—the interest of parents in the care, custody, and control of their children—and alter family bonds for hundreds of thousands of vulnerable children across the U.S. An ethical child welfare system demands that decisions about how to help struggling families be data-informed, yet states have been painfully slow to acknowledge the potential for existing and emerging technology tools to transform their operations.

This article starts from the premise that child welfare systems and the families they interact with could benefit immensely if cutting-edge, private sector technology innovation could be applied to the vast social science datasets generated over the life of a state child welfare case. However, to realize these benefits, Congress must update the multi-layered regime of federal laws governing child welfare data to require as a condition of funding that states increase data scientists' and researchers' access to this data. And importantly, the government must increase investment in the technological infrastructure that can enable artificial intelligence (AI)-enabled applications to identify the most effective interventions and reduce the currently enormous administrative burden on child welfare system workers. Moreover, state child welfare agencies must actively prepare to address the complex ethical and privacy questions raised by the inevitable introduction of AI-enabled technology into the practice of social work.

This article will (1) provide an overview of the data that is currently stored and collected by state child welfare systems; (2) describe the complex set of

† Jacqueline Schafer is the Founder and CEO of ClearBrief.ai, a legal technology company that leverages artificial intelligence to improve legal writing and the justice system. She previously served as Senior Director, Program Counsel for the national child welfare nonprofit Casey Family Programs, advising on Casey's data-informed consulting and agreements with state agencies across the country. Prior to that role, she served as an Assistant Attorney General in the Washington and Alaska Attorney General's Offices, where she advised the leadership of each state's child welfare agency and specialized in appellate practice and complex litigation. Jacqueline began her career as a litigation associate at the New York law firm of Paul, Weiss, Rifkind, Wharton, & Garrison. She is a founding board member of the Washington State CASA Association, which trains volunteer advocates for foster youth, and a board member of the Seattle Repertory Jazz Orchestra. She also serves as an advisor to two AI organizations focused on social impact AI innovation, X4Impact and Think of Us.

I am grateful to Marie-Amelie George, Peter Pecora, Jay Bartot, Ben Elowitz, Shelly Cano Kurtz, Dr. Ying Li, Luis Salazar, and Tovi Newman for their helpful comments and conversations.

federal laws that restrict sharing of this information and propose legislative and regulation changes that, if implemented, would foster technological innovation that could be life-changing for children and families; (3) suggest currently feasible machine learning applications that would benefit tech-optimized child welfare systems; (4) describe the practical steps needed to ready state child welfare agencies to implement technology innovation; and (5) analyze privacy considerations in adopting AI technologies.

Overall, this article provides a roadmap for the U.S. Department of Health and Human Services in complying with Section 5 of President Trump's February 11, 2019 Executive Order on AI, which commands all heads of federal agencies to "review their Federal data and models to identify opportunities to increase access and use by the greater non-Federal AI research community in a manner that benefits that community, while protecting safety, security, privacy, and confidentiality."[1] *Ideally, this analysis will inspire government, private sector, and nonprofit leaders to recognize the need for a coordinated investment in technological transformation that reflects the urgency of child welfare cases.*

TABLE OF CONTENTS

1. Exec. Order No. 13859, 84 C.F.R. § 3967 (2019), https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence.

INTRODUCTION

State child welfare systems are rich in data: they store narrative details about every aspect of a child and family's life, from the first hotline call about a potentially abused or neglected child, to the final social worker interactions when a case is closed. Admittedly, this data may reflect the individual biases of the child protection and foster care system employees tasked with documenting the details of an incredibly traumatic period for a struggling family. Although no dataset is perfect or unbiased, the data stored in unique state computer systems across the country holds significant promise for understanding how we can better serve the families who interact with state child welfare agencies. This article begins from the premise that if data scientists and researchers had easier access to these enormous, varied datasets collected by child welfare systems in each state, they could process the data and use it to unlock real-time insights. Those insights could improve how these systems engage with American families.

Accordingly, it is imperative for Congress to update the federal laws and regulations governing the confidentiality of state child welfare systems data in order to reference a single set of federal confidentiality requirements, explicitly include a requirement that states receiving federal funding must allow research access to de-identified child welfare data, and identify a uniform process for requesting such data that balances privacy concerns. These reforms, if appropriately supported with federal funds, would simplify state responses to information-sharing requests and help unlock the potential of emerging technology.

Specifically, if Congress allowed for greater research access, data scientists from the public and private sectors could apply artificial intelligence (AI) tools to these datasets with the goal of reducing social workers' administrative burdens and enhancing the positive impact of their efforts to heal families. The term AI is often used interchangeably with machine learning and predictive analytics, but it encompasses any computational process or product that appears to demonstrate intelligence and generate insights through non-biological processes.[2] AI is currently used throughout the world to perform critical tasks on large datasets, and is often called machine learning when describing "a method to uncover statistical correlations within a dataset [that] could range from simple linear regression to more complicated algorithms."[3] In other words, machine learning can detect patterns between input and output variables using a statistical and mathematical modeling approach.[4] The central questions that machine learning seeks to answer are, "how can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes?"[5]

One representative example of AI from the health care world is when advanced programs help radiologists interpret MRIs by identifying patterns in images of cancerous and non-cancerous breast lesions.[6] Researchers developed an AI diagnostic tool by training an algorithm on a vast health care database of images that allowed it to detect patterns that no individual human radiologist would be able to discern on her own.[7] The AI program—an example of AI being used to aid, not replace, human decision-making—led to a 39% reduction in missed breast cancers, as well as a 20% overall diagnostic improvement.[8] This is just one of numerous AI applications discussed *infra* in this article that utilize large datasets to solve our most pressing societal problems. Amid the growing chorus of academic leaders urging policymakers to view child maltreatment as

2. *Annotation and Benchmarking on Understanding and Transparency of Machine Learning Lifecycles (ABOUT ML)*, PARTNERSHIP ON AI, https://www.partnershiponai.org/wp-content/uploads/2019/07/Supplement-MLPrimer.pdf (explaining "AI can now do important tasks, including identifying and deciphering the objects in images ('computer vision' ('CV')), interpreting text in various ways ('natural language processing ('NLP')'), and controlling robots or game agents via strong feedback loops ('reinforcement learning')).

3. *Id*.

4. *Id*.

5. Jason Togyer, *Research Notebook: The Discipline of Machine Learning*, THE LINK, https://www.cs.cmu.edu/link/research-notebook-discipline-machine-learning.

6. Melissa Locker, *This AI Breast Cancer Diagnostic Tool Is the First to Get FDA Clearance*, FAST COMPANY (July 17, 2019), https://www.fastcompany.com/90377791/quantx-is-first-ai-breast-cancer-diagnostic-tool-cleared-by-fda.

7. *See* Koichiro Yasaka & Osamu Abe, *Deep Learning and Artificial Intelligence in Radiology: Current Applications and Future Directions*, 15 PLOS MEDICINE 1 (Nov. 30, 2018) (explaining that "[c]hest radiographs are one of the most utilized radiological modalities in the world and have been collected into a number of large datasets currently available to machine learning researchers.").

8. *Id.*

a public health issue,[9] innovative governments should examine how AI innovation could hold the same promise for child welfare as it does for healthcare.

This article posits that one of the key reasons that state child welfare leaders have not pressured the federal government to update the applicable data-sharing laws and provide increased support for technology innovation is that child welfare leaders and social workers have not been shown a clear picture of how AI could transform their work. This article attempts to glimpse into the future of how child welfare systems could be reimagined with AI tools that are currently in use in the private sector, and thus proceeds in four parts.

Part I gives an overview of how data is currently stored in state child welfare agencies, and the minimal uniform datasets that states currently share with the federal government. These discrete data points, which are stale by the time they are delivered to the federal government, and even more so by the time they are released publicly to researchers, currently underpin all federal child welfare policy decisions.

Part II demonstrates the challenge that every state child welfare agency currently faces when determining whether and for what purposes it can share the vast amounts of child welfare data in its possession. This section explains the overlapping set of federal laws that dictate how states approach confidentiality of child welfare system information and make it unnecessarily complicated to share data to improve operations.

Part III explores examples of how two states are starting to harness AI to improve government social services, and it also discusses examples of arguably the most problematic uses of AI in child welfare. This Part goes on to propose several novel applications of AI that could improve core child welfare operations and outcomes.

Part IV addresses the practical steps and challenges beyond changes to the data-sharing statutes that are needed to facilitate the widespread use of AI to improve social work.

And finally, Part V discusses the privacy issues at stake in this discussion and highlights emerging data science techniques that could be utilized to minimize risk of increasing research access to child welfare datasets. Part V also recommends formation of an advisory committee that could explore the issues raised in this article.

The federal government has touted the virtues of data-driven child welfare practice,[10] but as discussed below, it has not "walked the walk". Meaning the federal government has not taken the legal and funding steps that would make data-driven practice a reality for every state. Instead of requiring individual states to put forward nearly $50 million dollars of funding to access federal

---

9. *See, e.g.*, Todd I. Herrenkohl et al., *The Public Health Model of Child Maltreatment Prevention*, 17 TRAUMA, VIOLENCE, & ABUSE 363 (2016), https://doi.org/10.1177/1524838016661034 (highlighting the reasoning behind qualifying child maltreatment as a public health issue).

10. JAMES BELL ASSOCIATES, U.S. DEP'T OF HEALTH AND HUM. SERVS., GUIDE TO DATA-DRIVEN DECISION MAKING: USING DATA TO INFORM PRACTICE & POLICY DECISIONS IN CHILD WELFARE ORGANIZATIONS (Mar. 2018), https://www.acf.hhs.gov/sites/default/files/cb/guide_to_dddm.pdf.

matching funds to build a modern database capable of running AI tools,[11] the federal government, nonprofit organizations, and the private sector should work together with states to develop a prototype system and license it free of charge. The intent of this article is to focus the attention of federal and state policy leaders on what exciting innovations are possible in the future to motivate the U.S. to take the necessary steps in the present.

I. ONLY A SMALL FRACTION OF THE DATA THAT STATES POSSESS ABOUT CHILD WELFARE IS SHARED WITH THE FEDERAL GOVERNMENT AND MADE AVAILABLE TO RESEARCHERS.

Every state[12] child welfare agency collects and stores a vast array of information about child abuse and neglect cases, from the moment a call is made to a child abuse hotline to the moment a child ages out of the state foster care system or finds a permanent home.[13] It is important to understand the difference between the uniform data sets that states provide to the federal government for evaluation, which are also used to identify trends in child welfare outcomes and eventually become publicly available to researchers, and all of the remaining data stored in connection with these children and families as they interact with the state child welfare system.

Since 1993, state and tribal child welfare agencies have stored their information about child abuse, neglect allegations, and subsequent child welfare cases in databases the federal government calls Statewide Automated Child Welfare Information Systems (commonly called "SACWIS systems").[14] Despite the fact that by 2015, fewer than half of states were able to implement SACWIS systems certified as compliant with federal standards,[15] the federal government issued a final rule in 2016 encouraging states to replace SACWIS systems with new Comprehensive Child Welfare Information Systems (commonly referred to as "CCWIS systems").[16] CCWIS systems have updated

11. *See infra* Section IV(A).

12. For the purposes of this Article, I will refer mainly to state child welfare systems, although the principles described could be equally applied to county and tribal child welfare systems. Some tribes have begun to gather and store tribal child welfare data in ways that allow for greater analysis. *See* Malia Villegas ET AL., NAT'L CONG. OF AM. INDIANS, ROBERT WOOD JOHNSON FOUNDATION, DISAGGREGATING AMERICAN INDIAN & ALASKA NATIVE DATA: A REVIEW OF LITERATURE (July 2016), http://www.ncai.org/DataDisaggregationAIAN-report_5_2018.pdf (recommending and discussing tribal data collection).

13. *See* Jill Goldman & Marsha K. Salus, *A Coordinated Response to Child Abuse and Neglect: The Foundation for Practice*, U.S. DEP'T OF HEALTH & HUM. SERVS. (2003), https://www.childwelfare.gov/pubPDFs/foundation.pdf (outlining procedural steps between when a report is initially made and case closure).

14. Teresa M. Harrison et al., *A Tale of Two Information Systems: Transitioning to a Data-Centric Information System for Child Welfare*, DIGITAL GOV'T SOC'Y 1, 1 (June 1, 2018), https://www.ctg.albany.edu/media/pubs/pdfs/transitioningtoadata-centricinformationsystemforchild_welfare.pdf; *see also* 42 U.S.C. § 5106a(b)(2)(B)(xxiii) (2018) (requiring states to submit a state plan to be eligible for grants that provides for a statewide technology system that "track[s] reports of child abuse and neglect from intake through final disposition.").

15. Harrison, *supra* note 14; *SACWIS Status*, CHILD. BUREAU, U.S. DEP'T OF HEALTH & HUM. SERVS. (last updated Apr. 8, 2020), https://www.acf.hhs.gov/cb/resource/sacwis-status (noting various stages of SACWIS compliance as of 2015) [hereinafter *SACWIS Status*].

16. Comprehensive Child Welfare Information System, 81 Fed. Reg. 35,450 (June 2, 2016) (codified at 45 C.F.R. pt. 95); *SACWIS Status*, *supra* note 15 (noting various stages of SACWIS compliance as of 2015).

data requirements designed to improve the ability of states to share information with the federal government in connection with federal funding requirements, as well as with the state's other social services systems.[17] The major differences between CCWIS and SACWIS systems are that the former must be modular, or have built in modules that are easier to upgrade or change without huge expense.[18] CCWIS systems must also have the capacity to interface with other government databases, and they must be able to export information about the CCWIS software to the federal government to allow them to assist other states.[19] No state has yet implemented a CCWIS system, as discussed in greater detail below.[20]

Child welfare databases house data encompassing the entirety of a state's contacts with children and families. Caseworkers describe each interaction they have with the child and family until the case ends by adding case notes, detailed descriptions of physical and emotional abuse, photographs, recordings, medical information, descriptions of services provided, and countless other details about the child and family.[21] These data points are saved in the system in different native formats (for example, as typed narratives in case notes, or uploaded as photos or pdf documents).[22] While there are many common elements across states, each state's system is unique and collects information that is specific to state laws, practice models, and software design.[23]

The federal government, through its spending clause power wielded via funds provided to states under the Social Security Act (SSA) and the Child Abuse Prevention and Treatment Act, incentivizes states to gather only a small fraction of these data points from SACWIS systems and provide them to Health and Human Services twice a year in a standard format.[24] For the fiscal year

17. *CCWIS Final Rule Overview*, CHILD. BUREAU, U.S. DEP'T OF HEALTH & HUM. SERVS. (June 2016), https://www.acf.hhs.gov/sites/default/files/cb/ccwis_overview_presentation.pdf [hereinafter *CCWIS Final Rule*].

18. John Kelly, *Feds Chart Future for Antique Data Management Systems*, CHRON. OF SOC. CHANGE (Mar. 29, 2017), https://chronicleofsocialchange.org/subscriber-content/replacing-sacwis-feds-offer-funds-for-better-data-management-but-will-systems-buy-in/25755; 45 C.F.R. § 95 (2019).

19. Kelly, *supra* note18; 45 C.F.R. § 95.

20. *CCWIS Status*, CHILD. BUREAU, U.S. DEP'T OF HEALTH AND HUM. SERVS. (last updated Sept. 15, 2020), https://www.acf.hhs.gov/cb/resource/ccwis-status [hereinafter *CCWIS Status*].

21. *See, e.g.*, Letter from Gregory McKay, Director of Arizona Department of Child Safety, to Douglas A. Ducey, Governor of Arizona (Sept. 4, 2018) (on file with author) (describing Arizona's legacy SACWIS system, which houses "over 450 screen displays containing embedded logic to support the work functions of the Department . . . including the creation, control, and management of clients, intake functions, ongoing case management, development of new interfaces for data mining, mobility access options, visitation report entries, court record production, Business Intelligence (BI) processing capabilities, comprehensive reporting, decisions support processing, and general system enhancements.").

22. *See, e.g.*, Theo Douglas, *Washington Streamlines Child Welfare with New Apps, Portal and Hardware*, GOV'T TECH. (June 12, 2017), http://www.govtech.com/computing/Washington-Streamlines-Child-Welfare-with-New-Apps-Portal-and-Hardware.html (discussing variety of file types and information categories uploaded into Washington state's SACWIS system).

23. *See CCWIS Status*, *supra* note 20 (emphasizing disparities in meeting CCWIS standards across every major state and tribe); *see also SACWIS Status*, *supra* note 15 (reporting which states are SACWIS-compliant, which are in the process of complying, and which adopt a non-SACWIS model).

24. 42 U.S.C. § 5104(c)(1)(D) (2018); 42 U.S.C. § 5106(e) (2018) ("The Secretary shall prescribe regulations and make such arrangements as may be necessary or appropriate to ensure that there is effective coordination among programs related to child abuse and neglect under this subchapter and subchapter III and

2020, Congress allotted about $10.2 billion dollars of funding for state child welfare systems.[25] Because every state accepts federal funds for child welfare,[26] all 50 states, the District of Columbia, and Puerto Rico currently comply with the SSA requirement by providing these relatively small data sets to the federal government, resulting in the three main sources of standardized child welfare data accessible to public health researchers: the National Child Abuse and Neglect Data System (NCANDS),[27] Adoption and Foster Care Analysis and Reporting System (AFCARS),[28] and the National Youth in Transition Database (NYTD).[29]

NCANDS. The NCANDS dataset, which states are required to submit in connection with receiving federal Child Abuse and Treatment Act (CAPTA) grants,[30] is focused on reports of child abuse and neglect—in other words, the point when the state child welfare system might first become involved with a family. The major product of NCANDS is the annual federal Child Maltreatment report available on the Children's Bureau website, which is "the United States' primary source of information about maltreated children who were known to child protective services agencies."[31] NCANDS data is used by the federal government, states, nonprofits, and advocacy organizations as a measure to assess states' performance on national child welfare outcomes.[32] While aspects of this dataset describe individual, case-level, or personally identifiable information (such as date of birth, military status, and whether the state provided family preservation services in connection with a Child Protective

---

other such programs which are assisted by Federal funds."); Foster Care and Adoption Data Collection, 45 C.F.R. § 1355.40 (2019).

25. EMILIE STOLTZFUS, CONG. RESEARCH SERV., IF10590, CHILD WELFARE: PURPOSES, FEDERAL PROGRAMS, AND FUNDING (last updated Feb. 3, 2020) (noting Congress allotted state child welfare programs approximately $10.2 billion for 2020 via Div. A of P.L. 116-94, and as part of the Family First Transition Act (FFTA, Sec. 602, Div. N of P.L. 116-94), which provides one-time funds and temporary policy changes to help implement the Family First Prevention Services Act (FFPSA, Title VII, Div. E of P.L. 115-123)).

26. *See Title IV-E Foster Care*, CHILD. BUREAU, U.S. DEP'T OF HEALTH & HUM. SERVS. (lasted updated June 25, 2020), https://www.acf.hhs.gov/cb/resource/title-ive-foster-care (describing compliance reviews in the 50 states, the District of Columbia, and Puerto Rico).

27. *NCANDS Reporting System*, CHILD. BUREAU, U.S. DEP'T OF HEALTH & HUM. SERVS. (last updated June 12, 2019), https://www.acf.hhs.gov/cb/research-data-technology/reporting-systems/ncands.

28. *Foster Care Datasets*, NAT'L DATA ARCHIVE ON CHILD ABUSE & NEGLECT, https://www.ndacan.cornell.edu/datasets/datasets-list-afcars-foster-care.cfm (last visited Oct. 21, 2020).

29. *National Youth in Transition Datasets*, NATIONAL DATA ARCHIVE ON CHILD ABUSE & NEGLECT, https://www.ndacan.cornell.edu/datasets/datasets-list-nytd.cfm (last visited Sept. 12, 2020).

30. 42 U.S.C. § 5104(c)(1)(C)-(D) (2018); *see generally* 1988 Act to amend the Child Abuse and Treatment Act, Pub. L. 100–294, 102 Stat. 102 (1988) (codified as amended at 42 U.S.C. §§ 5101-5119(c)) (requiring creation of national data collection and analysis program to make available state child abuse and neglect reporting information).

31. U.S. DEP'T OF HEALTH & HUMAN SERVS., PRIVACY IMPACT ASSESSMENT*:* NAT'L CHILD ABUSES & NEGLECT DATA SYSTEM 4, (Oct. 17, 2016), https://www.hhs.gov/sites/default/files/acf-nationalchildabuseandneglectdatasystem.pdf (noting also that NCANDS data were incorporated into the Child and Family Services Reviews to ensure conformity with State Plan requirements in titles IV–B, and IV–E of the SSA) [hereinafter PRIVACY IMPACT ASSESSMENT].

32. *See, e.g.*, *Child Maltreatment*, CHILD TRENDS (May 7, 2019), https://www.childtrends.org/indicators/child-maltreatment; *State-Level Data for Understanding Child Welfare in the United States*, CHILD TRENDS (Feb. 26, 2019), https://www.childtrends.org/publications/state-level-data-for-understanding-child-welfare-in-the-united-states (using NCANDS data to assess child welfare outcomes).

Services report), states do not submit un-encrypted individual or case identifier numbers to the federal government.[33]

AFCARS. States submit the AFCARS data sets to the federal government every six months in connection with the funding they receive under Title IV-E of the SSA.[34] Under AFCARS, states "collect case-level information on all children in foster care for whom the title IV-E child welfare agency has responsibility for placement, care, or supervision and those who have been adopted with [T]itle IV-E agency involvement."[35] In other words, this data set is focused on what happens to youth and families once a state has opened a child welfare case.[36] It includes 103 discrete pieces of information about a child and the family, such as reasons for coming into care and whether the child/family receives Social Security Act payments and Medicaid.[37]

NYTD. This data set, also submitted by states in connection with SSA funding, is an attempt to understand outcomes when youth exit the foster care system,[38] although it could include data on non-foster youth who receive independent living services that are paid for by the state child welfare agency.[39]

For each of these datasets, the states submit data to the federal Administration for Children and Families (ACF), a division of Health and Human Services, via an internet portal established for secure transmission of state data.[40] A state can only navigate to its own state-specific site; and access to other state sites is blocked.[41]

The raw data submitted to ACF still contains some sensitive information that could potentially be used to re-identify the individuals who have had contact with the state child welfare system (such as county information, which combined with other data points could reveal individuals' identities in small counties).[42] Accordingly, the federal government contracts with Cornell University to host the National Data Archive on Child Abuse and Neglect (NDACAN), which manages the process of releasing a version of these datasets to researchers upon

33. *See* CHILD.'S BUREAU OF THE U.S. DEP'T OF HEALTH & HUM. SERVS. NATIONAL CHILD ABUSE AND NEGLECT DATA SYSTEM (NCANDS) CHILD FILE CODEBOOK, (Sept. 2015), https://www.acf.hhs.gov/sites/default/files/cb/ncands_child_file_codebook.pdf (fields # 4, 59); PRIVACY IMPACT ASSESSMENT, *supra* note 31.

34. U.S. DEP'T OF HEALTH & HUM. SERVICES, ADOPTION AND FOSTER CARE ANALYSIS AND REPORTING SYSTEM (AFCARS) GUIDE TO AN ASSESSMENT REVIEW, 2 (Dec. 2012), https://www.acf.hhs.gov/sites/default/files/cb/afcars_assessment_review_guide.pdf.

35. *Id.*

36. *See id.* (showing what happens to youth and families once a state opens child welfare cases).

37. NAT'L DATA ARCHIVE ON CHILD ABUSE AND NEGLECT, AFCARS FOSTER CARE ANNUAL FILE: CODE BOOK, (revised Mar. 25, 2019), https://www.ndacan.acf.hhs.gov/datasets/pdfs_user_guides/AFCARSFosterCareCodebook.pdf.

38. NAT'L YOUTH IN TRANSITION DATABASE, GUIDE TO THE NYTD REVIEW, 13 (Dec. 12, 2017) https://www.acf.hhs.gov/sites/default/files/cb/nytd_review_guide.pdf.

39. *See* 45 C.F.R. §§ 1356.80-86 (2019) (highlighting reporting requirements); *see also* NAT'L YOUTH IN TRANSITION DATABASE, GUIDE TO THE NYTD REVIEW, 5 (Dec. 12, 2017) https://www.acf.hhs.gov/sites/default/files/cb/nytd_review_guide.pdf (showing include data on non-foster youth who receive independent living services that are paid for by the State child welfare agency).

40. PRIVACY IMPACT ASSESSMENT, *supra* note 31.

41. *Id.*

42. NAT'L DATA ARCHIVE ON CHILD ABUSE & NEGLECT, *supra* note 37; NATIONAL CHILD ABUSE & NEGLECT DATA SYSTEM, *supra* note 33.

formal request.[43] Researchers receive slightly different datasets from Cornell than the ones uploaded by states to the federal government, with the pieces of raw data that could potentially be used to re-identify individuals removed.[44] There is significant lag of at least six months—and for certain datasets, years—between when the states submit the data to the federal government and when Cornell makes it publicly available to U.S. researchers.[45]

These datasets are the main window that researchers have into the heavily guarded public systems that manage child welfare cases for the nearly 450,000 children in foster care throughout the country.[46] Moreover, the discrete data points provided in these data sets are not particularly robust.[47] Many of the federal codebook requirements rely on individual social worker discretion and interpretation in analyzing aspects of a case and entering the data into the correct field in the SACWIS system, such as determining a child's race.[48] The reduced quality of the underlying data entry results in ambiguous datasets that can hinder researchers' ability to draw reliable conclusions from the data. One stark example is the total lack of data around the number of Indian Child Welfare Act cases and state compliance.[49] As tribes noted when commenting on proposed, and long-delayed regulations that would expand the number of data points about tribal-affiliated children, "ICWA has been law for 40 years but there has been little in-depth data . . . regarding this law."[50] The codebooks' limited categories result in significant undercounts for mixed race and Alaska Native/American Indian children in particular, which in turn affects the funding available for

43. *All Datasets*, NAT'L DATA ARCHIVE ON CHILD ABUSE & NEGLECT (last accessed Aug. 25, 2019), https://www.ndacan.cornell.edu/datasets/datasets-list.cfm.

44. *See, e.g.*, NAT'L CHILD ABUSE & NEGLECT DATA SYSTEM, *supra* note 33 (NCANDS maintains a separate NCANDS codebook for the restricted use NCANDS files it distributes. Field names and properties in the NCANDS codebook may differ from those in [the state] codebook due to modifications NCANDS makes to their distributed file.").

45. *Foster Care Datasets*, *supra* note 28 (showing most recent available AFCARS data is from 2017, NYTD data is from 2018, and NCANDS data is from 2017); *see also Frequently Asked Questions (FAQ)*, NAT'L DATA ARCHIVE ON CHILD ABUSE & NEGLECT, https://www.ndacan.acf.hhs.gov/ faq.cfm#q12 ("access to some datasets in our holdings requires affiliation with an institution that has an Institutional Review Board (IRB). Requests for data from users outside of the U.S. and territories are handled on a case-by-case basis") (last visited Oct. 21, 2020).

46. U.S. DEP'T OF HEALTH & HUM. SERVS., THE AFCARS REPORT: PRELIMINARY FY 2017 ESTIMATES AS OF AUGUST 10, 2018 - NO. 25, https://www.acf.hhs.gov/sites/default/files/cb/afcarsreport25.pdf. *See also* Michael Corrigan, *Building CCWIS is as Easy as 1,2,3*, CHRONICLE. OF SOC. CHANGE (Jan. 15, 2019), https://chronicleofsocialchange.org/child-welfare-2/building-comprehensive-child-welfare-information-system ("[w]ithout quality data, there is a 50/50 chance your data is not providing a reliable and valid approach to consistently measuring what you think you are measuring. Moreover, you could be missing critical data pieces related to a youth's previous challenges and experiences, which are essential to reducing the duration and breadth of services needed.").

47. NAT'L DATA ARCHIVE ON CHILD ABUSE & NEGLECT, *supra* note 37; NAT'L CHILD ABUSE & NEGLECT DATA SYSTEM, *supra* note 33.

48. *See e.g.*, NAT'L CHILD ABUSE & NEGLECT DATA SYSTEM, *supra* note 33 (including in Field numbers 16–21 data points requesting social worker to opine on a child's race and ethnicity by choosing from a limited menu of five vague options: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, and Unable to Determine).

49. *Id*; 84 Fed. Reg. 16,572 (Apr. 19, 2019) (to be codified at 45 C.F.R. § 1355) (discussing complete lack of data points about number of ICWA cases and details surrounding such cases despite comprehensive state compliance requirements).

50. *Id.* 84 Fed. Reg. 16,572.

culturally-relevant services that could particularly benefit these youth and families.[51]

Furthermore, there is surprisingly no federal requirement for state and local courts that handle child welfare cases to provide data to the federal government—another huge potential source of timely information that could be informing government policies and procedures.[52]

Thus, the minimal data points collected by the federal government from states fail to capture critical information that gives a clear and updated picture of a child's well-being.[53] The outcomes of child welfare cases may affect generations of Americans, resulting in an enormous human cost of inherited trauma, cycles of abuse, and lost potential for those who come in contact with the child welfare system.[54] Children who have experienced abuse and neglect are at increased risk for a number of lifelong developmental, health, and mental health outcomes, including learning problems, problems relating to peers and developing relationships, depression and anxiety, aggression, and post-traumatic stress disorder.[55] These children continue to show increased risk for psychiatric disorders, substance use, serious medical illnesses, and lower economic productivity as adults.[56] Frustratingly, the majority of the incredibly rich datasets gathered by states in the course of a child welfare case—which could be analyzed with modern technology to improve lives—remain inaccessible to

---

51. CHILD WELFARE INFO. GATEWAY, RACIAL DISPROPORTIONALITY AND DISPARITY IN CHILD WELFARE 10 (2016), https://www.childwelfare.gov/pubpdfs/racial_disproportionality.pdf [hereinafter RACIAL DISPROPORTIONALITY]; U.S. GOV'T ACCOUNTABILITY OFF., GAO-18-591, NATIVE AMERICAN YOUTH: INVOLVEMENT IN JUSTICE SYSTEMS & INFORMATION ON GRANTS TO HELP ADDRESS JUVENILE DELINQUENCY (2018) https://www.gao.gov/assets/700/694430.pdf (noting undercounting of tribal youth is an issue in juvenile justice system because while some states may inquire about tribal affiliation when youth come into contact with the state's justice system, they lack reliable a process to identify Native American youth, and such youth are often unlikely to share their ethnicity with state officials, or anyone outside of their community).

52. *See, e.g.*, Susan Robison, *Delivering On The Promise Promoting Court Capacity To Improve Outcomes For Abused And Neglected Children*, NAT'L CONFERENCE OF STATE LEGISLATURES, (2007), http://www.ncsl.org/print/cyf/deliveringpromise.pdf ("Performance of the child welfare system cannot be adequately assessed without measuring the performance of courts. Courts make decisions that directly affect the safety, permanency and well-being of children and should be held accountable for the timeliness and outcomes of those decisions. Monitoring performance also is important in planning for improvement, efficient court management, and effective budgeting and use of resources. . . . [T]hree national organizations developed a set of court performance measures that address child safety and permanency, due process for all parties, and timeliness of decision making. The Pew Commission on Children in Foster Care recommended that these 'Packard performance measures' be widely adopted. To do so, courts must have ongoing capacity to collect and analyze data. Courts can begin this process without a massive investment in sophisticated information systems."); *see* Duren Banks et al., *Fostering Innovation in the U.S. Court System: Identifying High-Priority Technology and Other Needs for Improving Court Operations and Outcomes*, RAND CORP. 5–17 (2016) (describing an overview of state court obstacles to technology adoption).

53. Melissa Jonson-Reid & Brett Drake, *Child Well-Being: Where Is It in Our Data Systems?*, 10 J. PUB. CHILD WELFARE 457, 457–65 (2016).

54. *See generally* CHILD WELFARE INFO. GATEWAY, DEVELOPING A TRAUMA-INFORMED CHILD WELFARE SYSTEM (2015) https://www.childwelfare.gov/pubPDFs/trauma_informed.pdf (showing children who have experienced abuse and neglect are at increased risk for a number of lifelong developmental, health, and mental health outcomes).

55. *See, e.g., New Directions in Child Abuse and Neglect Research*, NAT'L ACAD. OF SCIENCES (2014), https://www.ncbi.nlm.nih.gov/books/NBK195987/ (describing the different consequences of abuse and neglect in different fields).

56. *Id.*

researchers who could otherwise provide urgent policy insights for state and federal leaders.[57]

Notably, while it is valuable for the federal government to encourage states to submit uniform data to compare child and family outcomes across the country, data uniformity should not be the most important goal of reform in this area. If the federal government greatly expanded the existing state data collection requirements to necessitate enormous, uniform datasets about children and families, that could potentially raise constitutional concerns. Namely, that the federal government is commandeering state child welfare programs and interfering with states' traditional police powers.[58] Instead, the federal government should support state child welfare systems by giving them a simplified legal framework and appropriate technology tools that allow each state to make the best use of its own, complex data. And as a practical matter, if there were federal support for development of a model tech platform as discussed in Part IV below, states would have a much easier time providing a larger scope of uniform data on a voluntary basis for comparative analysis because they would all be using a similar platform for data entry.

The status quo of federal child welfare data collection is not working. In order for policymakers to address this critical issue, they must understand how overlapping and restrictive federal laws, combined by outdated technology infrastructure, currently inhibit states from partnering directly with researchers in the public and private sector to better utilize public agency data.

II. Multiple Federal Statutes Complicate the Ability of State Child Welfare Agencies to Share Data for Research Purposes.

Two key federal laws, Title IV-E of the Social Security Act (SSA) and CAPTA, as well as other federal statutes specific to certain types of data contained in child welfare files, restrict the use and disclosure of child welfare system data.[59] These overlapping laws and regulations, which have been largely ignored despite other federal efforts to support technological improvements in child welfare operations,[60] should be updated to better support the necessary data-sharing that supports innovation in child welfare systems. Specifically, each of the federal laws discussed below governing child welfare system data should explicitly include a parallel requirement requiring sharing of de-identified child welfare system data for research purposes, and significant federal funds should be provided to support that data-sharing.

57. Melissa Jonson-Reid & Brett Drake, *Child Well-Being: Where Is It in Our Data Systems?*, 10 J. PUB. CHILD WELFARE 457, 457–65 (2016).

58. U.S. Const. amend. X; *see, e.g.*, *Murphy v. Nat'l Collegiate Athletic Ass'n,* 138 S. Ct. 1461, 1476 (2018) ("[C]onspicuously absent from the list of powers given to Congress is the power to issue direct orders to the governments of the States.").

59. Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (1996); Education Amendments of 1974, Pub. L. No. 93-380, 88 Stat. 484, 571–74 (1974); Individuals with Disabilities Education Act, Pub. L. No. 91-230, 84 Stat. 175 (1970).

60. *Id*.

A. The Social Security Act Requires States to Explain in their Title IV-E State Plan that they Appropriately Restrict Sharing of Child Welfare Information per Federal Confidentiality Requirements.

Title IV-E of the SSA is a federal grant program that, among other things, assists states in providing foster care and adoption/guardianship funding assistance for eligible children (generally, children removed from low-income homes).[61] Title IV-E is the largest federal source of funding for state child welfare programs—approximately $8.6 billion in FY2018—with the Act's Title IV-B also contributing another $709M in 2018 grants for child welfare services.[62] To be eligible to receive SSA funds, states must submit a state plan for the Secretary of HHS' approval explaining how the state will meet the SSA's statutory requirements.[63] One requirements is that states must explain how they will "provid[e] safeguards which restrict the use of or disclosure of information concerning individuals assisted" under the state plan.[64]

The SSA allows a state to share information about individuals who receive assistance from state foster care or adoption systems only if the sharing satisfies certain listed exceptions, including "for purposes directly connected with" (excerpted in relevant part):

> (1) the administration of a state plan under IV-B or E of the Social Security Act (which includes the foster care and adoption assistance plan and child welfare services plan);
>
> (2) "the administration of any other Federal or federally assisted program which provides assistance, in cash or in kind, or services, directly to individuals on the basis of need;" or
>
> (3) "any audit or similar activity conducted in connection with the administration of any such plan or program by any governmental agency which is authorized by law to conduct such audit or activity[.]"[65]

The federal regulations elaborate to clarify that the term "administration" of a state plan (including a IV-E plan) includes "establishing eligibility, determining the amount of assistance, and providing services for applicants and recipients."[66]

Thus, a data scientist must be able to argue to a state child welfare agency's leadership (and their counsel) that it meets one of the narrow exceptions to

61. *See* 42 U.S.C. § 670 (providing that states supply assistance to certain cases of adoption, foster care, or guardianship); *see also* 45 C.F.R. §§ 1355.10 to 1357.50 (2019) (delineating the different regulations).

62. Emilie Stoltzfus, CONG. RSCH. SERV., R45270 CHILD WELFARE FUNDING IN FY2018 4 (2018).

63. 42 U.S.C. § 671(a)(8). Like other legislation passed pursuant to Congress' Spending Clause power under Article I, §8 of the Constitution, states may choose to comply with the Act's requirements or forgo the federal funding provided by the Act.

64. *Id*.

65. *See id.* (listing additional acceptable disclosures of information not relevant to this article). The statute also notes that states have option to restrict disclosures even more than specified.

66. 45 C.F.R. § 205.50 (2019); *see* 45 C.F.R. § 1355.30 (2019) (stating 45 C.F.R. § 205.50 (2019) also applies to IV-E and IV-B plans); *see also* 45 C.F.R. § 1355.21 (2019) (reiterating that Title IV-E and B plans "must provide for safeguards on the use and disclosure of information which meet the requirements" in 42 U.S.C. § 671(a)(8)).

receive access to state child welfare data. While academic researchers, nonprofits, and private sector companies interested in analyzing child welfare data and developing technology solutions to child welfare systems problems may be able to claim that they fit into one of these categories (perhaps that the sharing would further the administration of the state plan), gaining access to child welfare data remains a matter of individually negotiating with state agencies because there is no explicit federal legal requirement for states to allow research access.

B. The Child Abuse Prevention and Treatment Act (CAPTA) Requires There to be a State Statute that Specifically Allows Information Sharing.

To further complicate matters, CAPTA is a separate restriction on an agency's sharing of child welfare information with slightly different exceptions to the federal government's strict confidentiality requirements. CAPTA, the other major federal statute responsible for state child welfare grant funding, was originally enacted in 1988 but has been amended several times, most recently in 2003, and was last reauthorized in 2010 as the CAPTA of 2011 (and a CAPTA reauthorization bill is pending as of August 2019).[67] In 2018, states received approximately $158 million in CAPTA grants.[68]

A state plan submitted in connection with a CAPTA grant must contain "an assurance in the form of a certification by the Governor of the State" that the state has in effect and is enforcing a state law or program relating to child abuse and neglect that includes methods to preserve the confidentiality of child welfare information.[69] CAPTA defines that information to include "all records in order to protect the rights of the child and of the child's parents or guardians, including . . . reports and records made and maintained pursuant to the purposes of" receiving CAPTA grants and community based grants for child abuse and neglect prevention.[70] This definition is thus broad enough to encompass most data collected by child welfare systems.[71] The statute goes on to provide that this broadly defined class of records "shall only be made available to" child welfare system partners listed in the statute.[72] That includes only one category

67. Child Abuse Prevention and Treatment Act, Pub. L. No. 93-247, 88 Stat. 4 (1974) (codified as amended by the Keeping Children and Families Safe Act of 2003, Pub. L. No. 108-36, 117 Stat. 800 (2003), which amended Title I and replaced Title II, Community-Based Family Resource and Support Program with Community-Based Grants for the Prevention of Child Abuse and Neglect). CAPTA was reauthorized in 2010. *See* CAPTA Reauthorization Act of 2010, Pub. L. No. 111-320, 124 Stat. 3459 (2010). *See also* John Sciamanna, *Reauthorization Coming in Weeks, Maybe*, CWLA, https://www.cwla.org/capta-reauthorization-coming-in-weeks-maybe-days/ (last visited Oct. 21, 2019) (discussing 2019 pending CAPTA reauthorization).

68. Stoltzfus, *supra* note 62, at 4.

69. 42 U.S.C. § 5106a(b)(2)(B)(viii).

70. 42 U.S.C. 5106a(b)(2).

71. *See id.* (encompassing most data provided by child welfare systems).

72. 42 U.S.C. 5106a(b)(2)(B)(viii) (describing that a state's required CAPTA report to receive federal funding must explain how the state has "methods to preserve the confidentiality of all records in order to protect the rights of the child and of the child's parents or guardians, including requirements ensuring that reports and records made and maintained pursuant to the purposes of this subchapter and subchapter III shall only be made available to—(I) individuals who are the subject of the report; (II) Federal, State, or local government entities, or any agent of such entities, as described in clause (ix); (III) child abuse citizen review panels; (IV) child fatality review panels; (V) a grand jury or court, upon a finding that information in the record is necessary for the

relevant to the research access discussed in this article: "other entities or classes of individuals *statutorily authorized by the State to receive such information* pursuant to a legitimate State purpose."[73] Thus, CAPTA contains a critical limit on the ability to share information that even the SSA would allow states to share if it were to further the administration of state child welfare services in the Title IV-E plan.

Although several states have independently chosen to enact statutes allowing sharing for "bona fide research" purposes, this is not a requirement to receive federal funds.[74] And even in those states that do allow sharing for research purposes, the statutes do not mandate sharing (the decision would be at the discretion of an individual state agency), and there is no clear or uniform process around what data the researchers can access.[75] This leaves individual researchers to figure out on their own how to negotiate access to child welfare system data.

Leaving the decision about critical research access up to individual states—when the child welfare system is largely supported by federal funds in every state—is a mistake.[76] In its description of the purpose of CAPTA grants, the statute describes several goals that would suggest a need for less restrictive confidentiality requirements, including "supporting and enhancing interagency collaboration among public health agencies, agencies in the child protective service system, and agencies carrying out private community-based programs," and "developing, facilitating the use of, and implementing research-based strategies and training protocols for individuals mandated to report child abuse and neglect."[77] CAPTA also purports to help agencies improve agency operations and legal representation.[78] Yet, Congress failed to include in CAPTA any explicit requirement regarding data-sharing for research purposes.

As a result, individual states decide whether to adopt statutes that allow the public to conduct rigorous studies of child welfare data. States with poor outcomes may prefer to obscure the raw data that could result in political accountability. The result is a patchwork of data-sharing laws that leaves the work of these powerful public agencies mostly shielded from the public's view, and frustrates states' abilities to understand and optimize their own systems in the context of national data.

---

determination of an issue before the court or grand jury; and (VI) other entities or classes of individuals statutorily authorized by the State to receive such information pursuant to a legitimate State purpose. . . (ix) provisions to require a State to disclose confidential information to any Federal, State, or local government entity, or any agent of such entity, that has a need for such information in order to carry out its responsibilities under law to protect children from child abuse and neglect.").

73. 42 U.S.C. § 5106a(b)(2)(B)(viii)(VI).

74. *See* CHILD WELFARE INFORMATION GATEWAY, DISCLOSURE OF CONFIDENTIAL CHILD ABUSE AND NEGLECT RECORDS (2017), https://www.childwelfare.gov/pubPDFs/confide.pdf. (stating that The Child Welfare Information Gateway's list of such provisions, which is current as of 2017, is a starting point for determining what states allow sharing for "bona fide research" purpose.) [hereinafter DISCLOSURE].

75. *See id.* (demonstrating the lack of uniformity in data sharing between state agencies).

76. *See* DISCLOSURE, *supra* note 74, at 2 (discussing that every state receives federal funds).

77. 42 U.S.C. § 5106a.

78. *Id.*

C. The Recently Adopted Comprehensive Child Welfare Information Systems (CCWIS) Regulations Refer Back to SSA and CAPTA Confidentiality Requirements.

The CCWIS regulations, a much-needed modernization of state data system requirements adopted by the federal Administration for Children and Families in June 2016, expressly encourage greater sharing of child welfare information, yet they do not offer a revised confidentiality and data-sharing framework.[79] For states that have a CCWIS system, the regulations provide that the child welfare agency must use the child welfare system data described in CCWIS requirements to "[g]enerate, or contribute to, reports needed by state or tribal child welfare laws, regulations, policies, practices, reporting requirements, audits, and reviews that support programs and services described in [T]itle IV-B and [T]itle IV-E."[80] The broad language in this section would seem to support broader information sharing. However, the "data quality requirements" section of the CCWIS regulations simply require the agency to adhere to the confidentiality requirements from the SSA and CAPTA and their associated regulations.[81]

Agencies had 24 months from the date of the final rule's publication in 2016 to decide if an existing system will be transitioned to a CCWIS system.[82] While several states did indicate that they would be moving in this direction, few states have committed the necessary capital required to do so; the state of Arizona is one major exception as discussed in greater detail below.[83]

D. Other Federal Statutes Restrict Sharing of Certain Types of Child Welfare System Data.

While SSA and CAPTA remain the key barriers to state child welfare data-sharing, depending on the nature of specific data collected by social workers and stored in state child welfare systems, states may also need to undertake separate privacy analyses under other federal laws.

One key example is the Health Insurance Portability and Accountability Act of 1996 (HIPAA), which has its own privacy rule describing how protected health information can be used and disclosed by entities covered by HIPAA.[84] Many states store healthcare-related information in child welfare systems, but not all state child welfare agencies consider themselves to meet the definition of

---

79. 45 C.F.R. §§ 1355-56 (2018); *see CCWIS Final Rule*, *supra* note 17. (stating that CCWIS requirements focus on quality data and program outcomes, including that state agencies develop and implement a data quality plan, support new data exchanges, and implement a data exchange standard).

80. 45 C.F.R. § 1355.52(c) (2019); *see also id.* at § 1355.52(e) (discussing information-sharing with external systems).

81. 45 C.F.R. § 1355.52(d)(iii) (2019).

82. *CCWIS Final Rule*, *supra* note 17, at 5.

83. *See infra* Section IV.

84. Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (1996); *see* 45 C.F.R. § 160.103 (2019) (defining covered entity and other relevant terms); *see also* U.S. DEP'T OF HEALTH & HUM. SERVS., *Summary of the HIPAA Privacy Rule, OFF. FOR C.R.* (2013) https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html (describing how private information can be shared between entities covered by HIPAA) [hereinafter *HIPAA Privacy Summary*].

a "covered entity" under HIPAA such that every data-sharing agreement would require a HIPAA privacy analysis.[85] But notably, even child welfare agencies that deem themselves to be covered entities can enter into a business associate agreement to share data, including protected health information[86]—assuming the agency decides it has the authority to do so under the other federal and state laws discussed above. In addition, HIPAA does not place restrictions on the use or disclosure of de–identified information (which neither identifies nor provides a reasonable basis to identify someone).[87]

Child welfare agencies may also have to analyze federal education statutes because child welfare databases typically include some amount of data about every foster child's educational milestones. Notably, the Family Educational Rights and Privacy Act (FERPA) and the Individuals with Disabilities Education Act (IDEA)—which includes separate privacy restrictions—explicitly allow for information-sharing for research purposes.[88]

In sum, while child welfare agencies must think through whether and how these other federal statutes restrict research access to child welfare data, the SSA and CAPTA remain the main impediments to data-sharing for systems innovation.

III. Unlocking the Potential of AI in Child Welfare

Part III first highlights how certain states have recognized the benefits of harnessing state administrative agency data in real time and allowing data scientists to apply emerging technologies to this data with the goal of improving government policies. Then, to get policymakers excited about the untapped potential for systems improvement, Part III goes on to suggest potential child welfare applications for AI tools based on existing applications currently in use in the private sector. Part III then goes on to discuss one current use of predictive analytics that should be rethought as an initial focus of technology

85. *See* Alicia K. Davis, *Understanding HIPAA to Overcome Challenges in Child and Family Cases*, Trends in State Courts (2012) https://cdm16501.contentdm.oclc.org/digital/collection/accessfair/id/242 (noting that, "[p]ublic-health and social-service agencies have issued statements on whether or not they deem themselves to be covered entities.").

86. *See* Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (1996) (defining covered entity and other relevant terms); 45 C.F.R. § 160.103 (2019); *see also HIPAA Privacy Summary*, *supra* note 84.

87. *HIPAA Privacy Summary*, *supra* note 84, (citing 45 C.F.R. §§ 164.502(d)(2), 164.514(a) and (b). ("There are two ways to de-identify information; either: (1) a formal determination by a qualified statistician; or (2) the removal of specified identifiers of the individual and of the individual's relatives, household members, and employers is required, and is adequate only if the covered entity has no actual knowledge that the remaining information could be used to identify the individual.").

88. Education Amendments of 1974, Pub. L. No. 93-380, 88 Stat. 484, 571–74 (1974) (codified, in part, at 20 U.S.C. § 1232g); 20 U.S.C. § 1232g(b)(1)(F) (2018) (allowing information sharing for "organizations conducting studies for, or on behalf of, educational agencies or institutions for the purpose of developing, validating, or administering predictive tests, administering student aid programs, and improving instruction, if such studies are conducted in such a manner as will not permit the personal identification of students and their parents by persons other than representatives of such organizations and such information will be destroyed when no longer needed for the purpose for which it is conducted"); 34 C.F.R. § 99.31 (2019); Individuals with Disabilities Education Act, Pub. L. No. 91-230, 84 Stat. 175 (1970) (codified, in part, at 20 U.S.C. § 1400); *see also* 34 C.F.R. §§ 300.622, 303.414 (referring back to FERPA's 34 C.F.R. § 99 research exception to confidentiality rules).

transformation efforts. Importantly, the government, nonprofit, and private sector actors involved in any of these examples of potential data science applications for child welfare must rigorously analyze privacy implications before implementation (as discussed in more detail in Section V).

A. Examples of State Efforts to Apply Emerging Technology Tools to Government Social Services.

Two recent state developments focused on harnessing state data to generate insights in government social services demonstrate the possibilities if federal law supports this type of innovation. One fascinating example is the Rhode Island Innovative Policy Lab, now called Research Improving People's Lives (RIPL), which is a partnership between academic researchers and the State of Rhode Island.[89] Because Rhode Island's state laws allow data-sharing for bona fide research purposes,[90] RIPL was able to enter into a data-sharing agreement with the state to construct an anonymized database called RI 360 from the administrative records of nearly every Rhode Island state agency.[91]

> "The comprehensive scope of these data enable research projects across a wide range of policy challenges, such as lowering non-urgent emergency health care costs, curbing the opioid epidemic, improving worker training programs, creating tools to connect dislocated workers to benefits, helping families become more food secure, optimizing energy policy for low income families, helping children reach proficiency on reading and math tests, and closing the college achievement gap."[92]

RIPL recently published the results of one research project to emerge from this data-sharing arrangement, which used machine learning methods to examine whether the risk of future opioid dependence, abuse, or poisoning can be predicted before the first opioid prescription.[93] RIPL's machine learning models were able to accurately predict these outcomes and identify certain factors—some surprising and non–intuitive—that apparently serve as predictors for future opioid addiction, including prior non–opioid prescriptions, medical history, incarceration, and certain demographic variables.[94]

As another example, Los Angeles has recently focused on using technology to reduce the 2 million hours of social worker time spent on scheduling 4 million

---

89. Justine S. Hastings et al., *Integrating Administrative Data for Policy Insights*, R. I. INNOVATIVE POL'Y LAB, https://www.ripl.org/wp-content/uploads/2018/08/RIIPL_data_infrastructure_20180613_preprint.pdf (last visited Oct. 21, 2020).

90. *See* R.I. DEPT. OF CHILD., YOUTH AND FAMS., Confidentiality: Access to Restricted Information 100.0010 (A)(2)(d) (allowing bona fide research access by permission of Child Welfare Director or his/her designee) (1986), http://www.dcyf.ri.gov/policyregs/confidentiality__access_to_restricted_information.htm?ms=AAAAAAAAAAA%3D&q=cmVzZWFyY2g%3D&st=Mg%3D%3D&sct=MA%3D%3D&mw=MzYx.

91. Hastings et al., *supra* note 89, at 2.

92. *Id. at* 2–3.

93. Justine S. Hastings et al., *Predicting High-Risk Opioid Prescriptions Before They are Given* (Nat'l Bureau of Econ. Research, Working Paper No. 25791, 2019) https://www.ripl.org/initiatives/initiative-two-social-program-innovation-2/initiative-two-predicting-high-risk-opioid-prescriptions-nber.

94. *Id.*

hours of parent visitation, including all the hours the parties spend driving around the city to inconvenient visitation locations.[95] Using algorithms to analyze the map of travel patterns for visitation led the state agency to design an improved visitation system, as well as partnerships with private sector ride-sharing services, that minimize one frustrating aspect of child welfare operations.[96]

Another innovative state effort is underway in Georgia. In March 2019, Georgia passed a bill establishing a statewide Strategic Integrated Data System known as the Georgia Data Analytic Center (GDAC).[97] The GDAC Project is designed so that a central data repository overseen by the state Office of Planning and Budget will receive and maintain individually identifiable government administrative data, but will only transmit out de–identified data under the specific circumstances permitted by the new law and in a format that is secured to prevent disclosure of personally identifiable information.[98] This strategic and centralized coordination of state data is exactly the type of role that the federal government could and should be playing for all states to enable a straightforward path to large administrative data sets.

B. Promising Potential Use Cases for AI Applications in Child Welfare Systems

Because AI is able to detect patterns that human cognition cannot, AI could be used by social workers to inform the numerous decision-making points that arise in a child welfare case. The data could also be used at the policy and leadership level to understand broad trends in effectiveness of government services and operations. The following are several specific examples of practices that could be implemented using AI technology that currently exist and are in use in other sectors:

Diagnose key barriers to parental reunification and suggest the most effective prevention services for a given diagnosis by reviewing massive datasets. AI holds incredible promise for evaluating the effectiveness of child welfare programs and interventions because of its ability to employ advanced statistical analyses that can span multiple source datasets.[99] This capability could be particularly useful given the recently enacted Family First Prevention Services Act (FFPSA), which now allows states to receive some federal

95. Jeremy Loudenback, *Los Angeles Hopes Tech Investment Can Improve Visitation for Foster Children*, THE IMPRINT (Apr. 25, 2018) https://imprintnews.org/child-welfare-2/los-angeles-countys-dcfs-is-hoping-that-better-technology-will-improve-the-family-visitation-process/30605.

96. *Id.*

97. Gen. Assemb. H.B. 197, Reg. Sess. (Ga. 2019), http://www.legis.ga.gov/Legislation/20192020/183268.pdf; *2019-2020 Regular Session – HB 197*, GA. GEN. ASSEMBLY LEGIS. http://www.legis.ga.gov/legislation/en-US/Display/20192020/HB/197 (last visited Oct. 21, 2020).

98. Gen. Assemb. H.B. 197, Reg. Sess. (Ga. 2019) http://www.legis.ga.gov/Legislation/20192020/183268.pdf.

99. *See, e.g.*, Jennifer Bresnick, *Top 12 Ways Artificial Intelligence Will Impact Healthcare*, HEALTH IT ANALYTICS (Apr. 30, 2018) https://healthitanalytics.com/news/top-12-ways-artificial-intelligence-will-impact-healthcare (discussing AI applications in healthcare such as analyzing electronic health record data to help identify infection patterns to prevent evolution of superbugs).

reimbursement for state monies spent on prevention services.[100] As states begin to implement FFPSA, there is greater demand to know what services are most effective in helping parents reunify with their children to utilize them earlier.[101] States are seeking to collect data to meet FFPSA's requirement that programs be "evidence-based" to qualify such services for federal reimbursement.[102] This data in turn could be utilized on a broader scale to inform national policies.

AI algorithms can detect patterns in large datasets that human researchers may otherwise miss. Just as health care systems are beginning to understand the implications of employing AI to diagnose and understand health outcomes and uncover opportunities for early intervention, child welfare systems could be doing the same. Eric Topol explores these issues in detail in his book *Deep Medicine*, which opens with a discussion of how misdiagnosis is disconcertingly common, with about 12 million significant misdiagnoses a year.[103] Topol discusses how in most cases, prescribed drugs don't work because physicians "have not honed an ability to predict what sort of person will respond to a treatment or acquired adequate knowledge about an individual to know whether the patient is among those people who will respond positively to a treatment."[104] In addition, bias often clouds an individual doctor's treatment decisions as a result of the fact that doctors deal with patients one at a time, where their personal and recent experience treating other patients often overrides hard data derived from much larger samples of people.[105] Caseworkers face parallel limitations in assessing the parents and children assigned to their case load, who come to the social worker's attention after the families are already in crisis.

AI applications in medicine are helping to overcome diagnostic challenges,[106] and could similarly be applied to social worker "diagnoses" of parental challenges that inform parenting plan recommendations. One possibility is to utilize "collaborative filtering" technologies—in other words, algorithms that leverage what other social workers indicated they liked and found to work well for their clients.[107]

In addition, Topol writes about technology startups that are beginning to create AI diagnostic tools that go beyond lists of symptoms by asking back and forth questions, aiding with getting a second opinion via telemedicine and crowdsourcing, and suggesting treatments thanks to machine learning capabilities allowing apps to devour the 160,000-plus cancer research papers

100. Bipartisan Budget Act of 2018, Pub. L. No. 115-123, 132 Stat. 64, 170 (2018) (extending the Family First Prevention Services Act).

101. *See Program Instruction*, U.S. DEP'T. OF HEALTH & HUM. SERVS., FAMILY FIRST PREVENTION SERVICES ACT (Nov. 30, 2018) https://www.cwla.org/wp-content/uploads/2018/12/ACYF-CB-PI-18-09-State-FFPSA-Prevention-PI.pdf.

102. *Id.*

103. ERIC TOPOL, DEEP MEDICINE 26 (2019).

104. *Id.* at 37.

105. *Id.* at 46.

106. Prince Glover, *Various Implementations of Collaborative Filtering*, TOWARDS DATA SCI. (Dec. 28, 2017), https://towardsdatascience.com/various-implementations-of-collaborative-filtering-100385c6dfe0.

107. *Id.*

published per year.[108] Relatedly, AI tools excel at anomaly detection—in other words, finding outliers in data that would otherwise go unseen,[109] which could be applied to understand why certain families succeed in reunification where others fail. Eventually, the ability to analyze broad data sets will give rise to insights about prevention that can help states access federal FFPSA funds and reduce the number of children coming into foster care.[110]

Reduce administrative burden on social workers so that they can focus on high quality social work, including in-person visits. Just as the private sector has increasingly applied data science to minimize and automate administrative tasks such as e-mail, scheduling, answering routine questions, and connecting customers with information,[111] these technologies could be applied to reduce a social worker's administrative burden. As detailed in a recent study analyzing how social workers in Connecticut's Department of Children and Families Administration spent their work hours, "[w]orkers spend much of their work time on tasks that do not involve direct contact with children, families or stakeholders. They include, among other things, travel time, data entry, court preparation, and written communications.[112]

As the joke goes, "[t]here's an app for that."[113] AI suggestions for auto-filling data entry in CCWIS systems, converting speech to text, and identifying discrete data points from narrative text (including voice entry text) would vastly decrease a social worker's time spent on data entry and improve the ability of machine learning to understand what is actually occurring in a case and detect patterns.[114] Such innovations are already in use in the private sector and could be integrated into social worker workflows if the underlying CCWIS system supported it.[115] The health care sector is similarly facing huge burdens around data entry, and AI is currently being explored as a way to reduce the challenge of data entry.[116] Much of the data captured by social workers in the course of

108. TOPOL, *supra* note 103, at 51–56.

109. Varun Chandola et al., *Anomaly Detection: A Survey*, 9 ACM COMPUTING SURVS. 1 (2009) http://cucis.ece.northwestern.edu/projects/DMS/publications/AnomalyDetection.pdf.

110. *See* TOPOL, *supra* note 103, at 117–18 (discussing innovative health care applications for AI based on large volumes of medical scan data).

111. Dave Zielinski, *AI for Administrative Tasks Can Make Work Life Easier*, SHRM (June 24, 2019), https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/ai-for-administrative-tasks-can-make-life-easier-at-work.aspx.

112. DEPARTMENT *Time Study DCF Social Work Staff During March 2016*, DEPT. OF CHILD. AND FAMIS. CT. MONITOR'S OFF. (2017), https://portal.ct.gov/-/media/DCF/PIP/timestudypdf.pdf?la=en.

113. *See* John Brownlee, *Apple Gets a Trademark: There's an App for That*™, CULT OF MAC (Oct. 11, 2010), https://www.cultofmac.com/62892/apple-gets-a-trademark-theres-an-app-for-that/ (noting that many admire the slogan's memorable reference).

114. *See, e.g.*, Bresnick, *supra* note 99 (discussing voice recognition and dictation are helping to improve the clinical documentation process).

115. *See id.; infra* Section IV (discussing technical limitations of CCWIS systems).

116. *See* Danielle Ofri, Opinion, *The Business of Health Care Depends on Exploiting Doctors and Nurses*, N.Y. TIMES (June 8, 2019)), https://www.nytimes.com/2019/06/08/opinion/sunday/hospitals-doctors-nurses-burnout.html (citing Brian G. Arndt et al., *Tethered to the EHR: Primary Care Physician Workload Assessment Using EHR Event Log Data and Time-Motion Observations*, 15 ANNALS FAM. MED. 419, 420 (2017)) ("By far the biggest culprit of the mushrooming workload is the electronic medical record, or E.M.R. . . . There are many salutary aspects of the E.M.R., and no one wants to go back to the old days of chasing down lost charts and deciphering inscrutable handwriting. But the data entry is mind-numbing and voluminous. Primary-care doctors

their jobs is unstructured and will need a natural language processing system to process it so that it could be further analyzed by AI to further policy goals.[117] AI currently excels at performing such natural language processing that can allow the data to be further utilized by child welfare leaders.[118]

AI has also the potential to reduce the administrative burden in connection with child welfare system interactions with parents at risk of having their parental rights terminated. Child welfare judges put significant emphasis on a parent's participation in the court proceedings as well as their case plan, which typically involves the parent attending therapy, parenting classes, and other types of service appointments.[119] Reminding people to appear in court improves appearance rates,[120] and social workers and defense attorneys spend a significant amount of time trying to track down parents to remind them about appointments. Algorithms applied to court system databases could be used to issue text message reminders,[121] and even deploy additional follow-up for individuals at highest risk of missing a critical court appearance. And just as Chinese universities are using AI to monitor class attendance and follow up with those who missed class to learn why,[122] child welfare systems could experiment with using similar technology to verify parent attendance and participation in services without requiring social workers (or the parents themselves) to spend large amounts of time manually tracking and verifying parents' case plan compliance. And if parents know that an electronic system, not their individual social worker, is monitoring attendance in this way, it could even reduce their negative emotions towards the social worker as the "hall monitor."[123]

In addition, robotic process automation (RPA) is a technology that holds particular promise for helping social workers automate and streamline their daily

spend nearly two hours typing into the E.M.R. for every one hour of direct patient care. Most of us are now putting in hours of additional time each day for the same number of patients.").

117. Lindsey Getz, *Big Data's Impact on Social Services*, 14 SOC. WORK TODAY 28 (2014), https://www.socialworktoday.com/archive/031714p28.shtml.

118. Michael J. Garbade, *A Simple Introduction to Natural Language Processing*, BECOMING HUM.: A.I. MAG. (Oct. 15, 2018), https://becominghuman.ai/a-simple-introduction-to-natural-language-processing-ea66a1747b32.

119. *See* CHILD WELFARE INFO. GATEWAY, CASE PLANNING FOR FAMILIES INVOLVED WITH CHILD WELFARE AGENCIES 3 (2018), https://www.childwelfare.gov/pubPDFs/caseplanning.pdf#page=3&view=Contents%20of%20a%20case%20plan ("State requirements for case plans typically include descriptions of the problems that led to the family's involvement with the State agency and the services that will be provided to the parents to address those problems, as well as any services needed by the child and the child's caregivers. The plans also include goals and objectives that the parent(s) must meet in order to ensure that they can provide a safe home for the child and timeframes for achieving those goals.").

120. JENNIFER ELEK ET AL., NAT'L CTR. FOR STATE COURTS' PRETRIAL JUSTICE CTR. FOR COURTS, USE OF COURT DATE REMINDER NOTICES TO IMPROVE COURT APPEARANCE RATES 1 (2017).

121. *See, e.g.*, *id.* at 4 (noting New York City's use of text messaging consistently reduced individuals' failure to appear); Michael Houlberg, *Text and Email Notifications Keep Court Users Up to Date*, IAALS BLOG (Mar. 6, 2020), https://iaals.du.edu/blog/text-and-email-notifications-keep-court-users-date (discussing several state bills and pending initiatives to create text message reminders for litigants).

122. Sarah Dai, *Chinese University Uses AI to Check Class Attendance Rates and Find the Reasons Behind Absenteeism*, SOUTH CHINA MORNING POST (Mar. 18, 2019 11:53 AM) https://www.scmp.com/tech/policy/article/3002107/chinese-university-uses-ai-check-class-attendance-rates-and-find.

123. *Id.*

data entry burdens.[124] RPA platforms can be can used to record common data entry and other user execution patterns, and then automate those patterns to minimize data entry and duplicative work efforts.[125]

Locate placement options, including family finding. Finding immediate family placement options is a requirement of ICWA, as well as state child welfare laws that apply to all children.[126] State child welfare systems could interface with other state systems, such as Medicaid, child support, and state tax databases and use machine learning to identify family member connections that would otherwise require extensive caseworker time and effort to identify. One pertinent example pertains to the ICWA requirement that the case be treated as an ICWA case when there is "reason to know" that a child has an affiliation with an Indian tribe.[127] An algorithm could analyze census, geographical, and other data in connection with tribal member rolls (submitted by participating tribes) to help in making this determination early on so that states can fully comply with ICWA. The private sector currently uses AI in background check processes to gather verification data from a variety of global streams and social media platforms.[128]

Improve employee engagement among child welfare staff and provide them non-punitive feedback about their effectiveness. Public child welfare agencies experience front line worker turnover rates as high as 25% a year.[129] According to a 2019 Economist Intelligence Unit report analyzing a survey of 1,145 business executives across the UK, access to information is critical in engaging and empowering workers:[130] "Having ready access to the data and insights they need to do their jobs, wherever they are located, does more to influence employee engagement and productivity, and ultimately their overall experience, than other technology factors."[131] Child welfare leaders should take note that it is not just executives that demand access to data—it is an expectation of modern employees.[132] By using AI to create data dashboards and insights that are relevant to child welfare workers in all areas of the child welfare system, state systems can improve worker satisfaction and ideally retention.

---

124. Press Release, Susan Moore, Gartner, Gartner Says Worldwide Robotic Process Automation Software Market Grew 63% in 2018 (June 24, 2019), https://www.gartner.com/en/newsroom/press-releases/2019-06-24-gartner-says-worldwide-robotic-process-automation-sof.

125. *See Robotic Process Automation (RPA) Tutorial: What is, Tools & Example*, GURU 99, https://www.guru99.com/robotic-process-automation-tutorial.html (last visited Oct 21, 2020) (discussing example RPA application to process business invoices).

126. *See* Indian Child Welfare Act of 1978, 25 U.S.C. § 1915; CHILD WELFARE INFORMATION GATEWAY, PLACEMENT OF CHILDREN WITH RELATIVES (2018) (discussing state-by-state relative placement requirements).

127. 25 C.F.R. § 23.111 (2019).

128. Chiradeep BasuMallick, *6 Ways AI and ML are Improving Employee Background Verification*, HR TECHNOLOGIST (Oct. 16, 2018, 11:58 PM), https://www.hrtechnologist.com/articles/background-verification/6-ways-ai-and-ml-are-improving-employee-background-verification.

129. Melanie Dawn Sage, *Child Welfare Workforce Turnover: Frontline Workers' Experiences with Organizational Culture and Climate, and Implications for Organizational Practice* (Jan. 1, 2010) (Ph.D. dissertation, Portland State Univ.), https://pdxscholar.library.pdx.edu/open_access_etds/365/.

130. THE ECONOMIST INTEL. UNIT, THE EXPERIENCE OF WORK: THE ROLE OF TECHNOLOGY IN PRODUCTIVITY AND ENGAGEMENT 4 (2019), https://theexperienceofwork.economist.com/pdf/Citrix_The_Experience_of_Work_BriefingPaper.pdf.

131. *Id.*

132. *Id.*

AI also shows promise in providing feedback to child welfare system workers, particularly those providing mental health services, about how well their interactions with clients correspond with best practices.[133] For example, the tech startup Lyssn currently uses AI and natural language processing techniques to translate recordings of therapist conversations into data, generating transcripts and metrics that provide feedback to these mental health providers in a HIPAA—and FERPA—compliant way.[134] These technologies could be implemented to assess whether a mental health provider has interacted in a trauma-informed way with a parent or child,[135] helping the provider learn and improve without the need to hire evaluators to review session recordings manually.

Identify areas of overlap in services with other state agencies to reduce inefficiencies in government. States that have modular child welfare databases that can link to other state agency databases would have a significant advantage in providing services to individuals whose needs span systems. This amount of data would allow machine learning to identify which agency is best positioned to help a particular client. For example, South Carolina's integrated state databases have allowed state research staff to link data and "demonstrate to agencies how outcomes of critical importance are actually measured in other agency data systems. In other words, these data linkages helped them show the agencies how their program outcomes were tied to those across other agencies."[136]

Digitize and analyze older agency paper records, and improve the courts' and the parties' electronic access to case files. Many child welfare agencies still keep many of their files as paper records in addition to the data stored in SACWIS systems.[137] Products like Amazon Textract can address this common challenge by using machine learning to automatically extract text and data, including from tables and forms, in virtually any document without the need for manual review, custom code, or machine learning experience.[138] The software "goes beyond simple optical character recognition (OCR) to identify the contents of fields in forms, information stored in tables, and the context in

---

133. *See, e.g.*, David C. Atkins et al., Scaling *Up* the *Evaluation* of *Psychotherapy*: *Evaluating Motivational I*nterviewing *Fidelity* via *Statistical Text Classification*, *Implementation* SCI. (Apr. 24, 2014), https://implementationscience.biomedcentral.com/articles/10.1186/1748-5908-9-49 (using technology to evaluate behavioral interventions).

134. LYSSN, https://www.lyssn.io/ (last visited Oct. 21, 2020).

135. *See generally* CHILD WELFARE INFO. GATEWAY, DEVELOPING A TRAUMA-INFORMED CHILD WELFARE SYSTEM (2015), https://www.childwelfare.gov/pubPDFs/trauma_informed.pdf (explaining the importance of trauma-informed practice).

136. ERIKA M. KITZMILLER, THE CIRCLE OF LOVE: SOUTH CAROLINA'S INTEGRATED DATA SYSTEM 1 (2014), https://www.aisp.upenn.edu/wp-content/uploads/2015/08/SouthCarolina_CaseStudy.pdf.

137. *See, e.g.*, Memorandum from the State of Wisconsin Department of Children and Families to Area Administrators et al., Contents of a Child Welfare Case Record & Imaging (Oct. 6, 2010), https://dcf.wisconsin.gov/files/cwportal/policy/pdf/memos/2010-10.pdf (discussing Wisconsin's 2010 recommendation to case workers to continue to store certain child welfare paper records in addition to the state SACWIS database records).

138. Press Release, Amazon, AWS Announces General Availability of Amazon Textract (May 29, 2019, 6:00 PM), https://press.aboutamazon.com/news-releases/news-release-details/aws-announces-general-availability-amazon-textract.

which the information is presented, such as a name or Social Security number from a tax form."[139]

Technology like this could help unearth useful and meaningful historical information buried in the generations of paper documents stored by child welfare agencies; for example, many foster children who were adopted out in prior decades lack information about their birth family history.[140] State laws vary as to the extent that states can make adoption and birth family records available upon request by a current or former foster youth, but digitizing historical records is a necessary first step.[141] Digitizing and categorizing paper records can also help the agency comply with its records retention schedule, which can have important downstream impacts on background checks.[142]

Similarly, courts and the parties to child welfare cases struggle to manage the large amounts of paper records associated with a pending case, which is to everyone's detriment in understanding a child's status before important hearings.[143] The Shanghai High People's Court in China has successfully improved operations by incorporating remote access to court documents into their operations.[144] Given the high turnover among the attorneys and judges working on a given child welfare proceeding, it is vital for participants in the child welfare system to be able to quickly review the history of the family and the case before a hearing.[145] AI and the use of cloud technology can help child welfare transition away from a reliance on paper files that contributes to poor case outcomes.

Allow for seamless nationwide searches of multiple state agency databases of child abuse and neglect findings, and other diverse data sets. Before a state can approve a prospective foster or adoptive parent to accept placement of a child, IV-E requires states to do comprehensive background checks on prospective foster or adoptive parents as a condition of receiving federal child welfare funds.[146] This includes "check[ing] any child abuse and neglect registry maintained by the State for information on any prospective foster or adoptive parent and on any other adult living in the home of such a prospective parent, and request[ing] any other State in which any such

---

139. *Id.* 118.

140. CHILD WELFARE INFO. GATEWAY, SEARCHING FOR BIRTH RELATIVES (2018), https://www.childwelfare.gov/pubPDFs/f_search.pdf.

141. *See* CHILD WELFARE INFO. GATEWAY, ACCESS TO ADOPTION RECORDS (2016), https://www.childwelfare.gov/pubPDFs/infoaccessap.pdf (explaining how to obtain adoption records).

142. See generally *State Laws on Maintaining Child Abuse and Neglect Records*, CHILD WELFARE INFO. GATEWAY, .https://www.childwelfare.gov/topics/systemwide/laws-policies/can/maintaining/ (last visited Oct. 21, 2020) (explaining the helpfulness of digitizing records for agencies).

143. See Gayle DeRose, Child Protective Services: Paperwork Burdens are a Reality, L-Tron, https://www.l-tron.com/child-protective-services-paperwork-burdens-galore/ (last visited Oct. 21, 2020) (describing the burden of too much paperwork).

144. Nesta, The AI Powered State: China's Approach to Public Sector Innovation 38 (2020), https://media.nesta.org.uk/documents/Nesta_TheAIPoweredState_2020.pdf.

145. *See* BETSY GOTBAUM, PUB. ADVOC. FOR N.Y.C., A DANGEROUS CYCLE: ATTORNEY TURNOVER AT ACS LEAVES CHILD. UNPROTECTED (2006), http://www.nyc.gov/html/records/pdf/govpub/2726a_dangerous_cycle_attorney_turnover_at_acs_leaves_children_u.pdf (illustrating the causes of attorney turnovers and the impacts it has on children).

146. 42 U.S.C. § 671(a)(20)(B).

prospective parent or other adult has resided in the preceding 5 years."[147] IV-E also requires states to comply with such requests received from another state as a condition of receiving federal funds.[148] Currently, it is necessary for each state to separately communicate and use another state's clunky, often paper-heavy process to request a search of that unique state registry for child abuse and neglect findings.[149] Large categories of employers, such as daycares, also have to manually communicate with each state agency database of administrative findings of child abuse and neglect to receive these results, which do not show up on a typical criminal background check.[150] Furthermore, almost every state does not automatically notify current employers if there is an updated criminal history or child abuse and neglect finding on an employee (California being a rare exception, which even then only notifies employers about California-specific arrests).[151]

If states entered into background check data-sharing agreements—ideally facilitated by the federal government—then AI solutions could be designed to gather automatically the relevant pieces of information from each unique state database without requiring the creation of a unified database. Private sector companies have developed AI applications for background check processes that gather data from government records and a vast array of publicly available sources.[152] Notably, improvements in the ability to search state databases of administrative child abuse and neglect findings should be accompanied by clear notices to the subject of the background check, with opportunities to contest and seek expungement of old or incorrect database records.

Once aggregated data and search functions are in place, there are additional research applications that become possible. One example currently in development is the use of AI to remove personally identifiable information (PII) from large datasets of child forensic interviews performed after an alleged report of child abuse, allowing researchers to develop improved techniques to treat post-traumatic stress disorder.[153] There may be similar insights available to researchers based on the details described in unsubstantiated vs. substantiated child abuse and neglect findings and documentation that could improve the investigation of such reports.

Help creating and improving individual treatment plans. Expanding upon the potential around health care diagnosis tools discussed above, one

---

147. 42 U.S.C. § 671(a)(20)(B)(ii).

148. *Id.*

149. *See* NAT'L RES. CTR. FOR PERMANENCY FAMILY& FAM. CONNECTIONS, STATE CHILD ABUSE REGISTRIES (2014), http://www.hunter.cuny.edu/socwork/nrcfcpp/downloads/policy-issues/State_Child_Abuse_Registries.pdf (including charts describing each state's process, which often includes filling out and faxing paper forms).

150. *See* CHILD WELFARE INFO. GATEWAY, REVIEW AND EXPUNCTION OF CENTRAL REGISTRIES AND REPORTING RECORDS (2018), https://www.childwelfare.gov/pubPDFs/registry.pdf (discussing state-by-state classifications of child abuse and neglect findings).

151. *See* CAL. DEP'T JUST., CONTRACT FOR SUBSEQUENT ARREST NOTIFICATION SERVICE (2018), https://oag.ca.gov/sites/all/files/agweb/pdfs/fingerprints/forms/subarr.pdf (describing opt-in process for employers to receive background check update notifications on employee arrests that occur in California).

152. BasuMallick, *supra* note 128, at 1.

153. Amazon Web Services, *Innovation in Technology for Public Interest #Tech4PI*, YOUTUBE (Aug. 14, 2019), https://www.youtube.com/watch?v=HERzKY-F1AI&ab_channel=AmazonWebServices.

promising example of an AI tool in development could be one that helps parents and children with assessing mental health risk and suicide prevention.[154] One such tool helped a major metropolitan area demonstrate a four-to-one return on investment for its treatment of prisoners with mental health issues, and led to an eighty percent crisis diversion rate.[155] Such tools synthesize a wide array of data streams and match them to one person, including mental health issues, criminal data, housing information and medical records, and use AI to help identify who is likely to follow a care plan or be readmitted.[156]

The main caution around these mental health tools, as discussed *infra* in Part III(C), is that such tools should be thoughtfully deployed only with specific legal safeguards in place: their least problematic use is for governments to make helpful services available to individuals, as opposed to being used as evidence in court to support removing children from their parents.[157]

Forecasting child welfare system needs. Another area where AI tools could be particularly helpful is in demand forecasting,[158] such as to help answer questions like how many social workers a state is likely to need in a given area in the next ten years, or what the most common questions are that foster children have at different stages of a child welfare case. Given the well-documented fact that federal census data undercounts children, which in turn reduces the available funding for schools and other government services aimed at helping children,[159] harnessing the unique information available in state child welfare systems could help improve system planning and provide data to back up funding requests.

C. Use of AI to Predict who will Commit Child Abuse and Neglect, or to Identify Statistically Rare Cases of Financial Fraud, Should not be a Focus of the Technology Movement in Child Welfare Absent a Design Process that Considers Bias in Underlying Data and the Exceptionally High Costs of False Positives.

To use AI effectively, decisionmakers need to understand that AI is simply a reflection of a mathematical model trained on data. Perhaps the most critical thing to understand about AI and data science in general is the maxim "garbage in, garbage out"—the results from an AI tool will only be as good as the data the

154. Robin Mayerhoff, *HarrisLogic Uses Technology to Help Slow the Rising Tide of Suicide Rates*, FORBES (Nov. 15, 2018, 10:00 AM), https://www.forbes.com/sites/sap/2018/11/15/harrislogic-uses-technology-to-help-slow-rising-suicide-rates/#75a39aad1572.

155. *Id.*

156. *Id.*

157. *See, e.g.*, *Should A Mental Illness Mean You Lose Your Kid?*, PROPUBLICA (May 30, 2014, 5:45 AM), https://www.propublica.org/article/should-a-mental-illness-mean-you-lose-your-kid (describing an instance where a mental health issue was used to remove a child from her parents).

158. Steve Banker, *Machine Learning and Artificial Intelligence in Demand Planning*, FORBES (Dec. 8, 2017, 7:51 AM), https://www.forbes.com/sites/stevebanker/2017/12/08/machine-learning-and-artifical-intelligence-in-demand-planning/#3dc0d054e83d.

159. *The Undercount of Young Children*, U.S. CENSUS BUREAU, https://www.census.gov/programs-surveys/decennial-census/2020-census/research-testing/undercount-of-young-children.html (last updated June 13, 2019).

model is trained on.[160] And rather than revealing insights about which families are most likely to abuse and neglect their children, historical child welfare data more accurately reflects that minority and impoverished communities experience much higher rates of policing. For example, a 2019 study showed that "police file more reports of child abuse and neglect in counties with high arrest rates, and that policing helps explain high rates of maltreatment investigations of American Indian-Alaska Native children and families."[161] The police are more likely to discover and arrest domestic violence perpetrators who live in row houses, semi-detached homes, and apartment buildings because the police typically receive domestic violence reports by neighbors who call in to complain.[162] Therefore, if an algorithm is trained using police data, it "will overpolice people in homes with shared walls (who tend to be poorer), and underpolice people in detached homes (who tend to be richer)."[163] In another example of how problematic data can reinforce racist algorithmic results, the UK recently agreed to stop using a "streamlining" algorithm to help decide visa applications after it came to light that one of the software's key inputs was the nationality of the visa applicant.[164] The algorithm relied on a government list of nationalities that should be red-flagged for additional intense screening; those individuals were then highly likely to be denied a visa.[165] The government then used the "data" from those visa denials to determine which nationalities should be on the red flag list,[166] contributing to the cycle of bad data.

As a result, when using AI to aid in government decision-making, the implementation team must understand and evaluate how problematic it would be to have a particular data science model produce what are known as "Type 1" and "Type 2" errors: false negatives and false positives.[167] If it is known that it would be particularly devastating for an algorithm to produce false positives (for example, for child protective services to mistakenly declare a hotline call deserving of an in-person investigation, potentially unnecessarily causing a CPS worker to err on the side of caution and remove a child from the home), the data

---

160. *Garbage In, Garbage Out: How to Prepare Your Data Set for Machine Learning*, CIKLUM (Jan. 11, 2019), https://www.ciklum.com/blog/garbage-in-garbage-out-how-to-prepare-your-data-set-for-machine-learning; *see also* Gary Smith, *Why Genuine Human Intelligence is Key for the Development of AI*, FAST CO. (July 30, 2019), https://www.fastcompany.com/90381653/why-genuine-human-intelligence-is-key-for-the-development-of-ai (noting most common pitfall for data science is using bad data).

161. Frank Edwards, *Family Surveillance: Police and the Reporting of Child Abuse and Neglect*, 5 RSF: THE RUSSELL SAGE FOUND. J. SOC. SCIS. 50, 50 (2019), https://www.rsfjournal.org/content/rsfjss/5/1/50.full.pdf.

162. Cory Doctorow, *Garbage In, Garbage Out: Machine Learning Has Not Repealed the Iron Law of Computer Science*, BOING (May 29, 2018, 6:17 AM), https://boingboing.net/2018/05/29/gigo-gigo-gigo.html (citing study by Human Rights Data Analysis Group).

163. *Id.*

164. *Home Office Drops 'Racist' Algorithm from Visa Decisions*, BBC (Aug. 4, 2020), https://www.bbc.com/news/technology-53650758.

165. *Id.*

166. *Id.*

167. *See* Amitav Banerjee et al., *Hypothesis Testing, Type I and Type II Errors*, 18 INDUS. PSYCH. J. 127 (2009), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2996198/#:~:text=A%20type%20I%20error%20(false,actually%20false%20in%20the%20population ("A type I error (false-positive) occurs if an investigator rejects a null hypothesis that is actually true in the population; a type II error (false-negative) occurs if the investigator fails to reject a null hypothesis that is actually false in the population.").

scientists could design the algorithm to make false positives less likely. Accordingly, data scientists need to engage with the individuals most impacted by Type 1 and Type 2 errors—child welfare system participants—to assess the risks of each type of error.

Government use of AI in decision-making also raises concerns that because the most sophisticated AI technology is capable of identifying patterns that are outside the ability of humans to readily understand, decisions may not be explainable as AI becomes more accurate.[168] Therefore, decisionmakers who recognize the need for government use of AI to be explainable should use caution in introducing AI models that defy explanation, even if they appear to produce more accurate results.[169]

Failure to respect these limitations of AI in government decision-making is evident in the current use case for predictive analytics that has received the most attention from child welfare leaders interested in technology innovation: using an algorithm to make a recommendation about whether state hotline calls reporting child abuse and neglect should be "screened in" for investigation.[170]

While publications including the New York Times wrote admiringly of Allegheny County, Pennsylvania's algorithm, the Allegheny Family Screening Tool (AFST),[171] others sounded the alarm about AFST's underlying data sources.[172] AFST works by assigning a risk score to a particular family "based on a statistical analysis of four years of prior calls, using well over 100 criteria maintained in eight databases for jails, psychiatric services, public-welfare benefits, drug and alcohol treatment centers and more."[173] Critics aptly pointed out that many of the algorithm's data points mainly reflect a family's poverty, including use of state benefits programs intended to supplement income and provide food security.[174] In response to this criticism, Allegheny County's child welfare director responded that AFST "simply augments the human decision whether to investigate a call alleging abuse or neglect by quickly distilling information already available to the call screener[.]"[175] However, the fact that the system apparently automatically triggers the family to be "screened in" if the AFST risk score is above a certain level—a result that can only be overridden

168. *See* Jesus Rodriguez, *Interpretability vs. Accuracy: The Friction that Defines Deep Learning*, LINKEDIN (June 6, 2018), https://www.linkedin.com/pulse/interpretability-vs-accuracy-friction-defines-deep-jesus-rodriguez (describing the difficulty in understanding deep learning models).

169. *Id.*

170. Dan Hurley, *Can an Algorithm Tell When Kids Are in Danger?*, N.Y. TIMES MAG. (Jan. 2, 2018), https://www.nytimes.com/2018/01/02/magazine/can-an-algorithm-tell-when-kids-are-in-danger.html.

171. Hurley, *supra* note 170; Naomi Riley, *Can Big Data Help Save Abused Kids?*, REASON (Feb. 2018), https://reason.com/2018/01/22/can-big-data-help-save-abused#comment.

172. Richard Wexler, *Pittsburgh Misuses Big Data to Target Poor Children for Abuse Investigations*, YOUTH TODAY (Mar. 28, 2018), https://youthtoday.org/2018/03/pittsburgh-misuses-big-data-to-target-poor-children-for-foster-care; Virginia Eubanks, *A Child Abuse Prediction Model Fails Poor Families*, WIRED (Jan. 15, 2018, 8:00 AM) https://www.wired.com/story/excerpt-from-automating-inequality.

173. Hurley, *supra* note 170.

174. Wexler, *supra* note 172; Eubanks, *supra* note 172.

175. Marc Cherna, *We Will Use All Resources to Keep Children Safe*, PITTSBURGH POST-GAZETTE (Aug. 25, 2019), https://www.post-gazette.com/opinion/letters/2018/03/23/We-will-use-all-resources-to-keep-children-safe/stories/201803230094.

by a supervisor—seems to compromise the ability of an individual worker to make a decision about a case that is untainted by the AFST score.[176]

Using AI to attempt to predict child abuse and neglect via historically problematic data sources does not exemplify the best use of technology innovation in the child welfare system, and may actually have the unintended effect of discouraging government investment in emerging technology. To date, these predictive algorithms have not demonstrated promising results,[177] perhaps because they were trained on problematic data sources. Just as civil liberties advocates have pointed out in response to government use of algorithmic decision-making and facial recognition in the criminal justice context, the technology is prone to discriminating against women and minorities due to the nature of the data the algorithms are trained on.[178] Just as in the criminal justice system, "biases in data sets will not only be replicated in the results, they may actually be exacerbated"—one key example being that criminal justice data reflects the disproportionate number of arrests of people of color,[179] also a well-known issue in the child welfare system.[180] Social justice organizations have aptly argued that the issue is not just the biased results of applying an algorithm to problematic data, but that these algorithms give an undeserved halo of

176. *See* Stephanie K. Glaberson, *Coding Over the Cracks: Predictive Analytics and Child Protection*, 46 FORDHAM URB. L. J. 307, 333–34 (2019) (describing how the AFST score is used before a worker is able to evaluate a case).

177. Memorandum from Judge Michael Nash (Ret.), Exec. Director, Office of Child Protection to Supervisor Mark Ridley-Thomas et al., 10 (May 4, 2017) (on file with author) (noting the extremely high false positive rate of Los Angeles County's algorithm: "While the tool correctly detected a high number of children (171 cases) at the highest risk for abuse, it also incorrectly identified an extremely high number (3,829 cases) of false positives (i.e., children who received high risk scores who were not at risk for a negative outcome)."); David Jackson & Gary Marx, *Data Mining Program Designed to Predict Child Abuse Proves Unreliable, DCFS says*, CHI. TRIBUNE (Dec. 6, 2017), https://www.chicagotribune.com/investigations/ct-dcfs-eckerd-met-20171206-story.html (discussing how Illinois abandoned predictive analytics hotline project because initial algorithm missed most serious cases, including ones resulting in child deaths).

178. *See* LEADERSHIP CONFERENCE ON CIVIL & HUMAN RIGHTS, THE USE OF PRETRIAL "RISK ASSESSMENT" INSTRUMENTS: A SHARED STATEMENT OF CIVIL RIGHTS CONCERNS, http://civilrightsdocs.info/pdf/criminal-justice/Pretrial-Risk-Assessment-Full.pdf (opposing, in letter signed by over 100 civil rights organizations, government use of risk assessment algorithms in pretrial decision-making even as alternative to problematic money bail practice) [hereinafter LEADERSHIP CONFERENCE]. *See also* Brief for American Civil Liberties Union et al. as Amici Curiae Supporting Petitioner at 6–7, Lynch v. Florida, 2019 WL 3249799 (No. SC2019-0298); Inioluwa Deborah Raji & Joy Buolamwini, *Actionalby Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products*, ASSOCIATION FOR THE ADVANCEMENT OF ARTIFICIAL INTELLIGENCE (2019), http://www.aies-conference.com/wp-content/uploads/2019/01/AIES-19_paper_223.pdf (concluding that "the potential for weaponization and abuse of facial analysis technologies cannot be ignored nor the threats to privacy or breaches of civil liberties diminished even as accuracy disparities decrease"); Karen Hao, *AI is Sending People to Jail – and Getting It Wrong*, MIT TECH. REV. (Jan. 21, 2019), https://www.technologyreview.com/s/612775/algorithms-criminal-justice-ai/?utm_campaign=the_algorithm.unpaid.engagement&utm_source=hs_email&utm_medium=email&utm_content=69523284&_hsenc=p2ANqtz-_8kaef_OGYyHIH3kD5Fq6mfYa4AuGl5EwD0GvwldCYVZoSYcMVE8Xx9bvzkU4ZAkuzSwRovx7PgrNVS2_GoIF0HUjO4Q&_hsmi=69523284 (discussing range of criminal risk assessment algorithms).

179. LEADERSHIP CONFERENCE, *supra* note 178.

180. *See* RACIAL DISPROPORTIONALITY, *supra* note 51 (discussing prevalence of racial disproportionality in the child welfare system); *see also* Frank Edwards, *Family Surveillance: Police and the Reporting of Child Abuse and Neglect*, 5 THE RUSSELL SAGE FOUNDATION J. OF SOCIAL SCIENCES 50 (2019), https://muse.jhu.edu/article/720075 (discussing the relationship between racial inequities in policing and frequency of child welfare investigations).

impartiality around government systems that desperately need fundamental change.[181]

Similarly, prioritizing AI applications that purport to detect fraud in child welfare systems would be an unhelpful use of limited technology resources given the potential for devastating false positives and generally low statistical rates of fraud.[182] Consider a $47 million automated fraud detection system adopted by the state of Michigan in 2013, which made about 48,000 fraud accusations against unemployment insurance recipients (a five-fold increase from the prior system).[183] The automated system lacked any human intervention mechanism before it garnished wages, levied bank accounts, and intercepted tax refunds of the poor—and a state review later determined that 93% of the fraud determinations were wrong.[184]

Accuracy aside, at least one international court that has weighed the government's use of AI to detect financial fraud against individuals' privacy rights has ruled that it violates human rights. The June 2020 decision by the Hague District Court held that an AI system used by the Dutch government to detect social benefit, tax, and other forms of financial welfare fraud violated Article 8 of the European Convention on Human Rights that protects the right to respect for private and family life, home, and correspondence.[185] The Court's decision rested on evaluating a number of facts about how the AI functioned against the intrusion on an individual's private life, and noted that it was unable to assess the truth of the government's characterization of the AI because the government did not disclose any specifics about the risk model.[186] The opinion demonstrates the importance of transparency about the factors the AI is evaluating, and of the need for governments to engage in a meaningful analysis of the potential privacy impacts *before* introducing AI decision-making tools.

To that end, instead of focusing limited resources on predicting child abuse and neglect or preventing financial fraud at this stage of the U.S. child welfare system's technology journey, it would be strategic to achieve some early technology "wins" by applying data science tools to three key areas: Governments should consider prioritizing investment in AI tools that can reduce the administrative burden child welfare systems face, recommend services for families, and improve our understanding of the effectiveness of the agency's own actions.[187] While these strategies still require careful implementation, they

181. *Id.*

182. Gilman, Michele, *AI Algorithms Intended to Root out Welfare Fraud Often End Up Punishing the Poor Instead*, CHI. REPORTER, (Feb. 20, 2020) https://www.chicagoreporter.com/ai-algorithms-intended-to-root-out-welfare-fraud-often-end-up-punishing-the-poor-instead/.

183. *Id.*

184. *Id.*

185. Den Haag [Hague District Court] Feb 5, 2020, UITSPRAKEN. https://uitspraken.rechtspraak.nl/inziendocument?id= ECLI:NL:RBDHA:2020:1878.

186. *Id.* at 6.49.

187. *See supra* Section III(B) (advocating for different priorities for use of new technology in child welfare); *see, e.g.*, Annie Sweeney, *Can Police Data Predict How "Bad Apple" Officers Influence Their Fellow Cops? New Study Says Yes*, CHI. TRIBUNE (Aug. 1, 2019), https://www.chicagotribune.com/news/criminal-justice/ct-predicting-bad-police-behavior-20190801-xumudeezmjalbbpmqwyvh26tdi-story.html (describing University of Chicago study using predictive analytics to understand officer use of force).

are less likely to raise the same level of ethical concerns in the event of Type 1 and Type 2 errors.[188] Moreover, reducing the administrative burden on social workers and other state child welfare workers is a powerful strategy to modernize the profession and attract a new generation of talent to the field.

D. Openness and Explainability are Key Values in Implementing any AI in Child Welfare.

For all of the potential AI applications discussed above, it is critical that governments put a high value on openness about why and how the algorithms are being used, and what data sources are being used to produce the results.[189] Recent scholarship in the context of autonomous driving and robotic systems have characterized this concept as "explainable agency," stating that "[a]n intelligent system exhibits explainable agency if it can provide, on request, the reasons for its activities."[190] That paper suggests that data scientists should incorporate this value in algorithmic design, including designing the ability to retrieve an audit trail.[191] Notably, the European Union General Data Protection Regulation requires companies to provide users with an explanation for decisions made by automated systems,[192] and the core recommendation of the AI Now Institute's 2017 report was that any high stakes matters, including public welfare, should not rely on "black-box" AI (an AI model that is not readily understandable by humans).[193]

Government leaders interested in exploring some of the technology applications discussed above should embrace thoughtful algorithmic design practices and implement a mandatory ethical review, which could be based on model guidelines issued by several AI think tanks.[194] By proactively thinking through potential sources of bias in the data and the algorithm's processes,

188. Banerjee et. al., *supra* note 167.

189. Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249 (2007); Paul Schwartz, *Data Processing and Government Administration: The Failure of the American Legal Response to the Computer*, 43 HASTINGS L.J. 1321 (1992); Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633 (2016) (noting ability to reconstitute decisions demonstrates procedural regularity in critical decision processes and supports trust in automated systems even when they remain partially non-disclosed).

190. Pat Langley, *Explainable, Normative, and Justified Agency*, ASSOCIATION FOR THE ADVANCEMENT OF ARTIFICIAL INTELLIGENCE (2019), https://pdfs.semanticscholar.org/d8ed/29985999ca9642e7940fbf7484f544df54e5.pdf.

191. *Id.*

192. Bryce Goodman & Seth Flaxman, *EU Regulations on Algorithmic Decision-Making and A "Right to Explanation,"* 2016 ICML WORKSHOP ON HUMAN INTERPRETABILITY IN MACHINE LEARNING 26 (2016), http://metromemetics.net/wp-content/uploads/2016/07/1606.08813v1.pdf.

193. ALEX CAMPOLO ET.AL, AI NOW 2017 REPORT (Andrew Selbst & Solon Barocas eds. 2017); *see also* Jay Stanley, *Pitfalls of Artificial Intelligence Decisionmaking Highlighted In Idaho ACLU Case*, ACLU (June 2, 2017, 1:30 PM), https://www.aclu.org/blog/privacy-technology/pitfalls-artificial-intelligence-decisionmaking-highlighted-idaho-aclu-case (discussing *K.W. v. Armstrong*, a class action lawsuit brought by ACLU representing Idahoans with developmental and intellectual disabilities receiving Medicaid assistance; when state claimed trade secrets prevented disclosure of algorithm that cut benefits by 30 percent, plaintiffs sued to gain access and discovered significant coding errors).

194. *See, e.g.*, Sam Brown, *An Agile Approach to Designing for the Consequences of Technology*, MEDIUM, https://medium.com/doteveryone/an-agile-approach-to-designing-for-the-consequences-of-technology-18a229de763b (last visited Oct. 21, 2019); Alix, *Working Ethically at Speed*, MEDIUM (May 7, 2018) https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.

governments can help avoid some of the initial mistakes that have been made by the criminal justice system.[195]

One nuance to the caution about relying on black-box AI is that in light of the tension between accuracy and explainability,[196] there may be some contexts in which it could be appropriate to use black-box algorithms, provided that they do not impact any substantive rights of individuals involved in the child welfare system. For example, there may be some applications and tools to simplify social worker administrative tasks that would not affect any substantive individual rights. But those decisions to use black-box AI in a particular circumstance must themselves be transparent and explainable.[197] For example, they must be documented clearly in a centralized and regularly updated policy that tracks agency-wide use of AI tools and the datasets on which they rely.

In addition, one key takeaway for child welfare leaders to understand is that avoiding bias in algorithms is not as simple as just training the algorithms to ignore data points revealing race, color, religion, gender, disability, or family status.[198] In a recent paper describing how data scientists can optimally adjust an algorithmic predictor to eliminate discrimination, the researchers explain that this concept of "fairness through unawareness" is ineffective because of "redundant encodings, [and] ways of predicting protected attributes from other features."[199] In light of this mathematical reality, government leaders should embrace a deliberate and thoughtful planning process with safeguards—including algorithmic transparency, stakeholder notice, and the ability to provide feedback—before implementing AI in government social services.[200] The July 2020 publication by the National Security Commission on Artificial Intelligence provides a helpful overview and starting point for recommended core principles

195. *See, e.g.*, Jeff Larson et al., *How We Analyzed the COMPAS Recidivism Algorithm*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm (discussing pitfalls of recidivism algorithm); PARTNERSHIP ON AI, REPORT ON ALGORITHMIC RISK ASSESSMENT TOOLS IN THE U.S. CRIMINAL JUSTICE SYSTEM, https://www.partnershiponai.org/report-on-machine-learning-in-risk-assessment-tools-in-the-u-s-criminal-justice-system (last visited Oct. 21, 2020).

196. Rodriguez, *supra* note 168.

197. *See* Banerjee et al., *supra* note 167.

198. Moritz Hardt, et al., *Equality of Opportunity in Supervised Learning*, PROC. NEURIPS (2016), https://arxiv.org/pdf/1610.02413.pdf.

199. *Id.* at 1. However, note that this is not to suggest that AI could not have a place in generating a race-blind administrative record that a human could then review to reduce that person's bias in decision making. *See* Timothy Williams, *Black People are Charged at a Higher Rate Than Whites. What if Prosecutors Didn't Know Their Race?*, N.Y. TIMES (June 12, 2019), https://www.nytimes.com/2019/06/12/us/prosecutor-race-blind-charging.html?action=click&module=Latest&pgtype=Homepage (discussing July 2019 introduction of software designed by Stanford University researchers that will automatically redact suspects and victims' races and names, as well as locations where crimes were said to have been committed, before prosecutorial review. The only information prosecutors will initially have access to in making charging decision is officer's incident report, which typically includes the reason someone was stopped before arrest, evidence that crime was committed, and statements by witnesses and the suspect.).

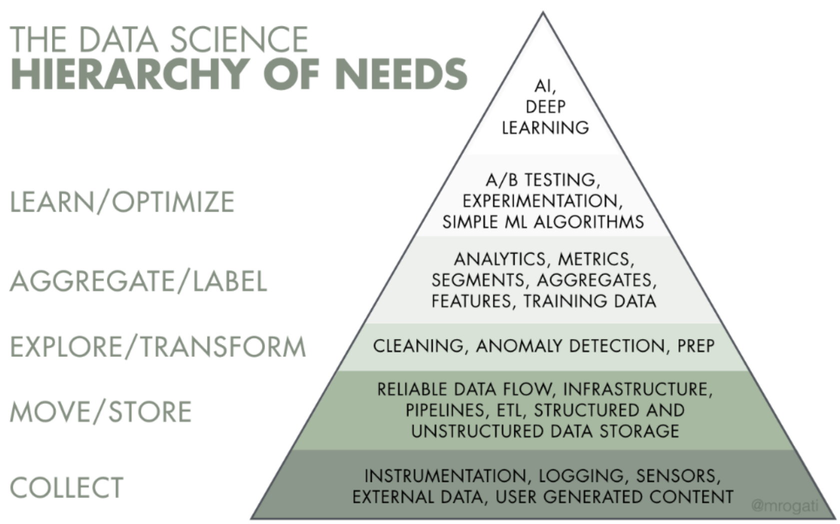
200. *See, e.g.*, D.J. Pangburn, *Washington Could Be the First State to Rein in Automated Decision Making*, FAST COMPANY (Feb. 8, 2019), https://www.fastcompany.com/90302465/washington-introduces-landmark-algorithmic-accountability-laws (discussing proposed but ultimately unsuccessful legislation in WA state to make public-sector automated systems and data sets they use "freely available by the vendor before, during, and after deployment for agency or independent third-party testing, auditing, or research.").

to guide AI implementation by government agencies, including the Department of Health and Human Services.[201]

IV. PRACTICAL OBSTACLES AND SOLUTIONS TO TECHNOLOGY TRANSFORMATION OF CHILD WELFARE SYSTEMS

A. Federal Investment in an Open-Source, Model CCWIS System Would Rapidly Move the Needle for Child Welfare Systems Across the Country.

One of the most fundamental steps in addressing the practical challenges of making child welfare data more accessible and useful to states, researchers, and the public is for all states to transition as soon as possible to a CCWIS system. The issue may be best explained by the following graphic, which describes the technology infrastructure that must be in place to allow for advanced technology tools such as AI:[202]



[203]

Child welfare systems are still at the bottom half of this pyramid. They struggle with reliable data flow and clean, structured data that are ready for more complex analysis. The need for a modular system with the capacity to support

201. *Key Considerations as a Paradigm for Responsible Development and Fielding of Artificial Intelligence*, NATIONAL SECURITY COMMISSION ON ARTIFICIAL INTELLIGENCE LINE OF EFFORT ON ETHICS AND RESPONSIBLE AI Quarter 2 Report*:* (Jul. 22, 2020), https://drive.google.com/file/d/1_zkNkT3Trz3rtFc8KVrEBNlg2R9MaUpi/view. *See also* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/890699/Guidelines_for_AI_procurement__Print_version_.pdf (UK Guidelines discussing procurement considerations for governments implementing AI).

202. Monica Rogati, *The AI Hierarchy of Needs*, HACKERNOON (June 12, 2017), https://hackernoon.com/the-ai-hierarchy-of-needs-18f111fcc007.

203. *Id.*

emerging machine learning algorithms that can make use of the data collected is too urgent to be left to individual states to develop at their own pace and at great cost to themselves.[204] Instead, the federal government would ideally sponsor a public-private partnership to create a demonstration CCWIS system that is capable of producing parallel data sets from every state that would support the use of inventive AI applications. Making this data available at the federal level is critical because currently, AI requires large amounts of data to detect the broad patterns that lead to the type of insights that would most benefit state child welfare systems.[205]

The federal government has offered to pay up to half the cost for transitioning to a CCWIS system (for the aspects of the system that deal with IV-E eligible youth), but progress towards that outcome is painfully slow.[206] While forty six states as of August 2018 have declared their intent to transition towards a CCWIS system,[207] no state has yet achieved an operational CCWIS system, with the most advanced states still in the indefinite "pre-procurement planning" stage.[208] The wait-and-see approach may be due to states' experience in implementing SACWIS systems—while some states paid hundreds of millions of dollars to build SACWIS systems, others bought and adopted the models that other states built for far less.[209] Indeed, states appear to be watching and waiting for the state of Arizona to implement its CCWIS system, a nearly $100 million project with 50% of the project covered by matching federal funds, and the state legislature providing the rest.[210]

For most states, investing nearly $50 million in new child welfare technology is out of the question due to budget constraints.[211] One of the key challenges of building a new child welfare system while still maintaining the existing infrastructure is that the existing systems are often so outdated that upgrades and patches take large amounts of developer time to fix. The problems with SACWIS systems are not limited to the fact that they cannot easily share data. As Arizona's child welfare agency noted in explaining the issue to the legislature, "[d]ue to the age, complexity, and inefficiencies" of Arizona's SACWIS system, the agency cannot enhance the SACWIS system fast enough to provide important processing functions the agency has identified as crucial:

204. *See Arizona* discussion, *infra* Section IV (showcasing the inadequacies in Arizona's Child Welfare Information Technology's infrastructure).

205. *See* H. James Wilson et al., *The Future of AI Will Be About Less Data, Not More*, HARV. BUS. REV. (Jan. 14, 2019), https://hbr.org/2019/01/the-future-of-ai-will-be-about-less-data-not-more (discussing that AI currently requires large datasets, but that this may change in the future as neural networks become more sophisticated).

206. *CCWIS Status*, *supra* note 20.

207. *CCWIS Status at the End of Transition Period (July 31, 2018)*, CHILD. BUREAU: AN OFF. OF THE ADMIN. FOR CHILD. & FAMILIES (Mar. 6, 2019), https://www.acf.hhs.gov/cb/resource/ccwis-status-end-transition.

208. *CCWIS Status*, *supra* note 20.

209. Kelly, *supra* note 18.

210. Letter from Gregory McKay, *supra* note 21, at 60 (noting that anticipated project build cost estimated at $86,088,076, with 50% of the project cost covered by federal match). Microsoft is the private sector vendor working on this project with Arizona. *Id.* at 55.

211. *Id.* at 55.

> These functions include the creation, control, and management of clients, intake functions, ongoing case management, development of new interfaces for data mining, mobility access options, visitation report entries, court record production, Business Intelligence (BI) processing capabilities, comprehensive reporting, decisions support processing, and general system enhancements. Key system shortfalls identified include:
>
> - Poor system usability
> - Ineffective tools to support and drive DCS business processes
> - Deficient searching and matching functionality
> - Lack of a mobile platform
> - Limited reporting capabilities
> - Lack of DCS workforce management capability
> - Incomplete provider and service management
> - Lack of capability to produce required forms and notices
> - Inadequate collaboration with other agencies and system interfaces
> - Poor data quality and data integrity
> - Lack of compliance with new federal requirements which will limit cost reimbursement.[212]

Arizona's technology infrastructure inadequacies are not unique. Child welfare Information Technology departments typically find themselves in the position of devoting their entire budget to keeping the current system functioning. And it is challenging for child welfare system leaders to convince state legislatures of the need to invest millions of dollars in a new technology project when faced with other immediate system needs, such as social worker salaries or increased reimbursements to foster parents.

The most significant financial cost of creating a new system can mainly be measured in the hours of developer time needed to build a new system that meets the states' needs and federal CCWIS requirements, achieved by hiring developers as state employees or contracting the work out.[213] Therefore, the most effective way to shorten the time frame for creating a functional CCWIS system—and reflect the appropriate level of urgency given the number of children affected by our current systems—would be for the federal government, nonprofit organizations, and the private sector to work together to develop a model prototype CCWIS system and license it without charge to all states that wish to adopt it.

Instead of leaving it up to individual states to bear the enormous cost of this endeavor, producing vastly different systems that still may only leave us with the ability to compare outcomes across the minimal data points required by federal regulations, the federal government could sponsor the design of a cutting-edge system that would lay the foundation for deep cross-system

212. *See id.* (citing 2020 budget).
213. *See id.* at 55–58 (describing project requirements and including budget description overview).

collaboration. While an Obama administration 2016 press release highlighted a federal effort with Salesforce to design a prototype system, the project appears to have faded away in the Trump administration.[214] But as recently as March 2019, the federal government's U.S. Data Federation Project has expressed support for various data-sharing projects across federal agencies, explaining that "[i]ts goal is to make it easier to collect, combine, and exchange data from disparate sources."[215] Creating a federal template software model that could be customized to individual state needs without huge initial investment costs would be exceedingly helpful in speeding up the timeline to making this a reality.

In addition, there are important ethical implications of having the five major private sector companies (Google, Facebook, Microsoft, Apple, and Amazon—referred to by some ethicists as the "frightful five"[216]) that already control the majority of consumer data sell and license their products to individual states to create new CCWIS systems.[217] Having large, powerful private companies, as opposed to the federal government, negotiate data-sharing terms individually with each state is likely to result in some of these companies owning and/or having access to the data of struggling children and families, and limited access to the algorithms by the public that would want to evaluate their fairness.[218] Ideally, the federal government should be the entity playing that critical role in defining the relationship of new technologies like AI with the social services sector. As with the federal highway systems, development of an effective social services technology infrastructure that can be optimized as AI evolves is too critical to our nation's collective future to be left to an ad hoc process.

B. Training the Workforce and Handling Requests for Data Access

The other two key practical concerns in moving forward with technology innovation is addressing the impacts on the child welfare workforce, as well as the burden it would place to individual state agencies to process research requests.

For the technology transformation discussed in this article to be successful, the child welfare workforce must recognize the value that improved technology can bring to their practice. Government leaders should approach this challenge

214. Office of the Press Secretary, *FACT SHEET: First Ever White House Foster Care & Technology Hackathon*, OBAMA WHITE HOUSE ARCHIVES (May 26, 2016), https://obamawhitehouse.archives.gov/the-press-office/2016/05/26/fact-sheet-first-ever-white-house-foster-care-technology-hackathon.

215. Julia Lindpaintner, *The U.S. Data Federation Wants to Make It Easier to Collect, Combine, and Exchange Data Across Government*, U. S. DATA FED'N. (Mar. 5, 2019), https://18f.gsa.gov/2019/03/05/the-us-data-federation.

216. *See generally* Farhad Manjoo, *Tech's Frightful Five: They've Got Us*, N.Y. TIMES (May 10, 2017), https://www.nytimes.com/2017/05/10/technology/techs-frightful-five-theyve-got-us.html (discussing the threats tech giants have on society).

217. Paul Nemitz, *Constitutional Democracy and Technology in the Age of Artificial Intelligence*, ROYAL SOC'Y PHIL. TRANSACTIONS A 1, 3 (2018), https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2018.0089.

218. *See* Taylor R. Moore, *Trade Secrets and Algorithms as Barriers to Social Justice*, CTR. FOR DEMOCRACY & TECH. (August 2017), https://cdt.org/files/2017/08/2017-07-31-Trade-Secret-Algorithms-as-Barriers-to-Social-Justice.pdf (discussing legal ramifications within trade secret law with civil rights and other societal considerations due to a lack of social balancing mechanisms).

thoughtfully and with deep respect for the psychological and administrative burdens that caseworkers already experience in the course of their caseload.[219] One study that examined frontline caseworker retention in Oregon indicated that retention may be improved by maintaining an organizational culture and climate that is empowering to workers and that encourages workers to be a part of the change process."[220] Accordingly, child welfare leaders should involve a representative group of caseworkers with varying years of experience in the field when starting to develop a model CCWIS system as well as AI programs intended to modernize social work practice. In technology speak, governments must be aware of the importance of hiring user experience researchers.

Critically, child welfare systems must communicate that person-to-person interactions remain as important as ever, and the goal of this technology revolution is to increase such interactions.[221] In theory, using AI tools to reduce administrative burdens in particular, which eat up a significant amount of caseworker time, will allow social workers to increase patient satisfaction, adhere closer to recommendations and prescriptions, reduce anxiety and stress,[222] and increase empathy.[223] The medical context has been clinically shown to improve clinical outcomes. Furthermore, AI is not a replacement for a caseworker's judgment calls. As discussed in the book *Prediction Machines: The Simple Economics of Artificial Intelligence*, our society will actually place increased value in this new era on developing judgment about how to use the data signals generated by AI.[224] And experts have highlighted social worker jobs as some of the least likely jobs to be replaced by automation.[225] Data systems can help us make sense of huge amounts of available information, but they are not a substitute for human connection and understanding.

In addition, an important practical consideration in adopting the recommendations in this article is the increased burden it would place on state agencies to process data requests by researchers.[226] This issue could be addressed by centralized federal government assistance and management of requests for state-level child welfare data, similarly to how NDACAN currently

219. Brian E. Bride et al., *Correlates of Secondary Traumatic Stress in Child Protective Services Workers*, 4 J. OF EVIDENCE-BASED SOC. WORK 69, 77 (2007).

220. Sage, *supra* note 129.

221. *See* TOPOL, *supra* note 103, at 285–305 (discussing how AI will give doctors the gift of time, allowing for increased ability to demonstrate empathy and interact with patients).

222. Topol, *supra* note 103, at 290–91.

223. *See A Parent's Perspective on Child Welfare & Family Engagement*, BE STRONG FAMILIES (May 31, 2018), https://www.bestrongfamilies.org/news/2018/5/30/a-parents-perspective-on-child-welfare-family-engagement (explaining that parents at risk of losing parental rights "need someone who can help them navigate the system, preferably someone compassionate, empathetic and respectful who wants to build a positive, trusting relationship with them.").

224. AJAY AGRAWAL ET AL., *PREDICTION MACHINES: THE SIMPLE ECONOMICS OF ARTIFICIAL INTELLIGENCE 19–20, 15, (HARV. BUS. REV. PRESS* (2018)).

225. Paul Petrone, *A List of the Jobs That Are Most (and Least) Likely to Be Replaced by Robots*, LINKEDIN THE LEARNING BLOG (Dec. 6, 2017), https://learning.linkedin.com/blog/future-skills/a-list-of-the-jobs-that-are-most—and-least—likely-to-be-replac.

226. *See* 81 Fed. Reg. 35449 (June 2, 2016) (codified at 45 C.F.R. § 95) (addressing state comments about increased data-sharing burdens due to new federal CCWIS regulations).

manages the existing federally-required child welfare datasets.[227] The U.S. Data Federation program, an initiative of the federal GSA Technology Transformation Services 10X program, has indicated it is interested in addressing the need for better federal data collection, and state child welfare leaders and HHS should specifically reach out to participate in this effort.[228] Congress should provide additional funding to address this issue.

In addition, for unique situations not covered by the general research access requirement discussed in this article, HHS could also issue a sample data-sharing agreement for states to use and indicate that this would satisfy all federal requirements, similar to how it has provided sample HIPAA business associate agreements.[229] Congress could also providing funding to support state child welfare agencies' efforts to develop local data-sharing protocols, particularly with courts.

C. States Should Consider Leveraging Volunteers to Assist with Technology Transformation.

Just as states have integrated Court Appointed Special Advocates (CASAs), a cohort of trained volunteers who seek to provide extra support for foster youth, into the child welfare system, states should explore leveraging private sector technology workers' interest in volunteering. 2020 data from LinkedIn demonstrates a boom in volunteering, as workers added about 110,000 volunteer experiences and activities to their profiles each month—with projects relating to children's well-being leading the pack.[230] State child welfare agencies may want to explore partnering with technology companies to create volunteer programs (with appropriate background checks and security protocols) where trained technology professionals can volunteer to assist with data entry and other projects needed to prepare state agencies to adopt modern technology. This volunteer tech corps could help reduce agency costs while also giving tech workers a way to use their skills to benefit foster youth.

V. PRIVACY CONSIDERATIONS SURROUNDING THE SHARING OF DE-IDENTIFIED CHILD WELFARE DATA SETS

Changing federal law to require greater research access to child welfare data will likely face significant opposition from privacy advocates, the defense bar, and child advocacy organizations unless the government is able to address concerns about preserving an appropriate level of data privacy for individuals

227. *See* NAT'L DATA ARCHIVE ON CHILD ABUSE & NEGLECT, https://www.ndacan.acf.hhs.gov (last visited Oct. 21, 2020) (providing datasets of child maltreatment for scholarly use for community research and secondary analysis).

228. Lindpaintner, *supra* note 215.

229. *Business Associate Contracts: Sample Business Associate Agreement Provisions*, U.S. DEP'T. OF HEALTH & HUM. SERVS. (Jan. 25, 2013), https://www.hhs.gov/hipaa/for-professionals/covered-entities/sample-business-associate-agreement-provisions/index.html.

230. George Anders, *A Rare Bright Spot in Pandemic Life: Volunteers Rush to Provide Support*, LINKEDIN (Aug. 26, 2020), https://www.linkedin.com/pulse/rare-bright-spot-pandemic-life-volunteers-rush-provide-george-anders/?trackingId=FSQEPaiBQb%2BEqVvd1bhLGw%3D%3D.

whose data is stored on child welfare systems. The core privacy risk in increasing research access to high quality child welfare data is that bad actors could re-identify an individual from a released data set, thereby subjecting the vulnerable population served by these systems to stigma and even physical safety concerns.[231] For example, researchers found in a 2019 study that 99.98% of Americans would be correctly re-identified in any dataset using fifteen demographic attributes.[232] The paper elaborates on the "release and forget model" by describing several high profile examples of governments and companies that released to the public what they thought were de-identified, selective versions of data sets, but that ultimately turned out to inadvertently reveal the identities of individuals.[233]

Renter screening reports are one example of the real-world repercussions when AI applications utilize publicly available government data in a troubling way. Prospective tenants applying for housing are almost universally required to submit to background screenings performed by various private sector companies, which run automated searches based on partial names or incomplete dates of birth.[234] An algorithm reviews government data such as sex offender registries, criminal and housing court records, and terrorism lists, and produces a report that no human typically reviews before it is sent to the property owner.[235] A May 2020 investigative report uncovered the extensive problem of thousands of individuals in need of housing who lost out on apartments because of erroneous screening reports incorrectly labeling them as criminals.[236]

Addressing these types of legitimate privacy concerns that could result from expanding research access to child welfare data requires a multi-pronged approach. Policymakers should consider the use of innovative data techniques that minimize the ability to re-identify datasets, legislative measures to protect privacy, as well as a rethinking of whether there should be a reduced expectation of privacy in certain aspects of government child welfare data. Importantly, privacy concerns about increased research access to de-identified data must also be considered alongside the reality that there is significant privacy risk in maintaining the status quo and leaving siloed child welfare data to reside in aging SACWIS systems. For example, many SACWIS systems cannot easily restrict different users' access to different aspects of the database even if data is not shared outside the agency, making internal breaches more likely.[237] And the

231. *See* Luc Rocher et al., *Estimating the Success of Re-Identification in Incomplete Datasets Using Generative Models*, 10 NATURE COMMS. 1 (2019) (noting that in the wrong hands, sensitive data can be exploited for blackmailing, mass surveillance, social engineering, or identity theft.); Ian Lundberg et al., *Privacy, Ethics, and Data Access: A Case Study of the Fragile Families Challenge*, 5 SOCIUS 11-12 (2019) https://journals.sagepub.com/doi/pdf/10.1177/2378023118813023 (discussing most likely threats involved in releasing high quality social services poverty data to researchers).

232. *Id.*

233. *Id.*

234. Lauren Kirchner & Matthew Goldstein, *Access Denied: Faulty Automated Background Checks Freeze Out Renters*, THE MARKUP (May 28, 2020), https://themarkup.org/locked-out/2020/05/28/access-denied-faulty-automated-background-checks-freeze-out-renters.

235. *Id.*

236. *Id.*

237. *See* Noah Johnson et al., *Towards Practical Differential Privacy for SQL Queries*, 11 PROCEEDINGS OF THE VLDB ENDOWMENT 526 (2018) ("As demonstrated by recent insider attacks. . . allowing members of an

dated technology infrastructure of SACWIS systems leaves them highly vulnerable to attacks.[238] In light of this reality, federal and state investment in improved technology—and in mechanisms to ensure appropriate implementation—is not optional.

A. Cutting-Edge Data Techniques can Help Protect Highly Sensitive Datasets.

Several emerging data security techniques used by the government and the private sector in other contexts could be applied to child welfare system data. First, differential privacy, a statistical technique which the U.S. government is considering using in connection with the release of 2020 census data,[239] helps protect privacy by injecting statistical "noise" into the data sets to make it increasingly difficult to re-identify any individual data included.[240] While the technique remains popular in the private sector,[241] critics argue that, with respect to the U.S. government, "noisier census data would have a major ripple effect, degrading the quality of many other surveys that rely on census data to select their samples."[242]

Another promising technique is homomorphic encryption, which involves encrypting data so it can be manipulated by an external researcher without allowing that researcher the ability to read the original data.[243] Researchers have studied its use in analyzing health data in the cloud without giving the cloud access to the unencrypted source data.[244] Yet another creative approach involves training an AI program on real, personally identifiable information, then using the AI to generate "synthetic data sets"—statistically identical data points that cannot be traced to any individual, but allow researchers to conduct their analyses on highly similar datasets.[245] This technique is used in the private

organization unrestricted access to data is a major cause of privacy breaches. Access control policies can limit access to a particular database, but once an analyst has access, these policies cannot control how the data is used.").

238. *See* 2018 GOVERNMENT CYBERSECURITY REPORT, SECURITY SCORECARD, https://explore.securityscorecard.com/rs/797-BFK-857/images/2018%20Government%20Cybersecurity%20Report.pdf (noting that in nearly 60 percent of security incidents, it takes the government years to discover a breach).

239. Jeffrey Mervis, *Can a Set of Equations Keep U.S. Census Data Private?*, SCIENCE MAG. (Jan. 4, 2019).

240. *Id.*; *see also Differential Privacy*, HARVARD UNIV. PRIVACY TOOLS PROJECT, https://privacytools.seas.harvard.edu/differential-privacy (last visited Oct. 21, 2019) (explaining that "a crucial feature of differential privacy is that it defines privacy not as a binary notion of 'was the data of individual exposed or not,' but rather a matter of accumulative risk. That is, every time a person's data is processed her risk of being exposed increases. To this end, the definition of differential privacy is equipped with parameters. . . that quantify the 'privacy loss' – the additional risk to an individual that results from her data being used.").

241. *See, e.g.*, Noah Johnson et al., *supra* note 238 at 526 (discussing elastic sensitivity approach to differential privacy adopted by Uber for internal data analytics).

242. Mervis, *supra* note 240.

243. Alex Hern, *'Anonymised' Data Can Never Be Totally Anonymous, Says Study*, THE GUARDIAN (July 23, 2019), https://www.theguardian.com/technology/2019/jul/23/anonymised-data-never-be-anonymous-enough-study-finds.

244. Joppe W.Bos, Kristin Lauter & Michael Naehrig, *Private Predictive Analysis on Encrypted Medical Data*, 50 J. BIOMEDICAL INFORMATICS 234, 240 (2014).

245. *Id.*; *Pros and Cons of Machine Learning Algorithms with Fake Data*, INGEDATA (Aug. 10, 2018), https://www.ingedata.net/blog/machine-learning-algorithms-fake-data.

sector for applications as diverse as improving financial portfolio analysis and risk management, to understanding why truckers cancel scheduled pickups.[246] In addition, private sector companies are experimenting with blockchain applications that utilize algorithms and artificial intelligence models to "travel" to securely stored data, use it for training purposes, and then leave—without ever needing to transfer the sensitive source data.[247]

In a move that helped avoid more invasive government approaches to contact tracing, Apple and Google worked collaboratively in the spring of 2020 on a privacy-focused phone application to assist governments trying to minimize COVID-19 outbreaks.[248] The resulting app uses Bluetooth low energy signals to store contact event information in a decentralized way, without a centralized repository vulnerable to hacks, while also encrypting metadata associated with smartphones and randomly generating phone identification keys.[249] That collaboration demonstrates how proactively designing a public health technology and research effort around the express goal of minimizing privacy risks can contribute to the success of the overall project, as users are more willing to voluntarily participate.

Even with cutting-edge statistical and technical approaches, data-sharing across state lines and across agencies still holds risk. However, that risk must be weighed against the cost of states continuing to operate large, costly systems that are ineffective in helping children and families. The promise of increased data-sharing is the potential to discover more effective ways of investing public funds to help families and reduce suffering.

B. The CARES Act Modifications to HIPAA Requirements can Potentially Serve as a Model for Child Welfare.

In light of the notion of viewing child welfare through a public health lens, it is worth exploring how the recent CARES Act changes to HIPAA requirements could be applied in the child welfare context to address individual privacy interests. Before the coronavirus crisis prompted passage of the 2020 CARES Act, HIPAA required a substance use disorder patient to provide written consent each time a health care entity wanted to share that patient's medical information for purposes of treatment, payment, or health care operations.[250] Now, under the CARES Act, a covered health care program that receives health information pursuant to the patient's written consent may use or redisclose that

---

246. Jin Yuan & Xianghui Yuan, *A Monte Carlo Synthetic Sample Based Performance Evaluation Method for Covariance Matrix Estimators*, APPLIED ECONOMICS LETTERS, (Mar. 2020); *Data Science in the Trucking Industry*, DATACAMP (May 7, 2018), https://www.datacamp.com/community/blog/data-science-industry-trucking.

247. Luke Christou, *Ocean Protocol: Blockchain-Enabled Data Sharing for Better Artificial Intelligence*, VERDICT AI (May 2019), https://verdict-ai.nridigital.com/verdict_ai_may19/ocean_protocol_blockchain_data.

248. Khari Johnson, *Apple and Google Build More Privacy and Flexibility into Bluetooth Contact Tracing Tech*, VENTUREBEAT (Apr. 24, 2020), https://venturebeat.com/2020/04/24/apple-and-google-build-more-privacy-and-flexibility-into-bluetooth-contact-tracing-tech/.

249. *Id.*

250. Michaela Poizner, *Coronavirus: CARES Act Takes Significant Step Toward Modernizing Part 2*, JDSUPRA (Apr. 3, 2020), https://www.jdsupra.com/legalnews/coronavirus-cares-act-takes-significant-94874/.

information for purposes of treatment, payment, or healthcare operations without additional consent by the patient.[251] That consent remains valid until the patient revokes it in writing, which allows the information to be shared with other treatment providers—as long as the initial disclosing provider keeps a log of the disclosures, which the patient has the right to review.[252]

The Act combines this new practical measure with a powerful anti-discrimination clause that prohibits any entities receiving the information, whether intentionally or inadvertently, from discriminating against an individual on the basis of the information in:

- Admission, access to, or treatment for health care (42 U.S.C. § 290dd-2(i)(1)(A))
- Hiring, firing, or terms of employment or receipt of workers' compensation (42 U.S.C. § 290dd-2(i)(1)(B))
- The sale, rental, or continued rental of housing (42 U.S.C. § 290dd-2(i)(1)(C))
- Access to federal, state, or local courts (42 U.S.C. § 290dd-2(i)(1)(D))
- Access to, approval of, or maintenance of social services and benefits provided or funded by federal, state, or local governments (42 U.S.C. § 290dd-2(i)(1)(E)).[253]
- Providing services paid for by the federal funds (42 U.S.C. § 290dd-2(i)(2)).

This concept of having the disclosing entity maintain a log of the disclosures of this specific set of medical data, which can be reviewed by the subject in one place, combined with strong anti-discrimination provisions, could be an interesting starting point for child welfare technology transformation legislation. The individuals whose information is held by child welfare systems do not currently have a single place where they can understand how their information is being used and disclosed. Even without the ability of an individual to opt-in to disclosure for the research purposes discussed in this article, having strong anti-discrimination laws and a tracking function could help avoid the negative outcomes of increased research access to de-identified child welfare data.

A related safeguard that should be considered is creating a private right of action to sue over misuse of individual data. For example, the 2019 New York Privacy Act bill, which ultimately failed, included such a provision; a similar provision was removed from the legislation that ultimately passed as California's Consumer Privacy Act.[254]

251. *Id.*
252. *Id.*; 42 U.S.C. § 290dd-2(b)(1)(C) (2018).
253. Poizner, *supra* note 251; 42 U.S.C. § 290dd-2(i)(1) (2018).
254. New York Privacy Act, S.B. 5642 (2019); *see also* Issie Lapowsky, *New York's Privacy Bill is Even Bolder Than California's*, WIRED (June 4, 2019 7:00 AM), https://www.wired.com/story/new-york-privacy-act-bolder (comparing New York and California privacy acts). Notably, the New York privacy bill was not drafted to apply to state and local governments. *See* S.B. 5642 § 1101(2)(A) ("This article does not apply to state and local governments.").

C. The Federal Government Could Form an Advisory Committee Specifically to Study and Recommend Legislative Changes Addressing Privacy Implications of Technology Transformation in Child Welfare.

A critical first step to exploring the increased data-sharing suggested in this article would be for federal policymakers to form a committee to explore how lessons learned from these other legislative approaches to regulating individual privacy may apply to child welfare data. In addition to laying out the ideal parameters of child welfare data-sharing and outlining a path for funding for a model CCWIS system, the committee should also consider the practical concern of obtaining appropriate cyber insurance coverage for state AI systems. Such insurance policies represent an important way to mitigate harms that could potentially result from integrating AI into child welfare systems.[255] Given that the federal government insures risk for a wide variety of activities (for example, when private insurers were unable to offer affordable terrorism insurance post the events of September 11, 2001, the federal government created a terrorism insurance program[256]), new legislation could contemplate federal support for cyber insurance for child welfare systems.

This committee could also take a holistic look at the information stored by child welfare systems and assess whether it must all be treated exactly the same, or whether it is possible to segment certain types of less stigmatizing data. An underlying reason to mask individual involvement in child welfare cases and receipt of government services is the shame and stigma associated with a system that treats families with a quasi-criminal model.[257] While the public's impression that it is shameful to have any contact with a state child welfare agency may be slow to change, it may nonetheless be possible to approach agency data storage in a way that reflects the level of privacy individuals expect regarding different operations of the child welfare system.

For example, records of a family's attendance at government-sponsored family art classes, parenting classes, or library tutoring seem unlikely to be viewed as highly embarrassing even if they suddenly became publicly available due to a data breach. FFPSA funding is likely to result in state child welfare agencies making these types of preventive services more widely accessible to families even before the state initiates a formal child welfare case.[258] Child welfare system participants, privacy advocates, state technology leaders, and proponents of greater research access to data should collectively brainstorm

255. *See* Ram Shankar et al., *The Case for AI Insurance*, HARVARD BUS. REV. (Apr. 29, 2020), https://hbr.org/2020/04/the-case-for-ai-insurance (describing risks associated with AI systems must be understood and mitigated before AI is fully integrated into an organization's value creation process).

256. Terrorism Risk Insurance Act of 2002, H.R. 3210, 107th Cong. (2002).

257. *See, e.g.*, Jennifer Sykes, *Negotiating Stigma: Understanding Mothers' Responses to Accusations of Child Neglect*, 33 CHILD. & YOUTH SERVS. REV. *448, 451* (2011) (discussing mothers' responses to accusations of child neglect and its stigma); Poppi O'Donnell, *Stigma Associated with Youth in the Foster Care System*, THE IMPRINT (Aug. 5, 2016), https://chronicleofsocialchange.org/child-welfare-2/stigma-associated-youth-foster-care-system/20086 (discussing the stigma associated with the foster care system).

258. *See Family First Legislation*, NAT'L CONF. OF STATE LEGISLATURES (July 23, 2020), http://www.ncsl.org/research/human-services/family-first-updates-and-new-legislation.aspx (providing a list of states' Family First legislations).

about ways to segment data storage so that states maintain the most stigmatizing child welfare data separately and subject it to different restrictions for research access than arguably less sensitive data about families. Federal law does not currently allow for any such distinctions. These discussions should result in legislation that takes a modern approach to balancing the benefits that can be achieved with greater research access to child welfare data with the privacy concerns of the individuals most affected by the disclosures.

CONCLUSION

The debate around the use of technology in child welfare should be reframed to not only focus on the financial and privacy costs of implementing emerging technology, but also the costs of *not* implementing it. There is a very real cost to not understanding your own data. Child welfare systems affect fundamental human rights, as the "interest of parents in the care, custody, and control of their children is perhaps the oldest of the fundamental liberty interests recognized" repeatedly by the U.S. Supreme Court.[259] Child welfare operations directly impact family bonds for hundreds of thousands of vulnerable children. This article implicitly argues that in order for these operations to be ethical, they must be data-informed, and there must be modern data evaluation mechanisms in place to hold the system accountable for its outcomes. The current status of child welfare data-sharing and analysis fails to meet this basic standard. Indeed, a key, understudied question raised by those who argue for rethinking child welfare systems is whether the damage caused by removing children from an allegedly abusive home and family is more damaging to the child than doing nothing at all, and avoiding contact with the child welfare system.[260] We must build a data and technology infrastructure that allows us to answer this question.

The federal government is uniquely positioned to issue a national data-sharing requirement tied to federal funding so that verified, vetted researchers can access broader data sets that fuel innovation in helping children and families. Child welfare leaders at both the federal and state level should prioritize adoption of technology tools and databases that can lift the administrative burden off the shoulders of the child welfare workforce, freeing them to spend the majority of their time interacting on a human level with the individuals that these complex systems are trying to help. Above all, in redesigning what the modern child welfare system of the future might look like, governments must include representation and feedback from the children and families whose lives are at stake.

---

259. Troxel v. Granville, 530 U.S. 57, 65 (2000).

260. *See Shanta Trivedi*, *The Harm of Child Removal*, 43 N.Y.U. REV. L. & SOC. CHANGE 523, 528 (2019) (quoting Nicholson v. Williams, 203 F. Supp. 2d 153, 204 (E.D.N.Y. 2002)).