Testimony of Pam Dixon

Executive Director, World Privacy Forum

Before the US Senate Committee on Banking, Housing, and Urban Affairs

Data Brokers, Privacy, and the Fair Credit Reporting Act

June 11, 2019

Chairman Crapo, Ranking Member Brown, and Members of the Committee, thank you for the opportunity to testify today on this important subject of data brokers, privacy, and the Fair Credit Reporting Act. Today data brokers are selling unregulated predictions and scores to the financial industry. The financial industry is using these scores to market unfair and unjust products to consumers that limit or obscure their access to loans, credit, and financial services.

Today I offer you four core observations and two solutions:

1. **Credit scores and predictions are being sold that are not regulated by the FCRA,**
2. **The technology environment is facilitating more scores being used in more places in consumers' lives, and not all uses are positive,**
3. **These scores are created without due process for consumers,**
4. **These scores can cause consumers exceptional harm.**

Therefore, Congress must:

1. **Expand the Fair Credit Reporting Act to regulate currently unregulated financial scores that affect consumers,**
2. **Enact a standards law that will provide due process and fair standard setting in the area of privacy.**

By doing these things, Congress will protect consumers and allow them to act to fill in gaps where privacy harms are occurring, along with other stakeholders.

I am the founder and executive director of the World Privacy Forum (WPF).[1] WPF is a non-profit, non-partisan 501(c)(3) public interest research group, and we have been researching, documenting, publishing, benchmarking, and educating on privacy topics since 2002. We have done significant work in digital privacy in our key issue areas of health, data brokers, AI and machine learning (broadly, predictive analytics), identity, biometrics, governance models of complex digital ecosystems, and privacy and vulnerable populations, including children.

Data broker issues have been one of our core areas of work for almost two decades. In the area of data brokers, we have published multiple reports,[2] crafted and delivered education for consumers, and I have testified before Congress on the topic on three occasions. In 2007, I had my personal "AI moment" when I realized that more and more consumers were being placed in custom classifications by data brokers, and these classifications were being offered for sale, with the pitch that the classifications would more accurately predict consumer behavior.[3] This understanding led to an early discussion draft of a paper about predictive analytics, and then to a deeply researched report, published in 2014 with my co-author Robert Gellman, called the Scoring of America.[4] It was the first major report that benchmarked consumer scores and analyzed data broker activities in scoring in light of existing policy, particularly the Fair Credit Reporting Act.[5] It was among the catalysts of the time that sparked an ongoing conversation about AI and privacy.

Since that time, I have conducted extensive field work and research regarding AI and privacy in the area of identity, machine learning, biometrics, and privacy. In 2018, after publishing extensive, peer-reviewed original research on biometrics and EU-US policy in Springer-Nature,[6] I was invited to serve on the OECD's AI Expert Group, (AIGO) which was composed of leading

[1] World Privacy Forum, https://www.worldprivacyforum.org.

[2] Robert Gellman and Pam Dixon, Data Brokers and the Federal Government: A New Front in the Battle for Privacy Opens,Third report in a series on data brokers, October 30, 2013. Available at: http://www.worldprivacyforum.org/wp-content/uploads/2013/10/WPF_DataBrokersPart3_fs.pdf.

[3] An exemplar of this type of classification, sometimes called consumer segmentation, is Acxiom's Personicx Customer Segmentation. See: Acxiom, Personicx Home Page. Available at: https://www.acxiom.com/what-we-do/consumer-segmentation-personicx/.

[4] Pam Dixon and Robert Gellman, *The Scoring of America: How Secret Consumer Scores Threaten Your Privacy and Your Future*, World Privacy Forum, April 2014. Available at: https://www.worldprivacyforum.org/wp-content/uploads/2014/04/WPF_Scoring_of_America_April2014_fs.pdf.

[5] 15 U.S.C. § 1681 et seq.

[6] Pam Dixon, *A Failure to Do No Harm: India's Aadhaar biometric ID program and its inability to protect privacy in relation to measures in Europe and the U.S., S*pringer Nature, Health Technology. DOI 10.1007/s12553-017-0202-6. http://rdcu.be/tsWv. Open Access via Harvard- Based Technology Science: https://techscience.org/a/2017082901/.

AI and machine learning experts in OECD member countries.[7] I learned an extraordinary amount from my time drafting the OECD Global Guidelines on AI,[8] which are now adopted, published, and as of May 2019 have been ratified by the US Government.[9] What I learned convinced me of the need to do more on predictive analytics and update the Scoring of America report.

I will be highlighting three key points in my discussion today:

- What is new and different about scoring and data brokers in today's world?
- What are the key problems in consumer scoring and prediction, and how does the FCRA currently address these problems?
- What are potential solutions to risks and harms produced by unregulated consumer scoring activities?

Regarding solutions, in my testimony, I will discuss two key solutions. The first solution is ways in which the Fair Credit Reporting Act (hereafter FCRA) can accommodate advances in prediction techniques. I will also discuss a draft bill that law professor Jane Winn and I co-authored regarding the use of due process standard setting (voluntary consensus standards, as defined by OMB Circular A-119[10]). The bill facilitates setting fair, multi-stakeholder, due process standards in the areas of privacy that would benefit from specific, granular guidance.[11] The bill presents a way of using standards to fill in meaningful gaps in privacy protections.

[7] OECD AI Expert Group roster, OECD. Available at: http://www.oecd.org/going-digital/ai/oecd-aigo-membership-list.pdf.

[8] OECD, Recommendation of the Council on Artificial Intelligence, Adopted on 5/21/2019. Available at: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

[9] NTIA, The US Joins with OECD in Adopting Global AI Principles, May 22, 2019. Thus far, the US has been among 42 countries to approve the new international agreement on AI.

[10] OMB Circular A-119, "Federal Participation in the Development and Use of voluntary Consensus Standards and in Conformity Assessment Activities," 2016 Revision, 81 FR 4673 pages 4673-4674. Available at: https://www.nist.gov/sites/default/files/revised_circular_a-119_as_of_01-22-2016.pdf.

[11] The United States has a sectoral regulatory framework. Sector-based legislation is legislation that applies to just part of the economy, for example, the government sector, or the financial sector. "Sector" means "A part or subdivision, especially of a society or an economy." Harper Collins English Dictionary, "sector." Available at: https://www.collinsdictionary.com/dictionary/english/sector.

I. Introduction

To score is human, which is why prediction as a business model is proliferating today, as are the scores that function as a form of modern shorthand describing our preferences and even future inclinations and abilities.[12] We're all familiar with traditional credit scores, which are regulated by the Fair Credit Reporting Act. But credit scores have been joined by literally thousands of new, unregulated predictive scores ranging from financial scores (consumer lifetime value scores) to health scores (frailty scores) to educational scores (College Board's "adversity score").[13] The application of predictive analytics and scoring, when done properly, can introduce efficiencies in situations with high-velocity data in complex data ecosystems, for example, making better predictions in financial markets.[14] Many scores may be quite neutral in practice, and not all predictive scoring applies to human behavior. But when prediction is applied to individuals and groups and is used without appropriate guardrails, the predicting and scoring of Americans' preferences, skills, and imagined future can introduce meaningful harms to individuals, groups, and institutions. WPF calls this kind of unregulated scoring "consumer scores," which we define as follows:

> A consumer score that describes an individual or sometimes a group of individuals (like a household), and predicts a consumer's behavior, habit, or predilection. Consumer scores use information about consumer characteristics, past behaviors, and other attributes in statistical models that produce a numeric score, a range of scores, or a yes/no. Consumer scores rate, rank, or segment consumers. Businesses and governments use scores to make decisions about individual consumers and groups of consumers. The consequences can range from innocuous to important. Businesses and others use consumer scores for everything from predicting fraud to predicting the health care costs of an individual to eligibility decisions to almost anything.[15]

Due to advances in hardware and scoring systems algorithms, it is becoming easier and less expensive to create unregulated versions of credit scores that closely approximate the prediction quality of regulated credit scores.[16] This capacity has created new pressures on the efficacy of the Fair Credit Reporting Act, and it has created pressures on the public's trust regarding how their personal information, or information derived about them, is distributed and subsequently used by

[12] Scores in this testimony refer to numeric scores derived from the results of AI analysis built using predictive modeling. Predictive modeling uses copious amounts of information fed through analytical systems to predict future performance or activity, based on past information.

[13] For discussions regarding Consumer Lifetime Value scores and Frailty scores, see Appendix B in this report. The adversity score is discussed later in this testimony.

[14] Simulating financial markets on deep learning models, ReWork April, 2017. Available at: https://medium.com/@teamrework/simulating-financial-markets-on-deep-learning-models-39ccb7219fc6.

[15] Dixon and Gellman, The Scoring of America at 8.

[16] FICO,

third parties with whom the consumer has no relationship. There are now literally tens of thousands of consumer scores that have been created by data brokers and others to predict aspects of consumer behavior, group behavior, various types of risk, and more. Unfortunately, the observations we made five years ago in Scoring of America about consumer score secrecy and unfairness still hold true today: most consumer scores are secret, and it is very difficult to even learn when a score is being used in your life. The secrecy of some scores may have, in fact, gotten worse, as discussed in the case study section of this testimony.

After 20 years of conducting research on data brokers, I've seen the data broker industry evolve from paper-based lists of consumers to digitized lists of consumers. Now, a new evolution is taking place as data brokers move to AI models that predict consumer behavior. Predicting a behavior or intent is the core of the new data broker business model. Because of the new data broker business models, it is important that we work to define data broker activities in more focused ways that clearly articulate risks harms so as to craft appropriate and effective mitigations. Overbroad approaches that attempt to "boil the ocean" are unhelpful. However, more focused work can address the most significant issues.[17]

Even with thoughtful work, however, we have a challenging problem to solve here, because data broker activities have shifted to prediction, and are in some instances are creating new risks and harms that are difficult to readily define.

II. What is scoring today?

When we wrote the Scoring of America, we did not know it then, but in 2014 we were seeing the beginnings of a major shift. We were looking at the scoring issue from the bottom up, doing benchmarking research on what existed and was happening at the ground level at the time, as well as what companies were engaged in the activities. We interviewed companies, experts, data analysts, consumers, and thought leaders; we visited the FTC and spent many hours discussing the fine points of the FCRA with the fine attorneys there. We attended conferences, and read through wide swaths of literature. We didn't have an economic theory of AI, what we had was an idea of something important, and that this something may be more than what existing legal and regulatory structures could address. At this time, data brokers were still fairly focused on selling lists — some of them highly objectionable — of consumers. The predictive aspects came into

[17] The state of Vermont has enacted a thoughtful and incremental approach to data brokers. The focus of Vermont's statute was to prohibit discriminatory uses of data, and to create transparency around third party data brokers. See: Vermont Secretary of State, Data Broker Page. Available at: https://www.sec.state.vt.us/corporationsbusiness-services/data-brokers.aspx. Note Vermont's definition of data broker: "A Data Broker is a business, or unit or units of a business, separately or together, that knowingly collects and sells or licenses to third parties the brokered personal information of a consumer with whom the business does not have a direct relationship. 9 V.S.A. § 2430(4)(A)."

play when data brokers began to sell products that classified consumers and scored behaviors into predictions.

That was 2014. Now, in our updated research, what we have found is a rapid expansion of scoring activities, to the point that the entire data broker business model has radically shifted. Data brokers, as we used to know them, are barely recognizable today. There are entirely new data broker business models, and prediction capacity has radically changed the industry. The major trends such as AI, machine learning and its subsets like biometrics, all manner of large data sets and predictive analytics, the Internet of Things, mobile, cloud, and fully digital and dematerialized identity ecosystems are all emerging apace now.[18] These technologies are fusing and converging to create something quite complex. This is not the same world as the Internet as a General Purpose Technology. This "data fusion" is a world that is bringing new and novel tensions that legislative structures have not yet addressed.

Here are some key elements that have changed:

Scoring and prediction have advanced dramatically

AI and machine learning have undergone radical changes since 2014. Particularly in the last four years, machine learning techniques such as neural networks have transformed data modeling and predictive accuracy. Two areas, accuracy and speed, are important to understand here to get a picture of the full extent of the technological transformation.

Accuracy gains

"The accuracy argument" — that scores are patently harmful because they are inaccurate — is still important, but the argument has changed. This is a foundational point to understand, because if the predictions that data brokers are selling are largely accurate, then policy mechanisms need to shift to address what to do when there is an accurate score that can create risk, havoc, or diminished opportunities in a person's life. Two prediction accuracy use cases readily trace the arc of the advances I am describing.

In his book Prediction Machines: The Simple Economics of Artificial Intelligence, Ajay Agrawal explains that in the financial sector, anti-fraud analysis techniques achieved about an 80 percent rate of capture of fraud in the late 1990s. There was incremental improvement until recently, when machine learning techniques pushed the capture rate up to 99.9 percent.[19] Similarly in the field of facial recognition, which is a subset of AI, it, too, has achieved remarkably similar advances in accuracy in approximately the same time frame. In the fall of 2018, typically staid

[18] Pam Dixon, *Digital Identity Ecosystems,* Paper presented at Harvard Kennedy School Feb. 5, 2019, World Privacy Forum. Available at: https://www.worldprivacyforum.org/2019/02/digital-identity-ecosystems/.

[19] Ajay Agrawal et al, Prediction Machines at chapter 3.

and understated scientists at NIST characterized the advances in machine learning in facial recognition as a "technological revolution."[20] The algorithm testing NIST conducted found some of the most accurate facial recognition algorithms in history, with some of them achieving accuracy above the 99th percentile. Substantial accuracy gains are advancing through multiple areas of predictive systems.

Predictive speed: real time scoring and analysis

The speed of prediction and decision making is another change. Accurate, instant predictions are powerful tools. The scope of these advancements can be traced in the financial sector with ease, with documentation showing advances in real time prediction and analysis that were not imaginable pre-2015. In January 2019, the financial sector watchdog FINRA posted its analysis of 2018's market activity — it generated an historic processing volume of 66.7 billion electronic records per day, which was an 87.4 percent increase over daily volume in 2017. The most salient point from our perspective, though, was that the CIO noted that FINRA was sustaining very high volumes for "days and weeks at a time" while doing real-time threat analysis on 200 algorithmic patterns designed to search for 300 threat scenarios. FINRA is essentially predicting and deciding in real time, or near real time. The implications of accurate, high speed predictions of consumer behavior have not yet fully made their way through the existing regulatory process, but it has begun now in the financial and some technical sectors.[21]

Data as a commodity

In Congressional testimony and in our Scoring of America report, we documented many examples of third party companies selling lists of consumers. These kinds of lists still exist.[22] However, these lists are widely seen as commodity items now. Data itself has become a commodity as we have progressed from the early stages of digitization to today. Lists that used to cost hundreds or thousands of dollars are now seen as fodder for the predictive engines, not as ends in themselves. Additionally, data that used to be exclusive to data brokers now is more widely available, and is no longer the sole data stream AI experts use to craft predictive algorithms.

Data broker business models have shifted

[20] NIST, Facial Vendor Recognition Test, November 2018. Available at: https://nvlpubs.nist.gov/nistpubs/ir/2018/NIST.IR.8238.pdf.

[21] Again, biometrics provides another use example of highly accurate, real-time analysis. A Chinese company, Yitu, has pioneered real-time biometric facial prediction. Amanda Lentino, *This Chinese facial recognition start-up can identify a person in seconds.* May 17, 2019. Available at: https://www.cnbc.com/2019/05/16/this-chinese-facial-recognition-start-up-can-id-a-person-in-seconds.html.

[22] See the classic list finder, Nextmark Listfinder. Available at: https://www.nextmark.com.

The Digital Marketing Association, now part of the ANA,[23] has maintained a vendor marketplace for many years. In 2013, the data broker marketplace was segmented into list brokers, direct mail houses, public records specialists, and a modest roster of analytics firms.[24] These were largely third party companies that were conducting various tasks with data, from collection to distribution and many points in between. Typically, the business models involved third-party relationships of varying kinds, with varying kinds of privacy assurances, many being done at a contractual level.

Today, several key changes are apparent. The overall marketplace is now focused on AI and prediction, and the race is to build in the hardware other other infrastructure elements that will support advances in the technology. All told, yesterday's lists of consumers appear quaint next to the powerful predictive systems that are proliferating. Lists are commodities; now rapid consumer prediction is the goal. As prediction gets less and less expensive, we can expect predictive activities to jump out of the realm of traditional data brokers and into new and unexpected areas and places. Ajay Agrawal notes, "Economics offers clear insights regarding the business implications of cheaper prediction. Prediction machines will be used for traditional prediction tasks (inventory and demand forecasting) and new problems (like navigation and translation). The drop in the cost of prediction will impact the value of other things, increasing the value of complements (data, judgment, and action) and diminishing the value of substitutes (human prediction)."[25]

Already, it is clear that changes beyond prediction are occurring in the data broker business model; for example, some data brokers have been acquired and have become absorbed into larger businesses, as those businesses seek to own first party data.[26] This will eventually further undermine traditional data broker models and create something new. It's not that every business will "become a data broker"; there are still risks associated with some types of third party data uses. However, the old data broker models are fracturing and changing into predictive models that are more dispersed. We are seeing the edges of what this is looking like, and this is what we turn to next.

III. Key examples of modern scoring products

[23] DMA and ANA home page, https://thedma.org.

[24] Original research for the Scoring of America report, 2013 analysis of DMA Vendor Marketplace. PDFs available.

[25] Prediction Machines at Chapter 2.

[26] Angelina Rascouet, Publicis Surges as $4.4 Billion Epsilon Deal Deepens Data Push
April 14, 2019. Available at: https://www.bloomberg.com/news/articles/2019-04-14/publicis-to-buy-alliance-data-s-epsilon-unit-for-4-4-billion. See also Seb Joseph, With Epsilon deal, Publicis bets on first party data for survival, Digiday, April 15, 2019. Available at: https://digiday.com/marketing/publicis-epsilon-data/.

In the Scoring of America, you will find an extensive list of scoring products, many of which are still in existence. I will not repeat that list here. I would like to instead focus on some key products that serve as exemplars of the newest problems, and point to the pathways to solutions that will be necessary to create improvements. These products fall into two categories:

- Unregulated credit scores

- Use of consumer prediction scoring in educational eligibility circumstances

Unregulated credit scores: aggregated or "household" credit scores

Unregulated credit scores are predictive consumer scores that function like a regulated consumer score. They are seen by those who use them as unregulated because the scores exploit a loophole in the FCRA. Therefore, these scores do not fall under the FCRA and as such they are supposed to only be used for "marketing purposes." There are numerous exemplars of unregulated credit scores today. FICO offers a traditional regulated credit score, and they offer a separate unregulated credit score called an Aggregate FICO, which is offered through Equifax.[27] In the screenshot below is a description of the Aggregate FICO; it has numerous detailed financial metrics, and includes a neighborhood risk score. It can be used in meaningful financial contexts, as seen in the screenshot.

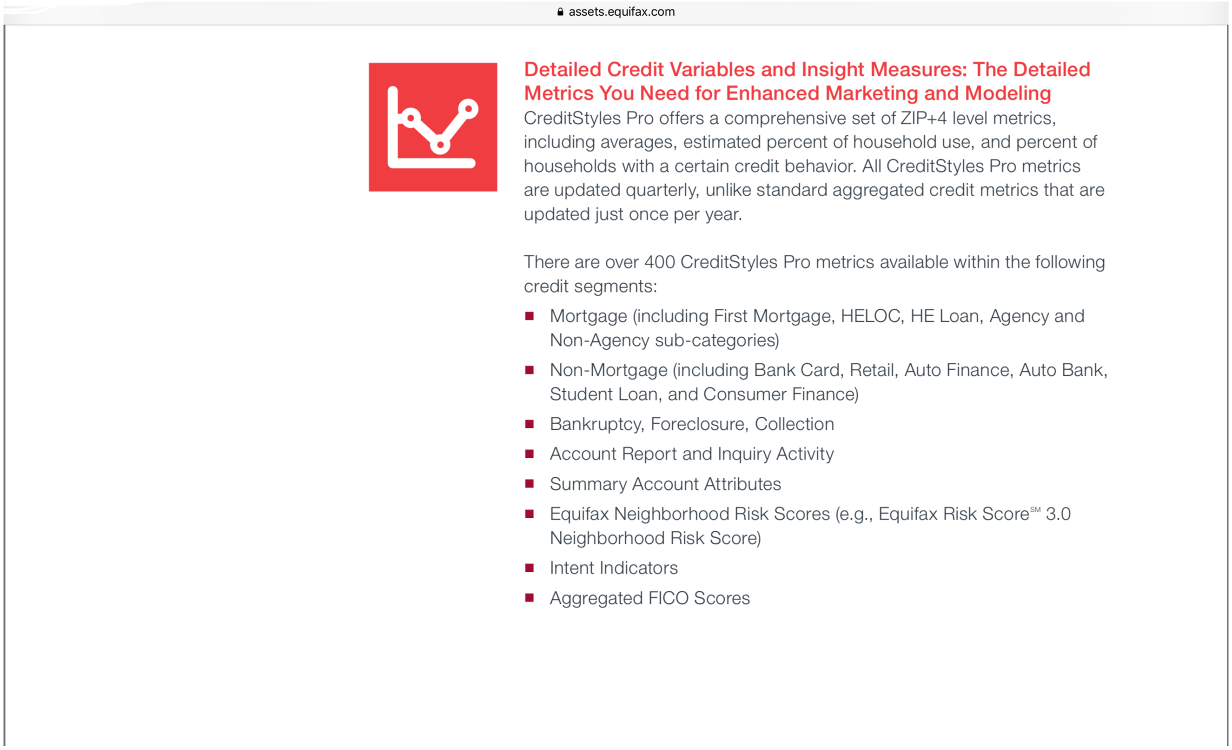
Aggregated FICO scores

Analytics IQ also offers several flavors of consumer scores, one of which appears to be an unregulated credit score.[28]

Why are some credit scores unregulated?

[27] Credit Styles Pro, Equifax. Available at: https://assets.equifax.com/assets/usis/creditStylesPro_ps.pdf.

[28] Analytics IQ, GeoCreditIQ, which is intended to be used for marketing purposes. Available at: https://analytics-iq.com/what-we-do/.



assets.equifax.com

Detailed Credit Variables and Insight Measures: The Detailed Metrics You Need for Enhanced Marketing and Modeling

CreditStyles Pro offers a comprehensive set of ZIP+4 level metrics, including averages, estimated percent of household use, and percent of households with a certain credit behavior. All CreditStyles Pro metrics are updated quarterly, unlike standard aggregated credit metrics that are updated just once per year.

There are over 400 CreditStyles Pro metrics available within the following credit segments:

- Mortgage (including First Mortgage, HELOC, HE Loan, Agency and Non-Agency sub-categories)
- Non-Mortgage (including Bank Card, Retail, Auto Finance, Auto Bank, Student Loan, and Consumer Finance)
- Bankruptcy, Foreclosure, Collection
- Account Report and Inquiry Activity
- Summary Account Attributes
- Equifax Neighborhood Risk Scores (e.g., Equifax Risk Score℠ 3.0 Neighborhood Risk Score)
- Intent Indicators
- Aggregated FICO Scores

The FCRA stands as one of the earliest and most important early implementations of Fair Information Practice principles. It is an extraordinarily well-designed law; deliberate and effective, it finds a balance between interests and gives all stakeholders clear roles, rules, and responsibilities. It is among the cornerstones of financial sector privacy regulation. The FCRA was created to solve problems of trust between consumers and the credit bureaus. Credit bureaus were collecting information from people, and using it in undisclosed ways. The FCRA stopped that in the 1970s when it was enacted, and it created transparency. We can see our credit report and correct it, we can see our credit score, and there are accuracy and other requirements for data furnishers.

However, the FCRA has some loopholes.

The most significant loophole is the "household loophole." The FCRA applies to *individuals*, not households. A credit score that applies to a household does not fall under the FCRA, especially if it also does not use regulated factors (such as information in a credit bureau report.) Forty years ago, companies building predictive credit scores needed to use the credit bureau report data, as it was the primary data available. Today, however, credit risk may be accurately predicted using a wide variety of newer factors, from purchase history to "intent indicators" to neighborhood risk scores. Unregulated credit scores may have a thousand factors instead of just a handful. But the factors in the unregulated credit score may potentially include marital status, or age, for example, both factors prohibited in financial sector laws like the Equal Credit Opportunity Act, depending on the product.[29] It is not possible to know, because without the transparency afforded by the FCRA, the unregulated credit scores are opaque.

If a consumer receives a marketing offer that falls short of a firm offer of credit or insurance by a micrometer, how is this unregulated offer fundamentally different in practice than a firm offer of credit or insurance as defined in the FCRA? If the unregulated credit scores end up achieving higher accuracy levels than the regulated credit scores (which is a possibility, if it has not already happened) then the market incentives for companies to use regulated credit scores is greatly diminished.

In an unregulated prediction marketplace, abuses can occur. Regulators should rightly be concerned about the uses of unregulated credit scores in the razor-thin line between the marketing of credit opportunity, and firm offers of credit.

For all unregulated credit scores, consumers should get to see their score, and should have the full complement of their rights under the FCRA. Congress should use its authority to make a deliberative investigation into the facts of unregulated credit scores and to conduct an analysis of how these products are being used, as well as neighborhood risk scores.

Consumer Scoring in Education: The College Board "Adversity score"

The College Board has launched a controversial new initiative to provide a "context" for student's academic performance. The College Board calls its program the "environmental context dashboard." Information about a student's home, neighborhood, and school background is given to participating colleges to view in a dashboard. In the sample dashboard, students receive an overall high school adversity score. The SAT score is a separate score based on performance on the SAT test. The adversity score is

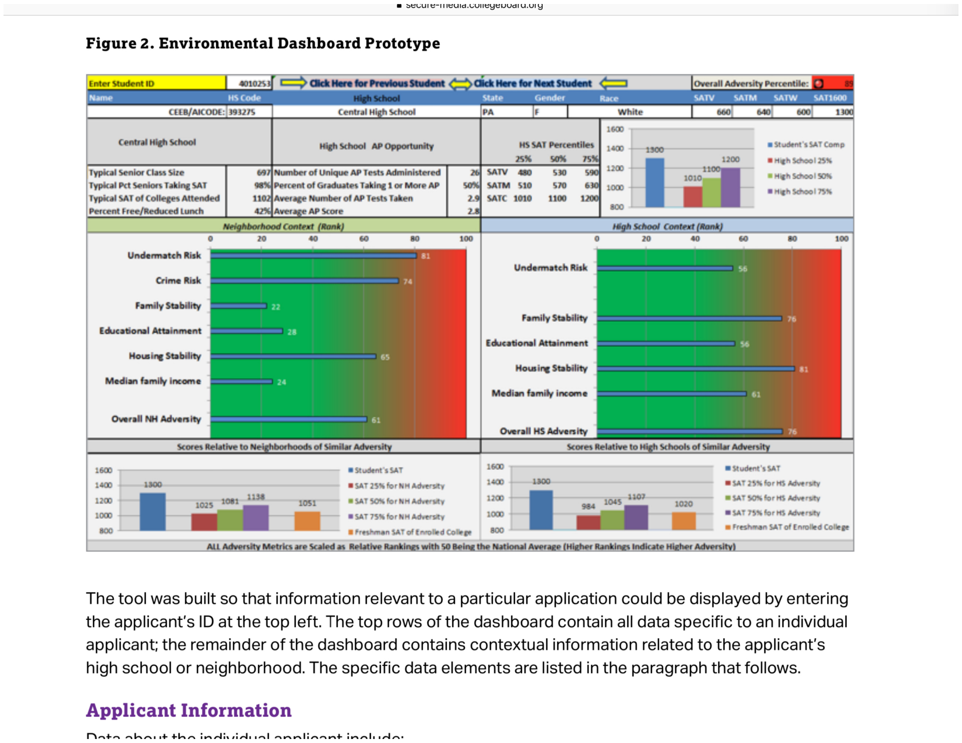


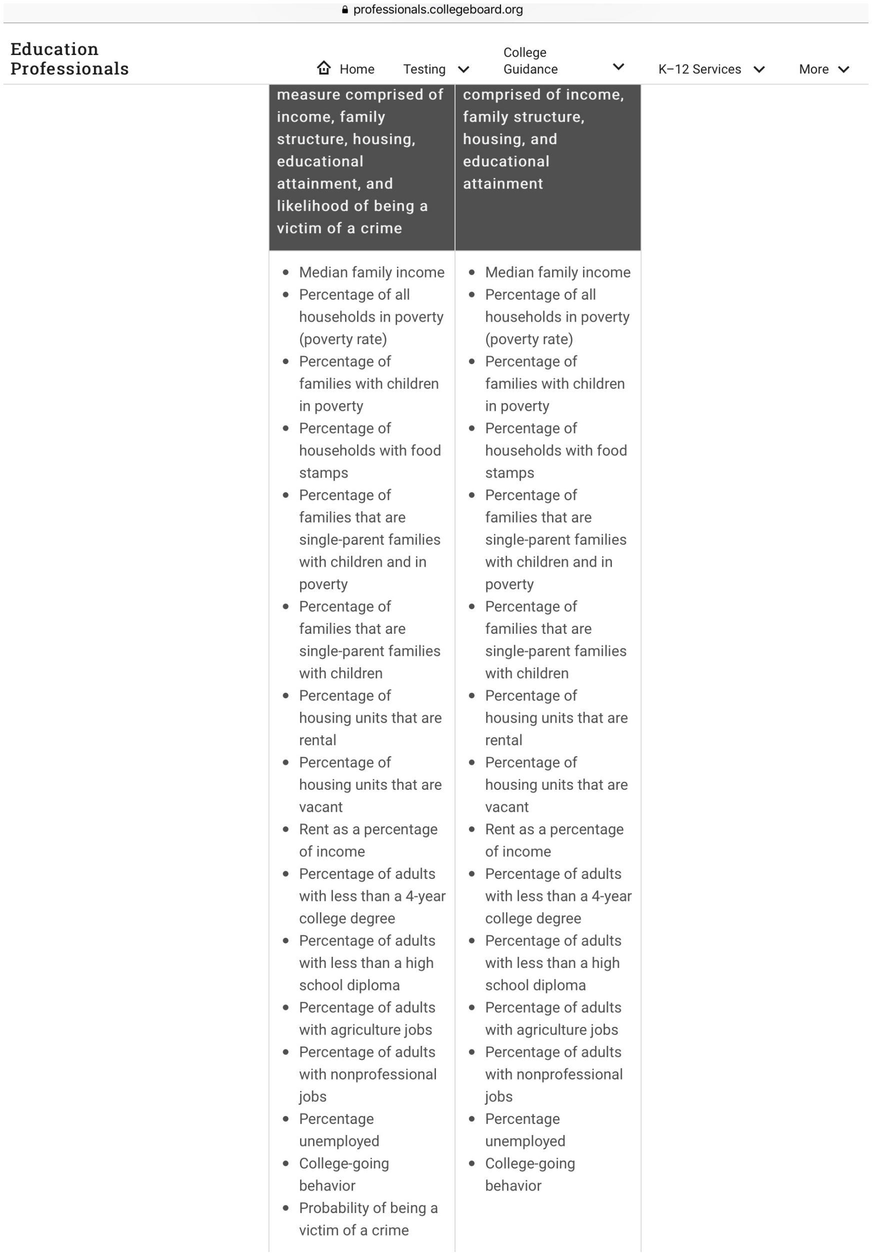
Figure 2. Environmental Dashboard Prototype

The tool was built so that information relevant to a particular application could be displayed by entering the applicant's ID at the top left. The top rows of the dashboard contain all data specific to an individual applicant; the remainder of the dashboard contains contextual information related to the applicant's high school or neighborhood. The specific data elements are listed in the paragraph that follows.

Applicant Information

Data about the individual applicant include:

based on a host of adverse risk factors in a student's environment. These factors look like the kind of data analysis seen on problematic third party data broker lists. For example, factors include if a child is likely to have been the victim of a crime, among many other factors.[30] In the screenshot below is a graph showing the *Data involved in the College Board adversity score.*



measure comprised of income, family structure, housing, educational attainment, and likelihood of being a victim of a crime	comprised of income, family structure, housing, and educational attainment
• Median family income	• Median family income
• Percentage of all households in poverty (poverty rate)	• Percentage of all households in poverty (poverty rate)
• Percentage of families with children in poverty	• Percentage of families with children in poverty
• Percentage of households with food stamps	• Percentage of households with food stamps
• Percentage of families that are single-parent families with children and in poverty	• Percentage of families that are single-parent families with children and in poverty
• Percentage of families that are single-parent families with children	• Percentage of families that are single-parent families with children
• Percentage of housing units that are rental	• Percentage of housing units that are rental
• Percentage of housing units that are vacant	• Percentage of housing units that are vacant
• Rent as a percentage of income	• Rent as a percentage of income
• Percentage of adults with less than a 4-year college degree	• Percentage of adults with less than a 4-year college degree
• Percentage of adults with less than a high school diploma	• Percentage of adults with less than a high school diploma
• Percentage of adults with agriculture jobs	• Percentage of adults with agriculture jobs
• Percentage of adults with nonprofessional jobs	• Percentage of adults with nonprofessional jobs
• Percentage unemployed	• Percentage unemployed
• College-going behavior	• College-going behavior
• Probability of being a victim of a crime	

[30] College Board ECD Detailed Data Description Page, Available at: https://professionals.collegeboard.org/environmental-context-dashboard/detailed-data-description.

Here are some of the problems with the score:

- The College Board adversity score not available to students or parents. Secret scores are not acceptable in this context; students need to be able to learn their scores, as do parents and household members, given that their activities are part of what is scored.
- There is no transparency in the score itself: what are the factors? What is the testing population? What is the model? How often is the model updated?
- There is no external government oversight for the score factors, development, and algorithmic fit.
- There are no due process standards created according to the ANSI Essential Guidelines for the use and application and interpretation of the score by colleges and universities, and other entities that may get the score.
- There are no controls on future uses. Will employers begin asking to see the adversity score? How long is the score kept? What is the policy regarding removal of a score? What is the policy regarding score accuracy?
- There is not a choice for students who may wish to have their socio-economic background be private and not considered when they apply to college. Why are we not giving students and parents the choice as to whether or not this score is even used and shared in the first place?

The adversity score raises profound privacy and ethical issues. It has an impact on a student's future profession, employment and life. This is an eligibility circumstance, and because this score concerns the reputation of a student, and perhaps even the student's parents or other family members, then this should be a score placed under the FCRA as an eligibility issue. Challenging ethical questions arise in the use of a contextual adversity score, and the challenges are heightened when a risk-based scoring categorization has contributed to a negative outcome. The FCRA would need to be extended to cover this new eligibility context.

What are the long term impacts of the adversity score? Will a student quit school due to discouragement at being classified in a certain way that demeaned them? Will students decide to not apply to college due to shame at their potential adversity score? These are just a few of the serious issues in the applications of AI techniques to the copious stores of learner data available for such analysis. A cohesive, non-secret policy in this area would be a good investment of effort and time.

Education has been for many decades in the United States a place where children from all walks of life can use the availability of a public school system to work hard and earn their place in the world. This is where the American dream can take place for kids who may not have been born in economic prosperity. However, prediction beyond academic achievement has entered education. The privacy implications of having life factors be provided to strangers without choice or transparency are no small matter.

I note here that the use of AI techniques applied to stores of learner data (learner analytics) has grown profoundly and is now an entire field of inquiry with substantial sophistication. Many studies exist on how to use AI on educational "big data," and there is little doubt that some of the education-focused analysis has proven invaluable. The privacy risks of these techniques have just begun to surface. Predictive categorization and / or classification of students based on their learning data is already sensitive. But far more problematic is the use of non-educational home data to score students regarding their life circumstance, and then to keep that score secret from them.

IV. Solutions

There are two key solutions available to solve the problems of data brokers and scoring.

FCRA- related solutions

First, unregulated forms of credit scoring, or household credit scoring, should be brought under the FCRA. If a marketing offer involves a financial product, such as a credit card, and the offer is based on a predictive score that functions as a credit score would in terms of predictive quality and accuracy, then those scores act in the same predictive way that credit scores would, and they should be treated as such. These marketing situations can have very meaningful impacts on consumers. Congress should investigate these products and determine with more specificity, and with information from industry, where the boundary lines are.

Currently, these products are opaque, and there are no regulatory requirements to ensure that the models are not biased, unfair, discriminatory, or otherwise constructed poorly.

Congress should call on these companies to immediately make consumers' unregulated credit scores available to them.

Second, Congress needs to launch a study commission to analyze and determine what new areas of eligibility need to be considered as falling under the FCRA. In our analysis, educational eligibility for college acceptance, when judged in part by a non-academic score such as the adversity score, should be covered under the FCRA. Although I discussed one primary exemplar, other scores of this type exist. This and other "new eligibility" scenarios will likely emerge in time as scoring gets more accurate and less expensive. But I request that Congress consider including educational eligibility circumstances, such as applying to college, in the definition of eligibility triggers in the FCRA.

Students and parents who are subject to adversity scores should immediately be able to see their scores. It is a deep unfairness that students cannot see their scores, when colleges are using this information to make important decisions affecting student's lives.

Standards - related solutions: Voluntary Consensus Standards

Due process standards have been a neglected aspect of solving complex privacy challenges. Voluntary consensus standards are a well-defined term of art, and law. A voluntary consensus standard is one that is developed or adopted by Standards Developing Organizations (SDOs), both domestic and international, according to strict consensus principles. Consensus standards contribute to regulatory quality because consensus-based SDOs must demonstrate adherence to the tenets of transparency, openness to participation by interested stakeholders, balance of representation, and due process, among other principles.[31]

In the United States, there are two critical definitional groundings for VCS:

1. The OMB Circular A-119: Federal Participation in the Development and Use of Voluntary Consensus Standards and in Conformity Assessment Activities,[32] (The National Technology Transfer and Advancement Act (NTTAA) codifies OMB Circular A-119.)
2. The ANSI Essential Requirements: Due Process requirements for American National Standards.[33]

In 1996, the National Technology Transfer and Advancement Act (NTTAA) (Pub. L. No. 104-113), codified OMB Circular A-119, Federal Participation in the Development and Use of Voluntary Consensus Standards and in Conformity Assessment Activities.[34] The NTTAA and OMB Circular A-119 established that Federal government agencies were to use voluntary consensus standards in lieu of government-unique standards except where voluntary consensus standards are inconsistent with law or otherwise impractical. The ANSI Essential Requirements set forth in detail the definitions and processes that comprise a "due process" standards setting body, and procedures.

[31] ANSI Essential Requirements: Due process requirements for American National Standards, Available at: https://share.ansi.org/Shared%20Documents/Standards%20Activities/American%20National%20Standards/Procedures%2C%20Guides%2C%20and%20Forms/ANSI-Essential-Requirements-2018.pdf. *See also*: U.S. Food and Drug Administration, Standards and Conformity Assessment Program, Available at: https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/standards-and-conformity-assessment-program-medical-devices#intro.

[32] OMB Circular A-119, "Federal Participation in the Development and Use of voluntary Consensus Standards and in Conformity Assessment Activities," 2016 Revision, 81 FR 4673 pages 4673-4674. Available at: https://www.nist.gov/sites/default/files/revised_circular_a-119_as_of_01-22-2016.pdf.

[33] ANSI Essential Requirements: Due process requirements for American National Standards, Available at: https://share.ansi.org/Shared%20Documents/Standards%20Activities/American%20National%20Standards/Procedures%2C%20Guides%2C%20and%20Forms/ANSI-Essential-Requirements-2018.pdf.

[34] National Technology Transfer and Advancement Act (NTTAA) (Pub. L. No. 104-113).

The most current definition of a standards body that creates voluntary consensus guidelines is as follows, as found in the 2016 revision of OMB Circular A-119:

> "Voluntary consensus standards body" is a type of association, organization, or technical society that plans, develops, establishes, or coordinates voluntary consensus standards using a voluntary consensus standards development process that includes the following attributes or elements:
>
> i. Openness: The procedures or processes used are open to interested parties. Such parties are provided meaningful opportunities to participate in standards development on a non-discriminatory basis. The procedures or processes for participating in standards development and for developing the standard are transparent.
>
> ii. Balance: The standards development process should be balanced. Specifically, there should be meaningful involvement from a broad range of parties, with no single interest dominating the decision-making.
>
> iii. Due process: Due process shall include documented and publically available policies and procedures, adequate notice of meetings and standards development, sufficient time to review drafts and prepare views and objections, access to views and objections of other participants, and a fair and impartial process for resolving conflicting views.
>
> iv. Appeals process: An appeals process shall be available for the impartial handling of procedural appeals.
>
> v. Consensus: Consensus is defined as general agreement, but not necessarily unanimity. During the development of consensus, comments and objections are considered using fair, impartial, open, and transparent processes.[35]

The idea of the FTC providing a safe harbor for business in the privacy sphere has continued to arise; but the FTC, and indeed all Federal agencies, must comply with the rules enshrined in the OMB Circular. Circular A-119 applies to all US Federal "agencies and agency representatives who use standards or conformity assessment and/or participate in the development of standards. "Agency" means any executive department, independent commission, board, bureau, office, government-owned or controlled corporation, or other establishment of the Federal government. It also includes any regulatory commission or board, except for independent regulatory commissions insofar as they are subject to separate statutory requirements regarding the use of

[35] OMB Circular A-119, "Federal Participation in the Development and Use of voluntary Consensus Standards and in Conformity Assessment Activities," 2016 Revision, 81 FR 4673 pages 4673-4674. Available at: https://www.nist.gov/sites/default/files/revised_circular_a-119_as_of_01-22-2016.pdf.

voluntary consensus standards. It does not include the Legislative or Judicial branches of the Federal government."[36]

The OMB Circular states that all Federal agencies[37] must use voluntary consensus standards (in lieu of government-unique standards) in procurement and regulatory activities, except "where inconsistent with law or otherwise impractical." Legislative and judicial branches of the federal government are not subject to OMB Circular A-119. However, the Circular does apply to all federal agencies, including law enforcement, national security, and other regulatory agencies such as the FBI, CIA, and NSA, HHS, the FTC, the FDA, and others. What is remarkable is not that such standards exist, but that in many if not most multistakeholder and legislative discussions around privacy, it has not been well-understood that they exist.

Our draft bill is included in its entirety in Appendix A.

VI. Conclusion

Rapid, widespread prediction of our behavior, intent, and even our future is on its way. If we are to have a trusted digital ecosystem, Congress will need to find effective solutions. Employing a combination of updated and expanded interpretations of eligibility under the FCRA to include application to educational institutions, as well as making the call that unregulated household credit reports are in fact credit reports would go far to begin to clarify the rules. And using a due process, fair standards setting process to determine specific guidance in hard-to-anticipate situations will help develop best practices and a more transparent dialogue between stakeholders.

The American public would like to enjoy the benefits of innovation secure in the knowledge that their personal information will not be misused by those who administer its collection, processing and dissemination. Recent experience in the U.S. has demonstrated that the American public's trust in a purely market-led approach to privacy is rapidly dwindling.[38] Affording all

[36] OMB Circular A-119, "Federal Participation in the Development and Use of voluntary Consensus Standards and in Conformity Assessment Activities," 2016 Revision, 81 FR 4673 pages 4673-4674. Available at: https://www.nist.gov/sites/default/files/revised_circular_a-119_as_of_01-22-2016.pdf.

[37] ANSI essential requirements can also fully apply to standards governing, for example, the FBI, CIA, and NSA in areas such as the voluntary sharing of information by businesses with law enforcement. The development of due process standards for this category of data flows and activity would be beneficial to all stakeholders, including the public, as these data flows are among the least understood aspects of today's data ecosystems.

[38] The Cambridge Analytica scandal that became headline news in 2018 highlighted consumer data privacy missteps such as data uses beyond what consumers understood, or potentially agreed to. See: Philip Bump. Everything you need to know about the Cambridge Analytica - Facebook debacle, Washington Post, March 19, 2018. Available at: https://www.washingtonpost.com/news/politics/wp/2018/03/19/everything-you-need-to-know-about-the-cambridge-analytica-facebook-debacle/?utm_term=.4172f3ec00cf.

stakeholders a "seat at the table" and meaningful input into standards of how data is used in a given ecosystem is a meaningful step toward rebuilding this trust. And trust is of central importance in digital ecosystems; without a basis for mutual trust, history has shown that deleterious consequences may ensue.

People care about how the systems that administer their personal information are governed, but they also want access to the economic prosperity that responsible innovation can provide. There is a meaningful opportunity to update the American tradition of transparent, accountable and inclusive “industrial legislatures” to insure its relevance in the world of knowledge governance.

That goal could be achieved by enacting privacy and information governance legislation that includes giving the FTC the power to recognize compliance based on voluntary, consensus standards within the OMB Circular A-119 framework as a tool to increase trust amongst stakeholders, encourage meaningful dialogue, and move privacy thought into a modern technological context, with much-needed protections.

Appendix A: . Full Text of Discussion Draft Bill

CONSUMER PRIVACY AND DATA SECURITY STANDARDS ACT OF 2019

PREAMBLE

Because information is the basis of knowledge, and knowledge is the basis of competitive advantage in local, national and global markets, this law establishes a fair, inclusive, and transparent process to govern the collection, use, maintenance, and disclosure of personal information.

In order for public and private sector institutions to fulfill their mandates to serve the citizens of the United States, these institutions must earn the trust of the American people by demonstrating that they access, use, maintain and disclose consumers’ personal information in a manner that respects reasonable consumer interests in privacy and data security.

Precisely what constitutes an appropriate balance in the interests of institutions and individuals regarding personal information varies, depending on the sensitivity of the personal information, the importance of the institutional need, and the context in which the information is used. At times, the appropriate balance can be reflected in sector-specific statutes and regulations. At other times, more context-specific and granular governance frameworks are needed.

The American system of voluntary consensus standards established by the private sector through recognized fair, inclusive, transparent, procedures that comport with due process, in which the interests of all principal stakeholders are accounted for, has provided effective solutions to similar problems for more than one hundred years.

When public and private sector institutions make effective use of voluntary consensus standards established through due process procedures to implement solutions to urgent problems, the benefits accrue not only to private and public institutions, but also to the American people.

Section 1. Definitions

(a) "Personal Information" refers to information that can be reasonably expected to distinguish or trace an individual's identity, either alone or when combined with other information that is linked or linkable to a specific individual.

(b) "Covered entity" refers to a person, partnership, association or organization over which the Federal Trade Commission has authority pursuant to section 5(a)(2) of the Federal Trade Commission Act (15 U.S.C. 45(a)(2)), and operates a website located on the internet or an online service and who collects, uses, maintains or discloses personal information from or about individuals, or on whose behalf such information is collected, used, maintained or disclosed, where such website or online service is operated for commercial purposes, including any entity that buys and sells consumer data without direct consumer interaction, and any entity offering products or services for sale through that website or online service. Notwithstanding the limitations in the Federal Trade Commission Act on Commission authority with respect to common carriers, a covered entity also includes common carriers subject to the Communications Act of 1934 (47 U.S.C. 151 et seq.) and Acts amendatory thereof and supplementary thereto.

(c) "Commission" refers to the Federal Trade Commission.

(d) "Standard" includes all of the following:

(1) Common and repeated use of rules, conditions, guidelines or characteristics for products or related processes and production methods, and related management systems practices.

(2) The definition of terms; classification of components; delineation of procedures; specification of dimensions, materials, performance, designs, or operations; measurement of quality and quantity in describing materials, processes, products, systems, services, or practices; test methods and sampling procedures; or descriptions of fit and measurements of size or strength.

(e) "Standard" does not include the following:

(1) Professional standards of personal conduct.

(2) Institutional codes of ethics.

(f) "Voluntary consensus standards" are due process standards developed or adopted by voluntary consensus standards bodies as set forth in this Act.

(g) "Voluntary consensus standards bodies" are organizations which plan, develop, establish, or coordinate voluntary consensus standards using agreed-upon due process procedures. A voluntary consensus standards body is defined by the following attributes:

  (1) Openness
  (2) Balance of interest.
  (3) Due process.
  (4) An appeals process.
  (5) Consensus, which is defined as general agreement, but not necessarily unanimity, and includes a process for attempting to resolve objections by interested parties, as long as all comments have been fairly considered, each objector is advised of the disposition of his or her objection(s) and the reasons why, and the consensus body members are given an opportunity to change their votes after reviewing the comments.

Section 2. Regulation of unfair and deceptive acts and practices in connection with collection and use of personal information.

(a) Acts Prohibited—In General—It is unlawful for a covered entity to collect, use, maintain, or disseminate personal information in a manner that violates the regulations prescribed by the Federal Trade Commission under subsection (d) of this Section.

(b) Enforcement—A violation of this Act or a regulation promulgated under this Act shall be treated as a violation of a regulation under section 18(a)(1)(B) of the Federal Trade Commission Act (15 U.S.C. 57a(a)(1)(B)) regarding unfair or deceptive acts or practices.

(c) Powers of Commission—Except as provided in subsection (a), the Federal Trade Commission shall enforce this Act and the regulations promulgated under this Act in the same manner, by the same means, and with the same jurisdiction, powers, and duties as though all applicable terms and provisions of the Federal Trade Commission Act (15 U.S.C. 41 et seq.) were incorporated into and made a part of this Act.

(d) Regulations—

(1) In general—Not later than 1 year after the enactment of this Act, the Commission shall promulgate under section 553 of title 5 regulations that require covered entities to collect, use, maintain and disclose personal information:

A. In accordance with reasonable security measures to protect its confidentiality, security, and integrity; and
B. In accordance with reasonable consumer interests in privacy.

(2) Such regulations may not impose direct or indirect liability on any covered entities for making a voluntary or compelled disclosure of personal information to a federal, state local or tribal law enforcement,

national security, regulatory or other governmental agency for an authorized governmental purpose.

(3) Before issuing a regulation for data security and privacy, or approving any voluntary consensus standard, the Commission shall consult with the Attorney General, and with other federal agencies, as appropriate, to ensure that the standard does not hamper competition, or restrict access to personal information for authorized law enforcement, national security, or other lawful, authorized governmental purposes.

(4) Enforcement—Subject to Section 3 of this title, a violation of a regulation prescribed under subsection (d) of this Section shall be treated as a violation of a rule defining an unfair or deceptive act or practice prescribed under Section 18(a)(1)(B) of the Federal Trade Commission Act, and any person, partnership, or corporation who violates a such a regulation shall forfeit and pay to the United States a civil penalty of not more than $10,000 for each violation, which shall accrue to the United States and may be recovered in a civil action brought by the Attorney General of the United States.

(5) Inconsistent State law—No State or local government may impose any liability for commercial activities or actions by a covered entity in connection with an activity involving personal information covered by the regulations promulgated by the Commission under this Section 2 of this Act or by a voluntary consensus standard approved by the Commission pursuant to Section 3 of this Act.

Section 3—Safe Harbors

(a) In prescribing regulations under this title, the Commission shall provide incentives for adoption of voluntary consensus standards, as set forth in this Act, by covered entities to implement the protections described in Section 2(d)(A) and (B) of this title.

(b) Deemed compliance—Such incentives shall include provisions for ensuring that a covered entity will be deemed to be in compliance with the requirements of the regulations issued under Section 2(d)(1) of this title if the covered entity follows a voluntary consensus standard, as set forth in this Act, that, after notice and comment, is approved by the Commission pursuant to the provisions of this Act, and found by the Commission to:

(1) meet the requirements of the regulations issued under Section 2(d)(1) of this title;

(2) be the result of due process procedures set forth in Section 4 of this Act; and

(3) appropriately balance the interests of all the stakeholders, including individuals and businesses, organizations, and other entities making lawful uses of the personal information.

(c) Expedited response to requests—The Commission shall act upon requests for safe harbor treatment within 180 days of the filing of the request, and shall set forth in writing its conclusions with regard to such requests.

(d) Appeals—Final action by the Commission on a request for approval of voluntary consensus standards, or the failure to act within 180 days on a request for approval of the voluntary consensus standard, submitted under subsection (b) may be appealed to a district court of the United States of appropriate jurisdiction as provided for in section 706 of title 5.

Section 4—Voluntary Consensus Standards

(a) Guidelines—A covered entity may satisfy the requirements of regulations issued under Section 2(d)(1) of this title by following a voluntary consensus standard, issued by the National Institute of Standards and Technology or by other voluntary consensus standards bodies, pursuant to this Act, and approved by the Commission under Section 3(a) and (b) of this Title.

(b) Voluntary Consensus Standards—Process—To be eligible for safe harbor status under Section 3(a) and (b), a voluntary consensus standard must be the result of a process:

(1) That follows the principles of consensus, due process and openness, depending heavily upon data gathering and compromise among a diverse range of stakeholders;

(2) That ensures that access to the standards setting process, including an appeals mechanism, was made available to anyone directly or materially affected by the standard under development;

(3) That provides all such stakeholders (including individuals, businesses, government agencies, and other entities such as consumer groups and civil society organizations), a reasonable opportunity to voluntarily contribute their knowledge, talents and efforts to the standard's development;

(4) That consistently adheres to essential due process procedures that served and protected the public interest in openness, balance, consensus and other due process safeguards;

(5) That is equitable, accessible and responsive to the requirements of all interested and affected parties;

(6) That includes a reasonable opportunity for broad-based public review and comment on draft standard, with consideration of and response to the comments submitted by voting members of the relevant consensus body and by public review of the comments, followed by incorporation of the approved changes into a draft standard; and

(7) That includes a right to appeal by any participant that believed that due process principles were not sufficiently respected during the standards development in accordance with the procedures of the standard setting organization.

(c) Voluntary Consensus Standards—To be eligible for safe harbor status in connection with regulations issued under Section 2(d)(1)(B), a voluntary consensus standard must

(1) Establish a clear nexus to the collection, use, maintenance and disclosure of the personal information it governs;

(2) Reasonably identify the interests of the stakeholders (including individual consumers, businesses and governments);

(3) Reasonably identify the benefits and material risks to the stakeholders arising from the proposed collection, use, maintenance and disclosure of the personal information involved;

(4) Reasonably ensure that the benefits from the proposed collection, use, maintenance and disclosure of the personal information outweigh risks, after such risks are mitigated by technological, operational or other means, presenting the supporting analysis for such assessment of costs and benefits fairly, symmetrically, and with an appropriate level of granularity;

(5) Reasonably addressing any alternatives, after disclosing all key assumptions, data and models;

(6) Reasonably addressing the requirements by the regulations promulgated under Section 2(d)(1)(B) of this Title by specifying routine uses for which consent is not required when the use and disclosure of the personal information is compatible with the purposes for which the information was collected, and non-routine uses, in which case procedures must be established to reasonably protect the interests of the individual, including as appropriate:

    (A) Written consent by the individual prior to use of the information for the non-routine purpose;

    (B) Transparency regarding information collection, use, maintenance, and dissemination;

    (C) Procedures for consumers to access and correct information material to decisions affecting their legitimate interests; and

    (D) Redress for actual damages caused by a business's failure to adhere to the standard.

(7) Establish reasonable internal controls and accountability to ensure effective implementation of the voluntary consensus standard by the covered entity.

.

Appendix B

The Scoring of America

(Attached)