Appendices

Produced by FUTURE OF PRIVACY FORUM FPF.ORG

PERSONAL DATA and the ORGANIZATION: STEWARDSHIP AND STRATEGY

A VISUAL GUIDE TO PRACTICAL DATA DE-IDENTIFICATION

Produced by FUTURE OF PRIVACY FORUM FPF.ORG
In collaboration with EY

What do scientists, regulators and lawyers mean when they talk about de-identification? How does anonymous data differ from pseudonymous or de-identified information? Data identifiability is not binary. Data lies on a spectrum with multiple shades of identifiability.

This is a primer on how to distinguish different categories of data.

DEGREES OF IDENTIFIABILITY

Information containing direct and indirect identifiers.

PSEUDONYMOUS DATA

Information from which direct identifiers have been eliminated or transformed, but indirect identifiers remain intact.

DE-IDENTIFIED DATA

Direct and known indirect identifiers have been removed or manipulated to break the linkage to real world identities.

ANONYMOUS DATA

Direct and indirect identifiers have been removed or manipulated together with mathematical and technical guarantees to prevent re-identification.

	EXPLICITLY PERSONAL	POTENTIALLY IDENTIFIABLE	NOT READILY IDENTIFIABLE	KEY CODED	PSEUDONYMOUS	PROTECTED PSEUDONYMOUS	DE-IDENTIFIED	PROTECTED DE-IDENTIFIED	ANONYMOUS	AGGREGATED ANONYMOUS
DIRECT IDENTIFIERS Data that identifies a person without additional information or by linking to information in the public domain (e.g., name, SSN)	INTACT	PARTIALLY MASKED	PARTIALLY MASKED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED
INDIRECT IDENTIFIERS Data that identifies an individual indirectly. Helps connect pieces of information until an individual can be singled out (e.g., DOB, gender)	INTACT	INTACT	INTACT	INTACT	INTACT	INTACT	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED
SAFEGUARDS and CONTROLS Technical, organizational and legal controls preventing employees, researchers or other third parties from re-identifying individuals	NOT RELEVANT due to nature of data	LIMITED or NONE IN PLACE	CONTROLS IN PLACE	CONTROLS IN PLACE	LIMITED or NONE IN PLACE	CONTROLS IN PLACE	LIMITED or NONE IN PLACE	CONTROLS IN PLACE	NOT RELEVANT due to nature of data	NOT RELEVANT due to high degree of data aggregation
SELECTED EXAMPLES	Name, address, phone number, SSN, government-issued ID (e.g., Jane Smith, 123 Main Street, 555-555-5555)	Unique device ID, license plate, medical record number, cookie, IP address (e.g., MAC address 68:A8:6D:35:65:03)	Same as Potentially Identifiable except data are also protected by safeguards and controls (e.g., hashed MAC addresses & legal representations)	Clinical or research datasets where only curator retains key (e.g., Jane Smith, diabetes, HgB 15.1 g/dl = Csrk123)	Unique, artificial pseudonyms replace direct identifiers (e.g., HIPAA Limited Datasets, John Doe = 5L7T LX619Z) (unique sequence not used anywhere else)	Same as Pseudonymous, except data are also protected by safeguards and controls	Data are suppressed, generalized, perturbed, swapped, etc. (e.g., GPA: 3.2 = 3.0-3.5, gender: female = gender: male)	Same as De-Identified, except data are also protected by safeguards and controls	For example, noise is calibrated to a data set to hide whether an individual is present or not (differential privacy)	Very highly aggregated data (e.g., statistical data, census data, or population data that 52.6% of Washington, DC residents are women)

Produced by FUTURE OF PRIVACY FORUM FPF.ORG

FINANCIAL DATA LOCALIZATION: CONFLICTS AND CONSEQUENCES

Modern banking customers are global, and expect on-demand, high-quality service from their financial institutions regardless of time or location, making 24/7 call centers and multi-national bank branches and service centers the norm. Similarly, regulators expect financial institutions to have a global understanding of their customers to assess and manage risk. Delivering on these expectations requires financial institutions to regularly move information between locations in support of business operations. Policy goals to ensure privacy and security are important and can coexist with the free flow of data. However, regulations that achieve these goals through localization cause conflict and complexity and can result in unintended consequences. Let's take a look:

HOW DATA FLOWS

Supporting business operations requires the regular movement of financial data between locations. Multi-national operations add complexity, as local governing regulations must be considered once a border is crossed.

UNINTENDED CONSEQUENCES OF LOCALIZATION

LEGAL TENSION

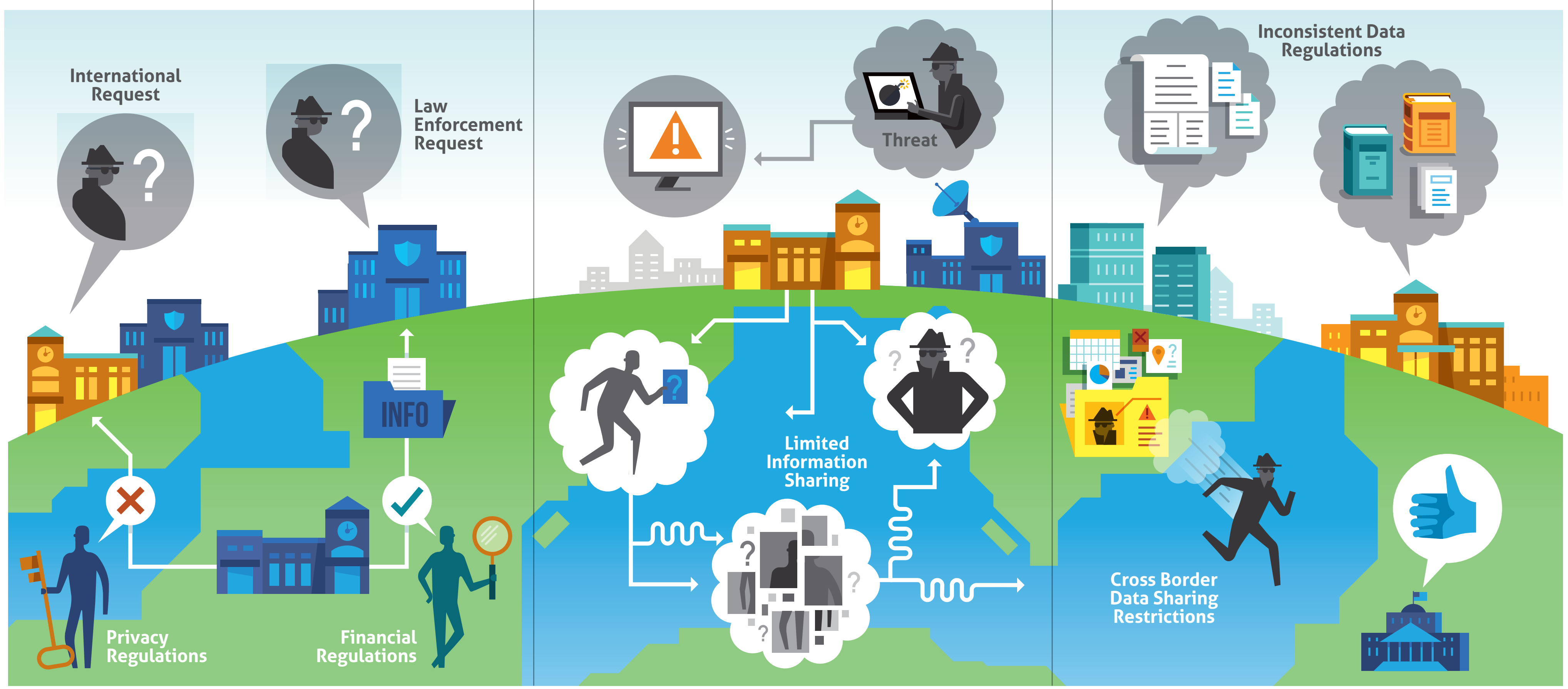
Banks have legal obligations to comply with both regulators and law enforcement agencies within their country. However, requests by law enforcement from other countries for access to data, even when narrowly tailored and proportionate, can often conflict with local regulations that seek to protect the privacy of citizens and the integrity of the financial system. These tensions are heightened by a lack of international, agreed upon principles or safe harbors.

HAMPERED THREAT RESPONSE

Data privacy and other cross border data transfer restrictions may limit the ability to share information from one country with peers and regulators in other countries so security threats may be slower to be identified. A legislative framework is needed for sharing threat information across borders while respecting local privacy and other rules.

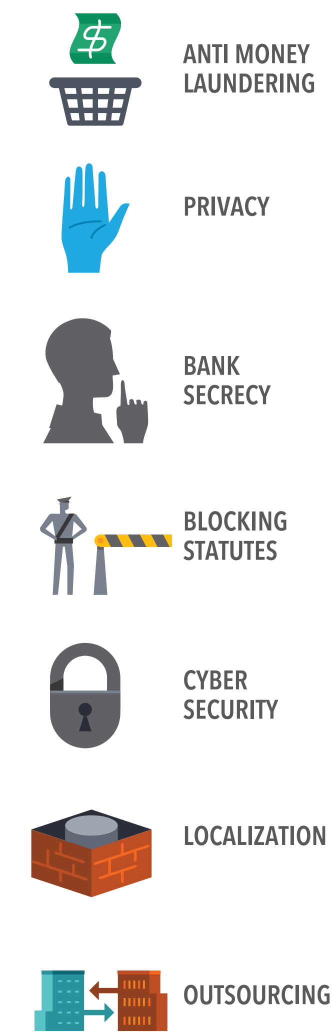
COMPROMISED REPORTING CAPABILITIES

The inconsistency of data regulations across countries erodes the opportunity for holistic reporting. For example, when considering criminal activity, regulations require criminal reports to be made locally. In addition, cross border data sharing restrictions often apply to sharing with affiliates, which increases the risk that a criminal rejected in one country can open an account in another country.



TYPES OF REGULATION

Many regulations exist to control access to information and protect privacy and security interests:



PERCEIVED DRIVERS FOR LOCALIZATION

INFORMATION SECURITY

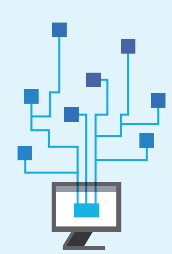
Perception: Localization provides better data security and protection.

Reality: Increased risk of cyber attacks as footprint grows and data becomes more diffuse.

PROTECTION OF PRIVACY VALUES

Perception: Localization protects data from over-broad law enforcement access abroad.

Reality: With narrowly tailored and proportionate laws we can accomplish better oversight and protect individual privacy.



TECHNOLOGY

Perception: Localization makes technology easier to manage.

Reality: It's more difficult to update applications and ensure consistency with increased end-points.

EFFICIENCY

Perception: Localization increases efficiency.

Reality: Redundancy of data centers and personnel reduces bank efficiency and increases cost.

LOCAL JOBS

Perception: Localization creates jobs and stimulates the economy.

Reality: Job creation is minimal, and localization can cause global financial companies to reduce their presence, limiting services and opportunities.

ACCESS TO DATA

Perception: Localization is the only way to ensure access to data during a crisis.

Reality: Contractual access can be granted to data stored outside a local jurisdiction to ensure regulators can perform regulatory and supervisory roles, even during a crisis.

Draft Legislative Definition of "Covered Data"
Future of Privacy Forum
last revised April 24, 2019

(1) In this Act, "Covered Data" means any data that: 1) is under the control of a Covered Entity; and 2) is linked or can practicably be linked to an individual by the Covered Entity or by an anticipated recipient of the data.

(2) "Covered Data" includes

a) "Identified Data" - information explicitly linked to a known individual.

b) "Identifiable Data" - information that is not explicitly linked to a known individual, but that can practicably be linked by the Covered Entity or intended recipients. [is not subject to access requests/portability etc. but is subject to all other restrictions]

c) "Pseudonymous Data" - information that cannot be linked to a known individual without additional information kept separately;

c) "De-Identified Data" - (i) data from which direct and indirect identifiers have been permanently removed; or (ii) data that has been perturbed to the degree that risk of re-identification is small, given the context of the data set. (iii) data that an expert has confirmed poses a very small risk that information can be used by an anticipated recipient to identify an individual

[Key substantive requirements for data to be classified as "de-identified data:"

- When subject to controls that are legal, administrative, technical, contractual, enforceable (public commitment/FTC), or some combination of such controls, the data is not subject to many requirements.
- The data cannot be made public.
- The data cannot be shared without controls that reasonably prevent identification. by anticipated recipients.
- The data is not subject to access/portability.
- Such de-identification is a **determinative** factor in assessing whether a use is "incompatible/out of context/subject to consent requirements" under a federal privacy law's substantive provisions.
- In many circumstances, the Act imposes different requirements regarding Identified Data and De-Identified Data; the Act incentivizes Covered Entities to de-identify Identified Data when appropriate.]

[Key impacts of data being classified as "pseudonymous data"

- Data cannot be made public
- Data cannot be shared without controls that reasonably prevent identification by anticipated recipients.
- Pseudonymization is an **important but not determinative** factor in assessing whether a use is "incompatible/out of context/subject to consent requirements" under a federal privacy law's substantive provisions.
- When pseudonymous data is shared and used by 3rd parties for personalization, targeting, profiling - the right to opt-out is applicable, unless the data is only used

in aggregate form (for analysis, research, ad reporting.) *This seeks to capture ad tech current self-regulatory practices.*

    - Important point: Data that has been pseudonymized, but for which a key is not available, or for which assurances are in place that prevent intended recipients from identifying users under c(iii) can be deidentified data.
- Access/portability requirements depend on technical feasibility

(3) Exceptions - The term “Covered Data” does not include:

(a) Publicly available information. “Publicly available” means information that is lawfully made available from federal, state, or local government records when that information is used for a purpose that is compatible with the purpose for which the data is maintained and made available in the government records.

(b) Data used by an employer solely in connection with an employee’s employment or post employment related status (retirement etc);

(c) Data used by a business in the context of business-to-business activities;

(d) Data deleted by a Covered Entity;

(e) Non-identifiable Data, which has been strongly de-identified (direct and indirect identifiers have been removed, or data has been significantly perturbed or highly aggregated and an expert assessment assures the data can be made public, shared (or shared a limited number of times) and presents no risk or very little privacy risk; and

(f) Data used to identify or mitigate cybersecurity threats; ensure the security and stability of a Covered Entity’s networks and/or physical infrastructure; or operate anti-fraud programs;

(g) Data used to prevent or detect criminal activity or child exploitation;

(h) Data used to comply with a legal requirement;

(i) Data regarding a deceased individual that does not reveal Covered Data regarding a living individual [e.g. genetic data].

(4) Section [t/k] of the Act authorizes a mechanism [t/k] by which [t/k] can revise or supplement the definition of “Covered Data” through [t/k administrative mechanism].

Draft Legislative Treatment of Consumer Control Provision

(1) Opt-Out Consent for Non-Sensitive Information

a) Individuals must be provided with clear, conspicuous, and readily available mechanisms to exercise choice. A covered entity must offer individuals the opportunity to choose (opt out) whether their data is used in the following ways:

(i) Identified and/or Identifiable Data is disclosed to a third party;

(ii) Identified, Identifiable Data or Pseudonymous data is used for a purpose that is materially different from the purpose(s) for which it was originally collected or subsequently authorized by the individuals;

- [De-identification is a **determinative** factor in assessing whether a use is materially different
- Pseudonymization is a **significant** factor in assessing whether a use is materially different.]

(iii) Identified, Identifiable Data or Pseudonymous data is used to contact, market to, or personalize services for an individual; or

(iv) Covered data is used in an incompatible manner that poses an unjustified harm to individuals.

- [Safeguards can mitigate potential harms.
- Benefits to the user can be considered to assess whether a harm is justified.
- Benefits to others (to other users, to research, to a community) can be considered.
- A user's reasonable expectation of how data is to be used is a significant factor. A user's reasonable expectation can be determined from factors including: prominence of disclosures; research and polling; and practices that are identified by self-regulatory processes.] [*key issue: defining "research"*]

(2) Exceptions - a covered entity is not required to offer individuals the opportunity to opt out if:

a) A third party that is acting as an agent to perform tasks on behalf of and under the instructions of the Covered Entity and the Covered Entity has entered into a contract with the third party requiring the third party to comply with the Covered Entity's commitments to individuals; or

b) Covered Data is used solely for analytics or research purposes.

(3) Opt-In Consent for Sensitive Information

A) For sensitive information (i.e., personal information specifying medical or health conditions, racial or ethnic origin, [political opinions], religious or philosophical beliefs, or information specifying the sex life of the individual), a covered entity must obtain affirmative express consent (opt in) from individuals if such data is to be used in the following ways:

i) Sensitive Identified and/or Identifiable Data is disclosed to a third party; or

ii) Sensitive Identified, Identifiable or Pseudonymous data is used for a purpose that is materially different from the purpose(s) for which it was originally collected or subsequently authorized by the individuals; or

- [De-identification is a **determinative** factor in assessing whether a use is materially different. Sensitivity of the underlying data is a risk factor that must be taken into account when assessing the risk of de-identification.
- Pseudonymization is a factor in assessing whether a use is materially different. (not a significant factor)]

(iv) Sensitive Covered Data is used in a way that poses an unjustified harm to individuals.

- [Safeguards can mitigate potential harms.
- Benefits to the user can be considered to assess whether a harm is justified.
- Benefits to others (to other users, to research, to a community) can be considered, but only by an independent ethical review process.
- A user's reasonable expectation of how data is to be used is a significant factor. A user's reasonable expectation can be determined from factors including: prominence of disclosures; research and polling; and practices that are identified by self regulatory processes, but only when those self regulatory processes meet [XYZ] standard.

(4) Exceptions - a covered entity is not required to obtain opt-in consent from individuals if:

a) A third party that is acting as an agent to perform tasks on behalf of and under the instructions of the Covered Entity and the Covered Entity has entered into a contract with the third party requiring the third party to comply with the Covered Entity's commitments to individuals; or

b) Covered Data is used solely for analytics or research purposes.

UNFAIRNESS BY ALGORITHM:

DISTILLING THE HARMS OF AUTOMATED DECISION-MAKING

December 2017

FUTURE OF PRIVACY FORUM

Overview

Analysis of personal data can be used to improve services, advance research, and combat discrimination. However, such analysis can also create valid concerns about differential treatment of individuals or harmful impacts on vulnerable communities. These concerns can be amplified when automated decision-making uses sensitive data (such as race, gender, or familial status), impacts protected classes, or affects individuals' eligibility for housing, employment, or other core services. When seeking to identify harms, it is important to appreciate the context of interactions between individuals, companies, and governments—including the benefits provided by automated decision-making frameworks, and the fallibility of human decision-making.

Recent discussions have highlighted legal and ethical issues raised by the use of sensitive data for hiring, policing, benefits determinations, marketing, and other purposes. These conversations can become mired in definitional challenges that make progress towards solutions difficult. There are few easy ways to navigate these issues, but if stakeholders hold frank discussions, we can do more to promote fairness, encourage responsible data use, and combat discrimination.

To facilitate these discussions, the Future of Privacy Forum (FPF) attempted to identify, articulate, and categorize the types of harm that may result from automated decision-making. To inform this effort, FPF reviewed leading books, articles, and advocacy pieces on the topic of algorithmic discrimination. We distilled both the harms and potential mitigation strategies identified in the literature into two charts. We hope you will suggest revisions, identify challenges, and help improve the document by contacting lsmith@fpf.org. In addition to presenting this document for consideration for the FTC Informational Injury workshop, we anticipate it will be useful in assessing fairness, transparency and accountability for artificial intelligence, as well as methodologies to assess impacts on rights and freedoms under the EU General Data Protection Regulation.

The Chart of Potential Harms from Automated Decision-Making

This chart groups the harms identified in the literature into four broad "buckets"—loss of opportunity, economic loss, social detriment, and loss of liberty—to depict the various spheres of life where automated decision-making can cause injury. It also notes whether each harm manifests for individuals or collectives, and as illegal or simply unfair.

We hope that by identifying and categorizing the harms, we can begin a process that will empower those seeking solutions to mitigate these harms. We believe that a more clear articulation of harms will help focus attention and energy on potential mitigation strategies that can reduce the risks of algorithmic discrimination. We attempted to include all harms articulated in the literature in this chart; we do not presume to establish which harms pose greater or lesser risks to individuals or society.

The Chart of Potential Mitigation Sets

This chart uses FPF's taxonomy to further categorize harms into groups that are sufficiently similar to each other that they could be amenable to the same mitigation strategies.

Attempts to solve or prevent this broad swath of harms will require a range of tools and perspectives. Such attempts benefit by further categorization of the identified harms, into five groups of similar harms. These groups include: (1) individual harms that are illegal; (2) individual harms that are simply unfair, but have a corresponding illegal analog; (3) collective/societal harms that have a corresponding individual illegal analog; (4) individual harms that are unfair and lack a corresponding illegal analog; and (5) collective/societal harms that lack a corresponding individual illegal analog. The chart includes a description of the mitigation strategies that are best positioned to address each group of harms.

There is ample debate about whether the lawful decisions included in this chart are fair, unfair, ethical, or unethical. Absent societal consensus, these harms may not be ripe for legal remedies.

Potential Harms from Automated Decision-Making

Individual Harms: Illegal	Individual Harms: Unfair	Collective / Societal Harms
Loss of Opportunity		
Employment Discrimination E.g. Filtering job candidates by race or genetic/health information	**Employment Discrimination** E.g. Filtering candidates by work proximity leads to excluding minorities	**Differential Access to Job Opportunities**
Insurance & Social Benefit Discrimination E.g. Higher termination rate for benefit eligibility by religious group	**Insurance & Social Benefit Discrimination** E.g. Increasing auto insurance prices for night-shift workers	**Differential Access to Insurance & Benefits**
Housing Discrimination E.g. Landlord relies on search results suggesting criminal history by race	**Housing Discrimination** E.g. Matching algorithm less likely to provide suitable housing for minorities	**Differential Access to Housing**
Education Discrimination E.g. Denial of opportunity for a student in a certain ability category	**Education Discrimination** E.g. Presenting only ads on for-profit colleges to low-income individuals	**Differential Access to Education**
Economic Loss		
Credit Discrimination E.g. Denying credit to all residents in specified neighborhoods (“redlining”)	**Credit Discrimination** E.g. Not presenting certain credit offers to members of certain groups	**Differential Access to Credit**
Differential Pricing of Goods and Services E.g. Raising online prices based on membership in a protected class	**Differential Pricing of Goods and Services** E.g. Presenting product discounts based on “ethnic affinity”	**Differential Access to Goods and Services**
	Narrowing of Choice E.g. Presenting ads based solely on past “clicks”	**Narrowing of Choice for Groups**
Social Detriment		
	Network Bubbles E.g. Varied exposure to opportunity or evaluation based on “who you know”	**Filter Bubbles** E.g. Algorithms that promote only familiar news and information
	Dignitary Harms E.g. Emotional distress due to bias or a decision based on incorrect data	**Stereotype Reinforcement** E.g. Assumption that computed decisions are inherently unbiased
	Constraints of Bias E.g. Constrained conceptions of career prospects based on search results	**Confirmation Bias** E.g. All-male image search results for “CEO,” all-female results for “teacher”
Loss of Liberty		
	Constraints of Suspicion E.g. Emotional, dignitary, and social impacts of increased surveillance	**Increased Surveillance** E.g. Use of “predictive policing” to police minority neighborhoods more
Individual Incarceration E.g. Use of “recidivism scores” to determine prison sentence length (legal status uncertain)		**Disproportionate Incarceration** E.g. Incarceration of groups at higher rates based on historic policing data

Potential Mitigation Sets

Harms	Description	Mitigation Tools
Individual Harms – Illegal		
Employment Discrimination **Insurance & Social Benefit Discrimination** **Housing Discrimination** **Education Discrimination** **Credit Discrimination** **Differential Pricing** **Individual Incarceration**	Existing law defines impermissible outcomes, often specifically for protected classes	• **Data methods** to ensure proxies are not used for protected classes & data does not amplify historical bias • **Algorithmic design** to carefully consider whether to use protected status inputs & trigger manual reviews • **Laws & policies** that use data to identify discrimination
Individual Harms – Unfair (with illegal analog)		
Employment Discrimination **Insurance & Social Benefit Discrimination** **Housing Discrimination** **Education Discrimination** **Credit Discrimination** **Differential Pricing** **Individual Incarceration**	Individual harms that could be considered illegal if they involved protected classes, but do not in this case	• **Business processes** to index concerns; ethical frameworks & best practices to monitor & evaluate outcomes • **Laws & policies** include tools like DPIAs to measure impact or enable rights to explanation
Collective/Societal Harms (with illegal analog)		
Differential Access to Job Opportunities **Differential Access to Insurance Benefits** **Differential Access to Housing** **Differential Access to Education** **Differential Access to Credit** **Differential Access to Goods & Services** **Disproportionate Incarceration**	Group level impacts that are not legally prohibited, though related individual impacts could be illegal	• Same as above section • **Laws & policies** should consider offline analogies & whether it is appropriate for industry to identify & mitigate
Individual Harms – Unfair (without illegal analog)		
Narrowing of Choice **Network Bubbles** **Dignitary Harms** **Constraints of Bias** **Constraints of Suspicion**	Individual impacts for which we do not have legal rules. Mitigation may be difficult or undesirable absent a defined set of societal norms	• **Business processes** to index concerns, ethical frameworks & best practices to monitor & evaluate outcomes • **Laws & policies** should consider whether it is appropriate to expect industry to identify & enforce norms
Collective/Societal Harms (without illegal analog)		
Narrowing of Choice for Groups **Filter Bubbles** **Stereotype Reinforcement** **Confirmation Bias** **Increased Surveillance of Groups**	Group level impacts for which we do not have legal rules or societal agreement as to what constitutes a harm	• Same as above section

Key

Loss of Opportunity	**Economic Loss**	**Social Stigmatization**	**Loss of Liberty**

Working Definitions: Harms

Automated Decision: The direct output or indirect result from an automated program analyzing individual or aggregate data. This includes pre-programmed algorithms and those that evolve via machine learning techniques.

Illegal: Examples in this category represent harms that are illegal under several U.S. civil rights laws, which generally protect core classifications—such as race, gender, age, and ability—against discrimination, disparate treatment, and disparate impact.

Unfair: Examples in this category represent actions that are typically legal, but nonetheless trigger notions of unfairness. Like the “illegal” category, some examples here may be differently classified depending on the legal regime.

Collective / Societal Harms: This category represents overall negative effects to society that are chiefly collective, rather than individual in nature.

Loss of Opportunity: This group broadly describes harms occurring within the domains of the workplace, housing, social support systems, healthcare, and education.

Economic Loss: This group broadly describes harms that primarily cause financial injury or discrimination in the marketplace for goods and services.

Social Detriment: This group broadly describes harms to one's sense of self, self worth, or community standing relative to others.

Loss of Liberty: This group broadly describes harms that constrain one’s physical freedom and autonomy.

Working Definitions: Mitigation

Individual Harms – Illegal: The harms in this category are those for which American law defines outcomes that are not legally permissible. These harms typically become legally cognizable because they impact legally protected classes in a manner that is defined as impermissible under existing law. Notably, disparate impact may be relevant to illegality regardless of intent in some areas.

Individual Harms – Unfair (with illegal analog): The individual harms in this category do not involve protected classes, but could be considered illegal if protected classes were implicated. For example, while price discrimination based on race could be illegal under the Fair Credit Reporting Act or Civil Rights Act, price discrimination based on computer operating system of the user is not protected under the law. Nonetheless, automated decision-making enables a growing number of personalized distinctions. Some may consider these distinctions unfair or unethical.

Collective/Societal Harms (with illegal analog): In this category, impacts at the group level may not be legally prohibited, but individual impacts could be illegal under different circumstances. While rules may prohibit disparate treatment of protected classes, differential treatment of groups that are not legally protected may not be considered illegal. For example, systematically failing to hire people of a certain race may be illegal, but systematically failing to hire Apple computer users or Red Sox fans is not protected under the law, though some may consider it unfair.

Individual Harms – Unfair (without illegal analog): This category applies to impacts on individuals for which we do not have legal rules. Some, such as narrowing of choice and network bubbles, may be harms that are newly enabled by the growth of technology platforms. Others, such as the the constraints of bias or the constraints of suspicion, have been challenges in the analog world for decades.

Collective/Societal Harms (without illegal analog): This category includes collective outcomes for which we do not have legal rules. As with the prior group, some of these harms—such as narrowing of choice for groups and filter bubbles—have become more frequent due to increased reliance on algorithmic personalization techniques. Stereotype reinforcement is as old as time, but can be compounded by the volume of information available online. Confirmation bias and increased surveillance of groups have been challenges in society for decades, if not since its inception.

Reviewed Literature

The alphabetized list below captures the literature FPF has reviewed to date for this effort. We welcome suggestions for further materials to review to lsmith@fpf.org.

- Aaron Reike, *Don't let the hype over "social media scores" distract you*, EQUAL FUTURE (2016).
- Alessandro Acquisti & Christina Fong, *An Experiment in Hiring Discrimination via Online Social Network*, presented at Privacy Law Scholars Conference (2016).
- Alethea Lange et al., *A User-Centered Perspective on Algorithmic Personalization*, presented at the Fed. Trade Comm'n PrivacyCon Conference (2017).
- Allan King & Marko Mrkonich, *"Big Data" and the Risk of Employment Discrimination*, 68 OKLA. L. REV. 555 (2016).
- Andrew Tutt, *An FDA for Algorithms*, 67 ADMIN. L. REV. 1 (2016).
- Aniko Hannak et al., *Bias in Online Freelance Marketplaces: Evidence from TaskRabbit*, presented at the Workshop on Data and Algorithmic Transparency (Nov. 2016).
- CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION (2016).
- Christian Sandvig et al., Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms, presented at the Int'l Comm'cn Ass'n Conference on Data and Discrimination: Converting Critical Concerns into Productive Inquiry (2014).
- Daniel Solove, *A Taxonomy of Privacy*, 154 U. PENN. L. REV. 3 (2016).
- Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1 (2014).
- EXEC. OFF. OF THE PRESIDENT, BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES (2014).
- EXEC. OFF. OF THE PRESIDENT, BIG DATA: A REPORT ON ALGORITHMIC SYSTEMS, OPPORTUNITY, AND CIVIL RIGHTS (2016).
- FEDERAL TRADE COMMISSION, BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION? (Jan. 2016).
- Frank Pasquale & Danielle Keats Citron, *Promoting Innovation While Preventing Discrimination: Policy Goals for the Scored Society*, 89 WASH. L. REV. 1413 (2014).
- Jennifer Valentino-Devries, Jeremy Singer-Vine, Ashkan Soltani, *Websites Vary Prices, Deals Based on Users' Information*, WALL ST. J. (Dec. 24, 2012).
- Joshua Kroll et al., *Accountable Algorithms*, 165 U. PENN. L. REV. 633 (2016).
- Juhi Kulshrestha et al., *Quantifying Search Bias: Investigating Sources of Bias for Political Searches in Social Media*, presented at the Workshop on Data and Algorithmic Transparency (2016).
- Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C.L. REV. 93 (2014).
- Latanya Sweeney, *Discrimination in Online Ad Delivery*, COMMC'NS OF THE ASS'N OF COMPUTING MACHINERY (2013).
- Lee Rainie & Jana Anderson, *Code-Dependent: Pros and Cons of the Algorithm Age*, PEW RESEARCH CENTER (2017).
- Mark MacCarthy, *Student Privacy: Harm and Context*, 21 INT'L REV. OF INFO. ETHICS 11 (2014).
- Mary Madden, Michele Gilman, Karen Levy & Alice Marwick, *Privacy , Poverty, and Big Data: A Matrix of Vulnerabilities for Poor Americans*, Wash. U. L. Rev ___ (forthcoming) (Mar. 2017).
- Megan Garcia, *How to Keep Your AI From Turning Into a Racist Monster*, WIRED (2017).
- Moritz Hardt, Eric Price & Nathan Srebro, *Equality of Opportunity in Supervised Learning*, presented at the Conference on Neural Info. Processing Sys. (2016).
- Motahhare Eslami et al., *Reasoning about Invisible Algorithms in the News Feed*, presented at the Ass'n of Computing Machinery Special Interest Gp. on Computer-Human Interaction (2015).
- Muhammad Zafar et al., *Fairness Beyond Disparate Treatment & Disparate Impact: Learning Classification without Disparate Mistreatment*, presented at the Int'l World Wide Web Conference (2017).
- Nanette Byrnes, *Why We Should Expect Algorithms to be Biased*, MIT TECHNOLOGY REVIEW (2016).
- NEW AMERICA & OPEN TECH. INST., DATA AND DISCRIMINATION: COLLECTED ESSAYS (S.P. Gangadharan, Ed. 2014).
- Omer Tene and Jules Polonetsky, *Big Data for All: Privacy and User Control in the Age of Analytics*, 11 Nw. J. Tech. & Intell. Prop.239 (2013).
- PAM DIXON & ROBERT GELLMAN, THE SCORING OF AMERICA: HOW SECRET CONSUMER SCORES THREATEN YOUR PRIVACY AND YOUR FUTURE, WORLD PRIVACY FORUM (2014).
- Pauline Kim, *Data-Driven Discrimination at Work*, 59 WILLIAM & MARY L. REV. ____ (2017).
- Peter Swire, *Lessons From Fair Lending Law for Fair Marketing and Big Data* (2014)
- PROPUBLICA, Machine Bias Investigative Series, https://www.propublica.org/series/machine-bias
- Sandra Wachter, Brent Mittelstadt, & Luciano Floridi, *Why a right to explanation of automated decision making does not exist in the General Data Protection Regulation* (2016).
- Solon Barocas & Andrew Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671 (2016).
- UPTURN, CIVIL RIGHTS, BIG DATA, AND OUR ALGORITHMIC FUTURE (2014).

FUTURE OF PRIVACY FORUM

1400 EYE STREET, NW | SUITE 450 | WASHINGTON, DC 20005 • FPF.ORG

202.768.8950

INFO@FPF.ORG | MEDIA@FPF.ORG

BIG DATA:
A Tool for Fighting Discrimination and Empowering Groups

FUTURE OF PRIVACY FORUM

Anti-Defamation League®

PREFACE

In May 2014, the Executive Office of the President concluded its 90-day study of Big Data and privacy and released a report entitled *Big Data: Seizing Opportunities, Preserving Values.* The report highlighted certain positive uses of Big Data, such as identifying health risks at an early stage, creating efficiencies in energy distribution, and uncovering fraud through predictive analysis. However, it also concluded that Big Data analytics could facilitate discrimination in housing, credit, employment, health, education, and a range of other markets. These potential benefits and drawbacks underscore the need to better understand how Big Data will shape our lives in years to come.

Recognizing the 50th anniversary of the Civil Rights Act and the challenges to fighting discrimination in the 21st century, the case studies included in this report show how businesses, governments, and civil society organizations are leveraging Big Data (and other data sets[1]) to protect and empower vulnerable groups, including by providing access to job markets, uncovering discriminatory practices, and creating new tools to improve education and assist those in need. While by no means an exhaustive list of Big Data's potential to uncover and fight discrimination, we offer these examples to show how Big Data already is redefining efforts to ensure equal opportunity for all.

We would like to thank the Anti-Defamation League for its partnership in preparing this report, Jared Bomberg and Julian Flamant at Hogan Lovells US LLP for providing essential research and drafting support, and members of the FPF Advisory Board for reviewing drafts and providing guidance.

We hope these examples will contribute to discussions about Big Data's impact on discrimination.

Jules Polonetsky
Executive Director and Co-Chair
Future of Privacy Forum

Christopher Wolf
Founder and Co-Chair
Future of Privacy Forum

[1] Some of the datasets being used by businesses, governments, and civil society organization will be considered by some to be more appropriately classified as "small data" as they are built, in some cases, from fixed and pre-existing datasets or rely on limited data inputs. We ask that our readers recognize the value of evolving uses and usefulness of data as exposed by these cases and imagine how that value will be compounded as applications of Big Data catch up to technological capabilities.

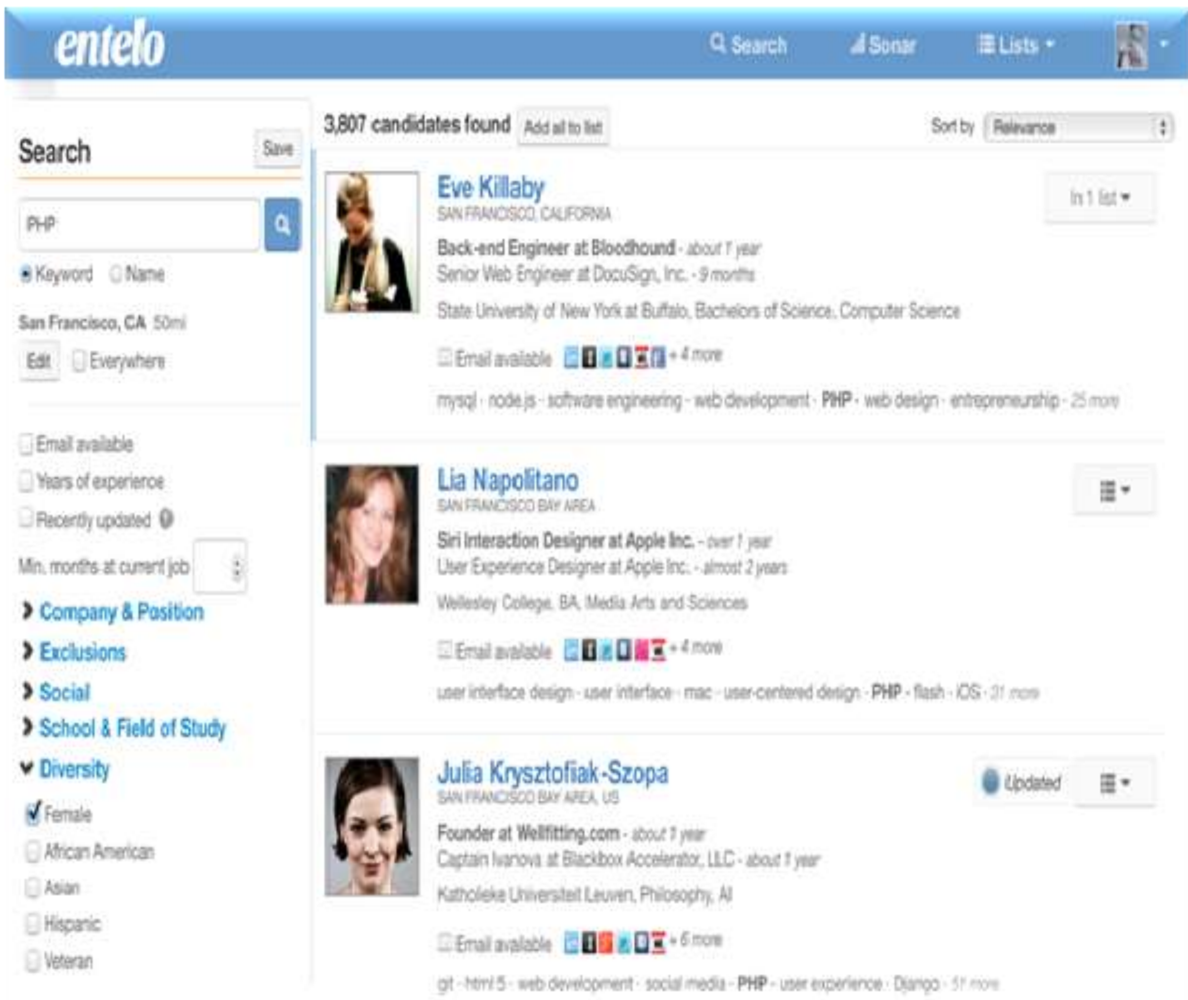
TABLE OF CONTENTS

I. SEEING BEYOND BIAS TO PROVIDE NEW OPPORTUNITIES

Case Study 1: Workplace Diversity (Entelo)

Entelo Diversity, a candidate recruiting platform launched in April 2014, is improving workplace diversity by empowering recruiters to search for job candidates from within underrepresented segments of the population. Using a proprietary algorithm, this workplace diversity tool sifts through publicly available data—pulled from social media platforms—to match recruiters with candidates who hold necessary qualifications, but also meet particular diversity requirements. The tool can filter candidates based on gender, race, and military history in five categories: Female, African American, Asian, Hispanic, and Veteran.

Example of an Entelo Search



Source: http://blog.entelo.com/company-news/announcing-entelo-diversity

Case Study 2: Opportunity for Advancement (Google)

A challenge for the technology industry is ensuring diversity in the workplace. Twitter has recently reported that 90% of its global “tech” employees are male and Google admits “[it’s] not where [it] wants to be when it comes to diversity,” with only 17% women among its tech workforce.

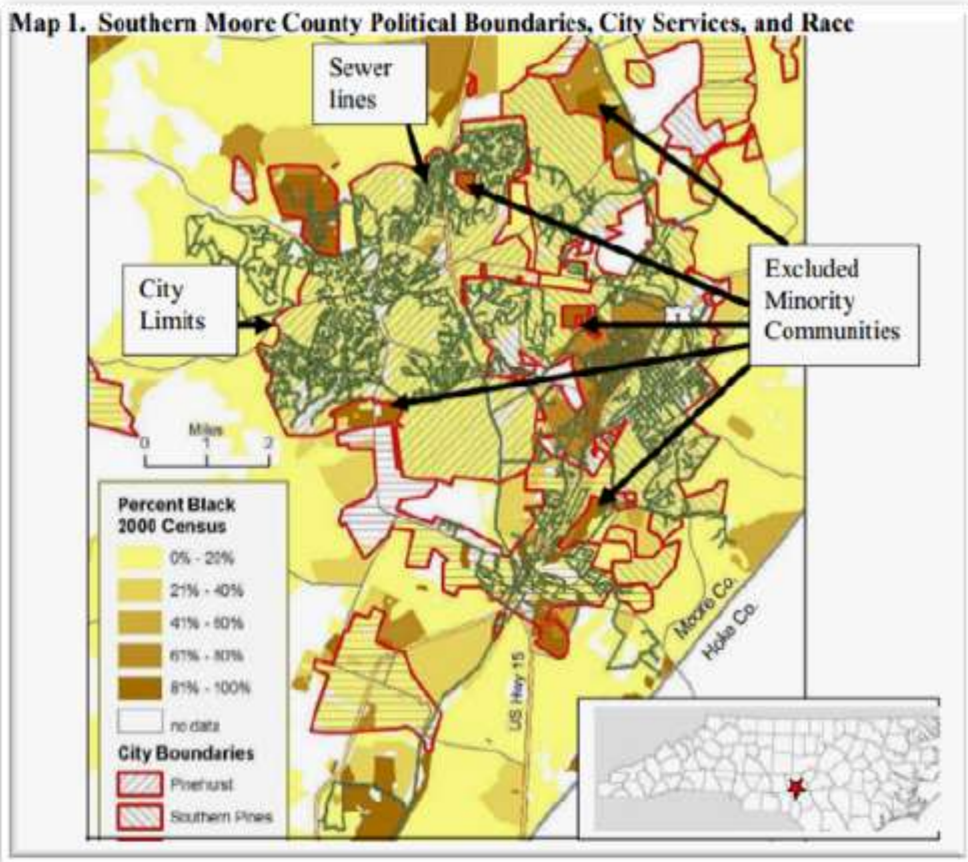
The challenge for Google is apparent within its management and leadership ranks where the workforce is dominated—to an extent—by men. Recognizing the value of a diverse workforce, Google is leveraging its data analytics capabilities to help change those numbers. Through analytics and research, the company identified that its employee advancement conventions, which in part call on employees to nominate themselves for promotions favor men, who are more likely to ‘raise their hands’ than women. Using the lessons gleaned from workplace analytics, Google has implemented programs to encourage women to apply for promotions and has reformed its hiring practices to ensure that female candidates meet female employees, with whom they are more likely to highlight their career achievements and credentials.

Source: http://www.google.com/diversity/at-google.html#tab=tech

Case Study 3: Allocation of Public Works (Cedar Grove Institute)

The Cedar Grove Institute for Sustainable Communities is a non-profit organization that leverages open data to explore disparities in allocation of geographic boundaries and public works of various communities. Cedar Grove uses a combination of demographic analysis, contextual investigation, housing and economic analysis, and geographic information systems to explore potentially discriminatory implications of public policy decisions. In a 2005 study, *Segregation in the Modern South: A Case Study of Southern Moore County*, Cedar Grove combined census data and other publicly available surveys and demographic information to explore the impact on community development of land annexation policies of Moore County North Carolina.

Map Showing Excluded Minority Communities



Source: http://www.cedargroveinst.org/

Case Study 4: Demographics of Health (State of New York)

In 2011, the Institute of Medicine, the health arm of the National Academy of Sciences, reported that lesbian, gay, bisexual, and transgender (LGBT) individuals have unique health experiences and needs, but as a nation, we do not know exactly what these needs are. The IOM also reported that clinicians and researchers are faced with incomplete information regarding the health status of LGBT individuals and that current research has not adequately examined subpopulations, particularly racial and ethnic groups and peoples' health needs based on age.

In response, the State of New York launched a coordinated, multi-agency effort to strengthen data collection regarding LGBT individuals in New York. The campaign will rely on data collected on a self-reporting basis by the New York's Department of Health, Department of Corrections and Community Supervision, Office for the Aging, Office of Mental Health, Office of Alcohol and Substance Abuse Services, Office of Temporary and Disability Assistance, Office of Children and Family Services, and Office for People with Developmental Disabilities. The data collected will be shared among the eight agencies to create a comprehensive method for identifying the needs of LGBT individuals. It is hoped that stronger data sets will empower the State and others to create more tailored approaches to reduce health disparities impacting LGBT individuals.

News Article
Wednesday, July 23, 2014

GOVERNOR CUOMO ANNOUNCES MULTI-AGENCY STATE EFFORT TO ADDRESS LGBT DISPARITIES

New York becomes first state in the nation with coordinated statewide strategy to improve LGBT data collection

Governor Andrew M. Cuomo announced that New York State is undertaking a coordinated, multi-agency effort to strengthen data collection for lesbian, gay, bi-sexual and transgender (LGBT) New Yorkers. Outlined in the first report by the State's Interagency LGBT Task Force, this statewide effort to include sexual orientation and gender identity information in data collections will allow the state to better tailor services to meet LGBT needs, ultimately improving the health and lives of thousands of New Yorkers. This effort makes New York the first state in the nation to employ a coordinated strategy to develop its data collection procedures for the LGBT community.

"New York State has a long history of advancing progressive ideals, and today we are continuing to lead the nation by identifying new ways to improve services and better meet the needs of the LGBT community," Governor Cuomo said. "By being more inclusive with how state agencies monitor the demographics of those they serve, we can address health and financial disparities, safety concerns, and a myriad of other issues that impact LGBT New Yorkers. This is another step forward for an important community in New York, and our administration will continue standing up for all New Yorkers, regardless of their sexual orientation or gender identity."

The Institute of Medicine in its March 2011 report, The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding, emphasized the need for collection of gender identity

Source: http://www.ocfs.state.ny.us/main/view_article.asp?ID=833

II. TRANSPARENCY IS A NECESSARY DISINFECTANT

Case Study 5: Hate Crime Report (Federal Bureau of Investigation)

The FBI's Uniform Crime Reporting program for hate crimes is a nationwide effort of more than 13,000 city, university and college, county, state, tribal and federal law enforcement agencies voluntarily reporting data on crimes brought to their attention. The data has become one of the country's primary methods of tracking, analyzing, and responding to hate crime violence. Hate crime incidents are broken down into various categories such as offense type, location, bias motivation, victim type, number of individual victims, number of offenders, and the race of the offenders. The streamlined and searchable nature of the data provides law enforcement and civil society groups an ability to monitor and analyze hate crimes and better direct training, advocacy, and legal efforts to reduce the number of hate crimes and improve the response to hate crime incidents.

Sources: http://www.fbi.gov/about-us/cjis/ucr/hate-crime/2012/hate-crime http://www.adl.org/press-center/press-releases/hate-crimes/adl-welcomes-fbi-hate-crime.html

Case Study 6: *McClesky v. Kemp*

McCleskey, an African American man, was sentenced to death after being convicted of armed robbery and the murder of a white police officer. In a writ of habeas corpus, McClesky argued that the Georgia capital sentencing process was administered in a racially discriminatory manner in violation of the Eighth and Fourteenth Amendments. In support of the claim, McClesky offered a statistical study (the Baldus study) to show disparities in the imposition of the death sentence in Georgia based on the murder victim's race and the defendant's race. The study was based on over 2,000 murder cases that occurred in Georgia during the 1970's, and involved data relating to the victim's race, the defendant's race, and the various combinations of such persons' races. The study found a consistent pattern of discrimination in the use of the death penalty against defendants who were charged with killing white victims compared to defendants who were charged with killing African American victims.

While the court ultimately found against McClesky, the case has been described as a turning point in the debate over the death penalty in the United States. The Baldus study has been replicated in numerous jurisdictions with similar findings. Race is now a powerful issue in debates over the death penalty because of studies like the Baldus study, which show that race can affect death penalty decisions.

Source: McClesky v. Kemp, 481 U.S. 279 (1987).

Case Study 7: Discrimination Complaint Data (EEOC)

The Equal Employment Opportunity Commission (EEOC) is responsible for enforcing federal laws that make it illegal to discriminate against a job applicant or an employee because of the person's race, color, religion, sex (including pregnancy), national origin, age (40 or older), disability, or genetic information.

In March 2013 the EEOC unveiled FedSEP, an electronic portal through which more than 325 federal agencies interact with the EEOC relative to their workforce and complaint data. This gateway provides each agency's staff with a single point of access to EEOC data collection systems and provides a new tool to collect and analyze government-agency data on workplace discrimination charges. By streamlining the submission process for so many documents and aggregating data from many sources, FedSEP allows EEO professionals greater ability to spot trends and uncover discrimination across the federal government.

Source: http://www.entellitrak.com/blog/detail/eeoc-cio-praises-partnership-with-micropact/

Case Study 8: *United States v. Sterling*

In 2006, Donald Sterling, his wife, and their family-trust real-estate company, *Beverly Hills Properties*, were accused of engaging in discriminatory practices in violation of the Fair Housing Act and Title VIII of the Civil Rights Act of 1968. The United States alleged, *inter alia*, that the defendants refused to rent portions of their 27-building development in the "Koreatown" neighborhood of Los Angeles to non-Koreans (*e.g., Hispanics and blacks)*. A study by Dr. Shelley Lapkoff used a database of tenant information released by the defendants to show that the number of Korean tenants across the 27 buildings had increased "significantly," from 64 percent to 83 percent, within the year following acquisition of those buildings by the defendants. Dr. Lapkoff's report also included an analysis of census data to determine whether the defendant's claim that the changing demographic distribution of Koreatown could explain the decreasing diversity of its tenants.

Dr. Lapkoff's analysis found that the overall demographics of Koreatown remained relatively stable during the period in question, and that Hispanics remained the dominant race in the area. The report concluded that, absent external changes, the increase in Korean tenants was consistent with the United States' allegation of housing discrimination. A later study by Dr. Lapkoff also used census data to show that there were no major shifts in household income or Korean households in the area that could explain the increase of Korean renters.

The studies helped lead to a settlement agreement that included a number of measures aimed at ending discriminatory renting practices in the 27 buildings owned by Beverly Hills Properties and required the defendants to pay $2,625,000 to be disbursed among aggrieved persons and a $100,000 civil penalty.

Table 3

Percentage of New Tenants Who Were Korean, Before and After Acquisition

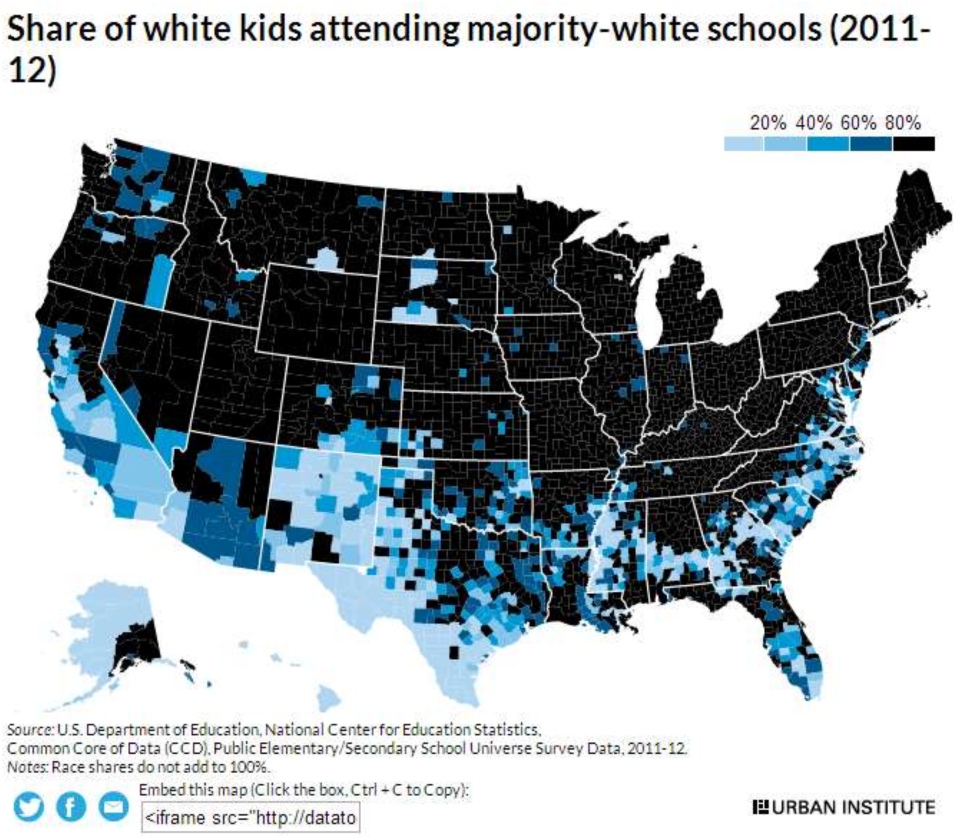
	Year Before Acquisition			Year After Acquisition		
Building #	Total	Korean	% Korean	Total	Korean	% Korean
Buildings Where Less than 80 percent of New Tenants are Korean at Time of Acquisition						
101	12	0	0%	16	4	25%
112	4	0	0%	6	6	100%
93	14	1	7%	18	13	72%
96	9	1	11%	11	9	82%
102	25	3	12%	22	8	36%
105	8	2	25%	7	5	71%
108	20	5	25%	30	28	93%
85	12	4	33%	14	13	93%
106	20	7	35%	18	11	61%
98	7	4	57%	11	11	100%
88	43	26	60%	41	34	83%
97	8	5	63%	10	8	80%
82	29	20	69%	27	25	93%
95	33	24	73%	44	38	86%
104	19	15	79%	28	25	89%
Subtotal	263	117	44%	303	238	79%
Buildings Where More than 90 percent of New Tenants are Korean at Time of Acquisition						
87	40	36	90%	45	43	96%
89	12	11	92%	13	13	100%
83	11	11	100%	9	9	100%
84	18	18	100%	17	14	82%
90	8	8	100%	19	17	89%
91	9	9	100%	4	4	100%
94	5	5	100%	10	10	100%
103	14	14	100%	18	16	89%
107	14	14	100%	10	9	90%
109	14	14	100%	20	18	90%
110	12	12	100%	7	6	86%
Total	157	152	97%	172	159	92%

Building # 86 Had No Rentals the Year Before Acquisition for Comparison

Source: United States v. Sterling, No. 2:06CV04885, (C.D. Cal. Nov. 12, 2009).

Case Study 9: Mapping Public School Segregation (Urban Institute)

Even as the country becomes more diverse – this year nonwhite students will account for the majority of public school students – black and Hispanic students often remain segregated from white students at historic levels. Drawing from the Department of Education's National Center for Education Statistics, the Urban Institute provides interactive county-level maps that track and visualize public-school segregation. The maps aggregate primary and secondary public-school enrollment by county and identify where white children predominantly attend majority-white schools and where minorities attend schools with predominantly minority classmates. The data is compiled using demographic information and a combination of five school surveys, covering the universe of all free public schools and school districts in the United States. It shows that despite the country's growing diversity, even extremely diverse regions of the country still have segregated school systems.



Sources: http://blog.metrotrends.org/2014/08/americas-public-schools-remain-highly-segregated/
http://www.vox.com/2014/8/19/6031279/majority-minority-public-schools
http://nces.ed.gov/ccd/pdf/psu12pgen.pdf

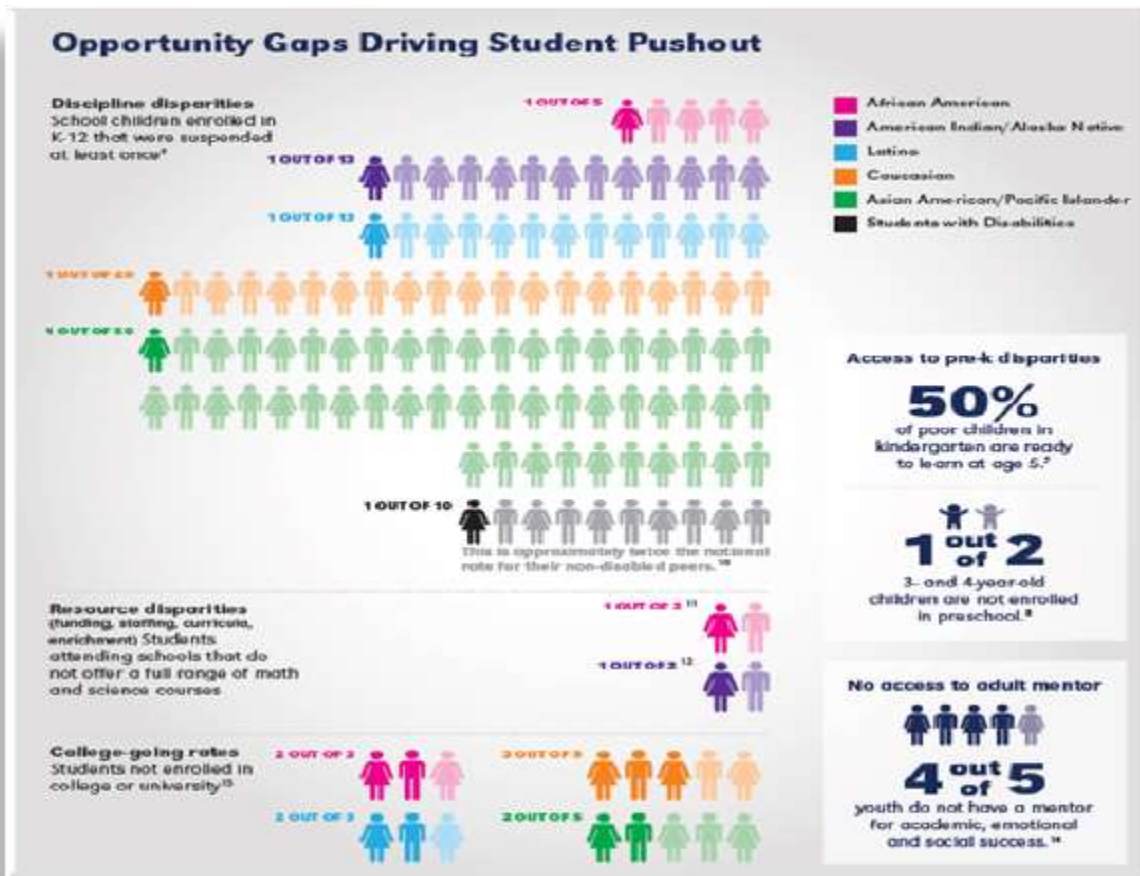
III. LONG-TERM PROBLEMS REQUIRE INNOVATIVE SOLUTIONS

Case Study 10: Education for All (NSBA)

A recent report by National School Boards Association (NSBA) offers novel policy solutions for increasing education rates in America. The report, *Partnerships, not Pushouts,* combines census data with data collected by various organizations to identify factors—known as "pushouts"—that may be responsible for driving young people away from education. Pushout factors can be more common among different segments of the population. For example, school suspensions—considered a major "pushout" factor, affect one out of five African American students and only one out twenty Caucasian students, which may partly explain the large discrepancy between graduation rates of those two groups.

To increase education levels among American youth, the NSBA proposes a variety student-centered "Personal Opportunity Plans" (POPs). To be effective, POPs are tailored to meet the needs of students on an individualized basis, addressing the pushout factor(s) most threatening to a particular student's academic success.

Discipline Disparities



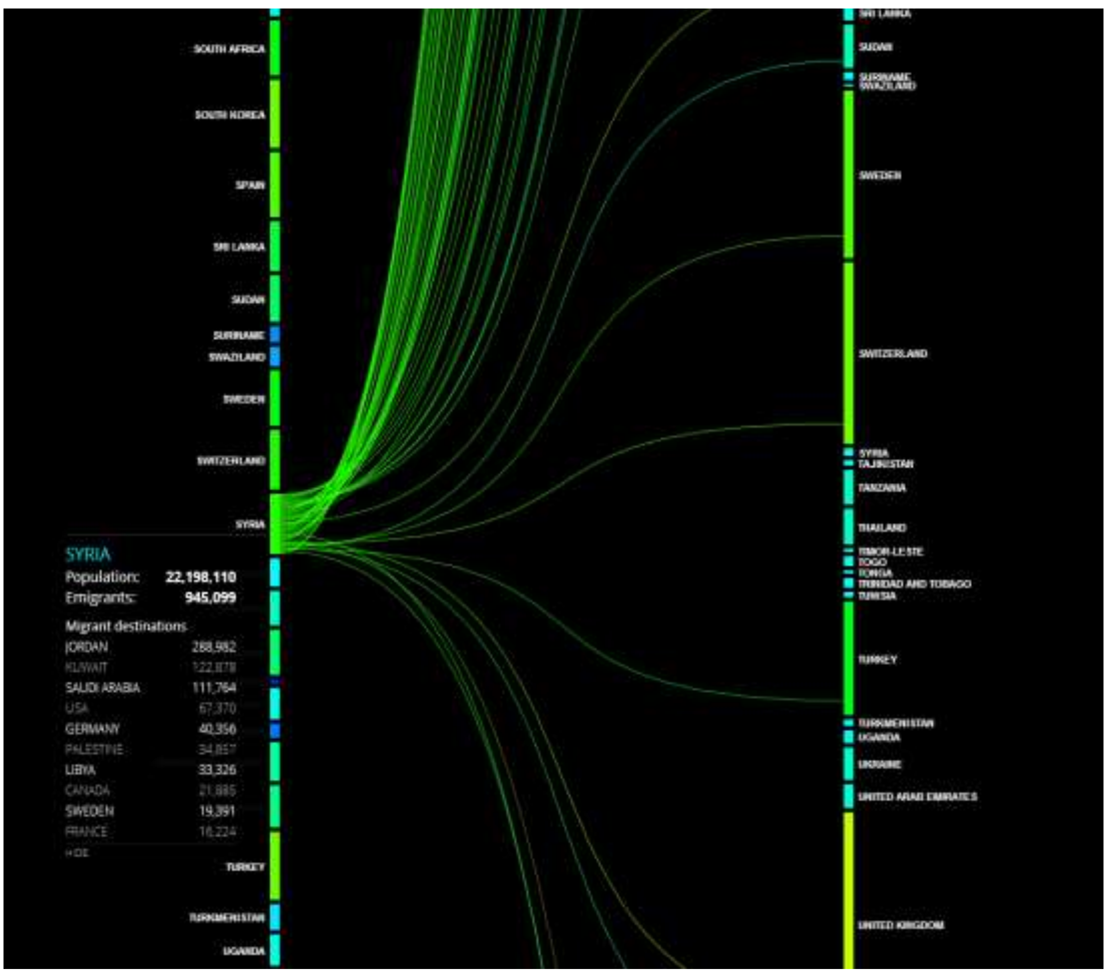
Source: http://www.nsba.org/sites/default/files/reports/Partnerships_Not_Pushouts_Guide.pdf

Case Study 11: Tracking Migratory Patterns (United Nations)

The United Nations has highlighted the social benefits of tracking migratory patterns of diverse peoples. For example, tracking the movement of displaced populations can empower humanitarian groups to provide better aid to those populations. As a new project under the *UN Global Pulse* banner, the organization is exploring new ways to track displaced populations using Big Data. The organization cites “significant shortcomings” with traditional methods of migratory benchmarking such as censuses, demographic and thematic surveys and administrative registers, which quickly become outdated.

In its review, Global Pulse highlights a number of studies that rely on Big Data collected from social media sites or open data initiatives to draw important conclusions about population movements. In the example below, PeopleMovin, repurposes “open” migration, refugee and asylum, and world population data to create an interactive tool allowing users to quickly identify international movement patterns and identify where relief efforts are most valuable.

Migration Patterns from Syria:



Source: http://peoplemov.in/#f_SY

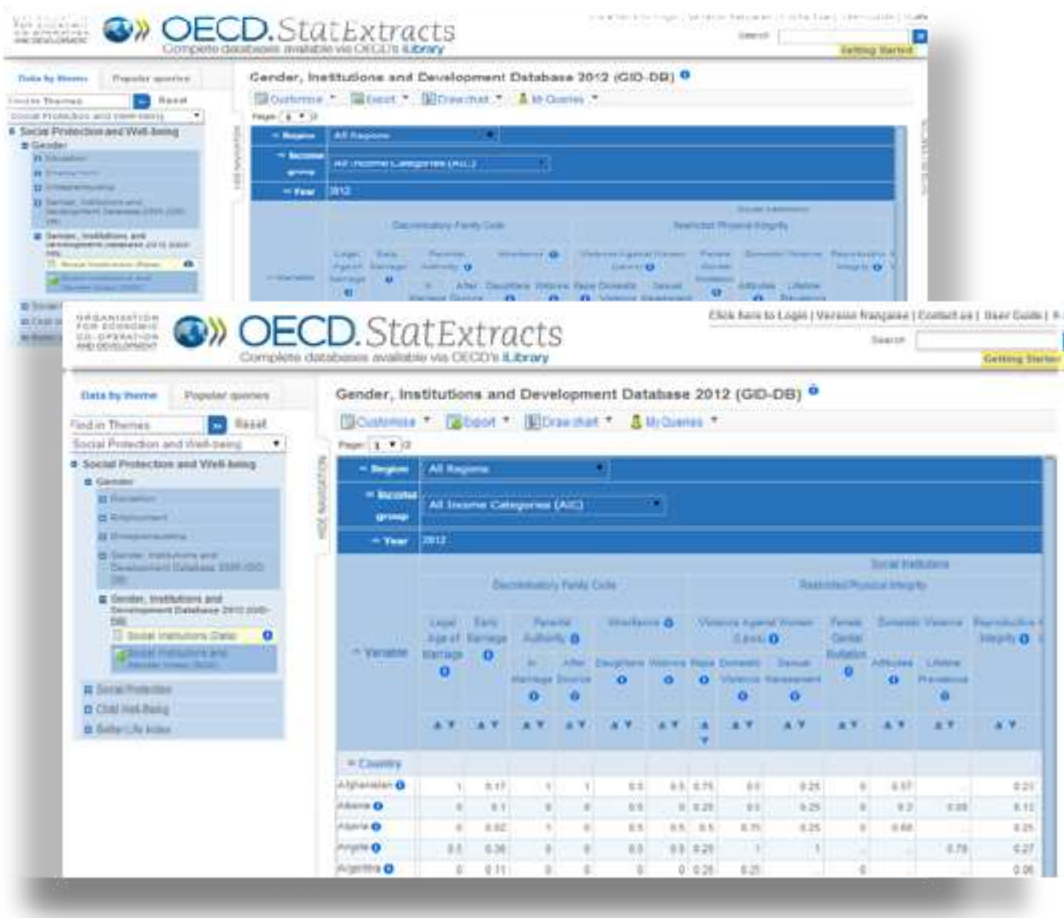
Case Study 12: Economic Development and Equality (OECD)

Since 2009, the Organization for Economic Co-Operation and Development has offered the publicly available *Gender, Institutions and Development Database.* The database compiles gender-discrimination data from 160 countries to provide researchers and policymakers with an analysis of 60 detailed variables, ranging from factors like "Discriminatory Family Code" to "Restricted Civil Liberties," that are likely to impact women's engagement in society and the economy.

A defining feature of the GID-DB is that, in addition to traditional quantitative analyses, the database uses an innovative scoring system to evaluate discriminatory institutional features. For example, while traditional studies of "early marriage" analyze rates of marriage among various age groups, the GID-DB has created a scaled system that combines rates of early marriage with an analysis of legal, traditional and religious customs to provide a much deeper look at gender discrimination. An example of the scaled system is provided below:

- 0: The law on the minimum age of marriage does not discriminate against women.
- 0.5: The law on the minimum age of marriage discriminates against some women, for example through customary, traditional and religious law.
- 1: The law on the minimum age of marriage discriminates against all women or there is no law on the minimum age of marriage.

Gender, Institutions and Development Database (GID-DB)



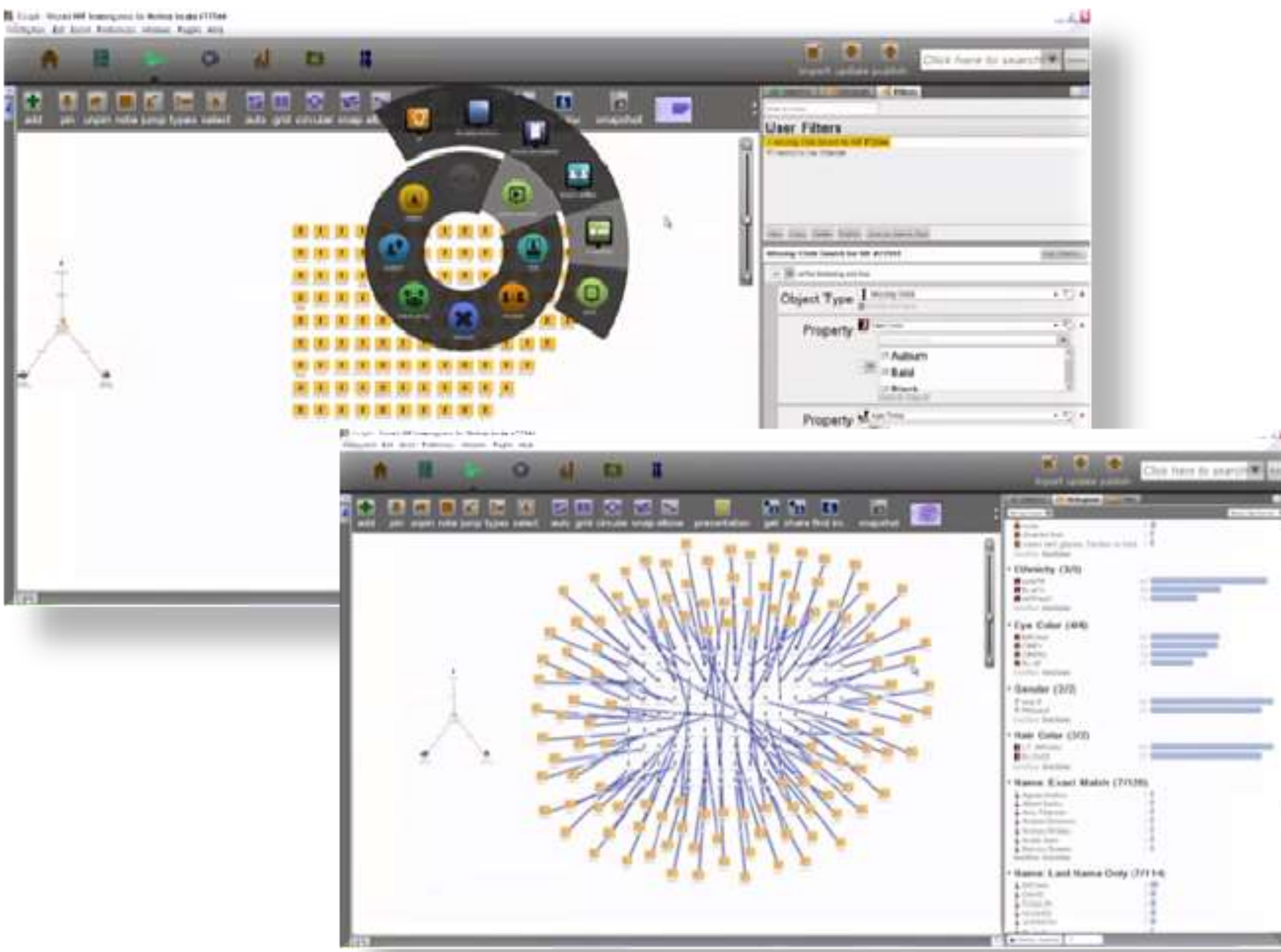
Source: http://stats.oecd.org/Index.aspx?DataSetCode=GIDDB2012

Case Study 13: Finding Missing and Exploited Children (Palantir/NCMEC)

The National Center for Missing & Exploited Children (NCMEC) compiles a wide variety of information from law enforcement, social media, and proprietary databases. Much of this information has traditionally been stored in siloed databases, requiring analysts to manually query each database when investigating a case. The Big Data analytics tool, developed in 2010 by Palantir, empowers NCMEC analysts to query a range of databases simultaneously.

The below case illustrates how the NCMEC uses Big Data to save children:

> A 17-year-old girl was reported missing and suspected of being a victim of sex trafficking. Through various searches, a NCMEC analyst was able to find multiple posts online that advertised this missing child for sex. Through information in the ads, the analyst was able to tie them to other posts from the same pimp. The analysis included over 50 advertisements, 9 different females, and a trail covering 5 states. A Link Analysis graph was created using Palantir that helped law enforcement to easily see the large scope of the ring. This insight helped law enforcement link the pimp to a multitude of other crimes and other girls that he victimized.



Sources: https://www.palantir.com/wp-assets/wp-content/uploads/2014/01/NCMEC-Impact-Study.pdf
https://www.youtube.com/watch?v=TKpam_iy3Fo

Case Study 14: Human Trafficking (Palantir/Polaris Project)

Human Trafficking is a problem facing tens of millions of people and their families around the world. According to the *Polaris Project*, each year 21 million people are enslaved worldwide to generate a profit of $32 billion for their captors. To combat this global problem, organizations like Polaris Project maintain extensive databases of information collected from various public and private sources. The organization reports that it may collect up to 170 different quantitative and qualitative variables per case record, including first-hand data obtained through its National Human Trafficking Resource Center Hotline. In 2013, the NHTRC received 31,945 phone calls, 1,488 e-mails, 1,669 tips from online form submissions, and 787 SMS threads. In order to leverage this vast amount of data, the organization uses the Palantir Gotham analytics platform to track trafficking rings, quickly identify discrete human-trafficking events, and mobilize appropriate response units.

Sources: http://www.polarisproject.org/resources/hotline-statistics/human-trafficking-trends-in-the-united-states https://www.youtube.com/watch?v=kdQrLMEF-Eg#t=82

FUTURE OF PRIVACY FORUM

Anti-Defamation League®