

**DATA BROKERS AND THE IMPACT ON FINANCIAL
DATA PRIVACY, CREDIT, INSURANCE, EMPLOY-
MENT, AND HOUSING**

HEARING
BEFORE THE
COMMITTEE ON
BANKING, HOUSING, AND URBAN AFFAIRS
UNITED STATES SENATE
ONE HUNDRED SIXTEENTH CONGRESS

FIRST SESSION

ON

EXAMINING DATA BROKERS' INDUSTRY PRACTICES AND STANDARDS
AND THE IMPACT THEY HAVE ON ACCESS TO, AND ELIGIBILITY FOR,
CREDIT, INSURANCE, EMPLOYMENT, AND HOUSING

—————
JUNE 11, 2019
—————

Printed for the use of the Committee on Banking, Housing, and Urban Affairs



Available at: <https://www.govinfo.gov/>

U.S. GOVERNMENT PUBLISHING OFFICE

WASHINGTON : 2021

COMMITTEE ON BANKING, HOUSING, AND URBAN AFFAIRS

MIKE CRAPO, Idaho, *Chairman*

RICHARD C. SHELBY, Alabama	SHERROD BROWN, Ohio
PATRICK J. TOOMEY, Pennsylvania	JACK REED, Rhode Island
TIM SCOTT, South Carolina	ROBERT MENENDEZ, New Jersey
BEN SASSE, Nebraska	JON TESTER, Montana
TOM COTTON, Arkansas	MARK R. WARNER, Virginia
MIKE ROUNDS, South Dakota	ELIZABETH WARREN, Massachusetts
DAVID PERDUE, Georgia	BRIAN SCHATZ, Hawaii
THOM TILLIS, North Carolina	CHRIS VAN HOLLEN, Maryland
JOHN KENNEDY, Louisiana	CATHERINE CORTEZ MASTO, Nevada
MARTHA MCSALLY, Arizona	DOUG JONES, Alabama
JERRY MORAN, Kansas	TINA SMITH, Minnesota
KEVIN CRAMER, North Dakota	KYRSTEN SINEMA, Arizona

GREGG RICHARD, *Staff Director*

JOE CARAPIET, *Chief Counsel*

BRANDON BEALL, *Professional Staff Member*

ALEXANDRA HALL, *Professional Staff Member*

LAURA SWANSON, *Democratic Staff Director*

COREY FRAYER, *Democratic Professional Staff Member*

CAMERON RICKER, *Chief Clerk*

SHELVIN SIMMONS, *IT Director*

CHARLES J. MOFFAT, *Hearing Clerk*

JIM CROWELL, *Editor*

C O N T E N T S

TUESDAY, JUNE 11, 2019

	Page
Opening statement of Chairman Crapo	1
Prepared statement	30
Opening statements, comments, or prepared statements of:	
Senator Brown	3
Prepared statement	31

WITNESSES

Alicia Puente Cackley, Ph.D., Director, Financial Markets and Community Investment, Government Accountability Office	4
Prepared statement	32
Responses to written questions of:	
Senator Menendez	163
Senator Warren	163
Senator Schatz	166
Senator Cortez Masto	169
Pam Dixon, Executive Director, World Privacy Forum	5
Prepared statement	49
Responses to written questions of:	
Senator Menendez	171
Senator Warren	176
Senator Schatz	183
Senator Cortez Masto	186

ADDITIONAL MATERIAL SUPPLIED FOR THE RECORD

Letter submitted on behalf of Acxiom by Jordan Abbott, Chief Ethics Office	202
Letter and responses to written questions of the Banking Committee sub- mitted by Bob Liodice, Chief Executive Office, Association of National Ad- visers	204
Letter submitted by CoreLogic	210
Letter submitted by Jim Nussle, President & CEO, Credit Union National Association (CUNA)	211
Letter submitted by Brad Thaler, Vice President of Legislative Affairs, Na- tional Association of Federally-Insured Credit Unions	213

DATA BROKERS AND THE IMPACT ON FINANCIAL DATA PRIVACY, CREDIT, INSURANCE, EMPLOYMENT, AND HOUSING

TUESDAY, JUNE 11, 2019

U.S. SENATE,
COMMITTEE ON BANKING, HOUSING, AND URBAN AFFAIRS,
Washington, DC.

The Committee met at 10:03 a.m. in room SD-538, Dirksen Senate Office Building, Hon. Mike Crapo, Chairman of the Committee, presiding.

OPENING STATEMENT OF CHAIRMAN MIKE CRAPO

Chairman CRAPO. This hearing will come to order.

Providing testimony to the Committee today are experts who have researched and written extensively on big data: Dr. Alicia Cackley, the Director of Financial Markets and Community Investment at the Government Accountability Office; and Ms. Pam Dixon, Executive Director of the World Privacy Forum. We appreciate both of you being here.

As a result of an increasingly digital economy, more personal information is available to companies and others than ever before. I have been troubled by Government agencies' and private companies' collection of personally identifiable information for a long time.

There have been many questions about how individuals' or groups of individuals' information is collected, with whom it is shared or sold, how it is used, and how it is secured.

Private companies are collecting, processing, analyzing, and sharing massive data on individuals for all kinds of purposes. Even more troubling is that the vast majority of Americans do not even know what data is being collected, when it is being collected, how it is being collected, by whom, and for what purpose.

In particular, data brokers and technology companies, including large social media platforms and search engines, play a central role in gathering vast amounts of personal information and often without interacting with individuals, specifically in the case of data brokers.

In 2013, the GAO issued a report on information resellers, which includes data brokers, and the need for the consumer privacy framework to reflect changes in technology in the marketplace.

The report noted that the current statutory consumer privacy framework fails to address fully new technologies and the growing marketplace for personal information.

The GAO also provided several recommendations to Congress on how to approach the issue to provide consumers with more control over their data.

In 2018, 5 years later, GAO published a blog summarizing its 2013 report, highlighting the continued relevance of the report's findings.

The Federal Trade Commission also released a report in 2014 that emphasized the big role of data brokers in the economy. The FTC observed in its report that "data brokers collect and store billions of data elements covering nearly every U.S. consumer," and that "data brokers collect data from numerous sources, largely without consumers' knowledge."

In her report "The Scoring of America," Pam Dixon discusses predictive consumer scoring across the economy, including the big role that data brokers play. She stresses that today no protections exist for most consumer scores, similar to those that apply to credit scores under the Fair Credit Reporting Act.

Dixon says, "Consumer scores are today where credit scores were in the 1950s. Data brokers, merchants, government entities, and others can create or use a consumer score without notice to consumers."

Dr. Cackley has also issued several reports on consumer privacy and technology, including a report in September 2013 on information resellers, which includes data brokers. She says in her report that the current consumer privacy framework does not fully address new technologies and the vastly increased marketplace for personal information. She also discusses potential gaps in current Federal law, including the Fair Credit Reporting Act.

The Banking Committee has been examining the data privacy issue in both the private and public sectors, from regulators to financial companies, to other companies who gather vast amounts of personal information on individuals or groups of individuals to see what can be done through legislation, regulation, or by instituting best practices.

Enacted in 1970, the Fair Credit Reporting Act is a law in the Banking Committee's jurisdiction which aims to promote the accuracy, fairness, and privacy of consumer information contained in the files of consumer reporting agencies. Given the exponential growth and use of data since that time and the rise of entities that appear to serve a similar function as the original credit reporting agencies, it is worth examining how the Fair Credit Reporting Act should work in a digital economy.

During today's hearing, I look forward to hearing more about the structure and practices of the data broker industry and technology companies, such as large social media platforms; how the data broker industry has evolved within the development of new technologies, and their interaction with technology companies; what information these entities collect, how it is collected, and whom it is shared with and for what purposes; what gaps exist in Federal privacy law; and what changes to Federal law should be considered to give individuals real control over their data.

I appreciate each of you joining us today and look forward to getting some further information about these questions.

Senator Brown.

OPENING STATEMENT OF SENATOR SHERROD BROWN

Senator BROWN. Thank you, Mr. Chairman. I appreciate your continuing these important, bipartisan efforts to protect Americans' sensitive personal information.

We are looking today at a shadowy industry known as "data brokers." Most of you probably have not heard of these companies. The biggest ones include names like Acxiom, CoreLogic, Spokeo, and ZoomInfo—and maybe one you have heard of, Oracle. According to some estimates, 4,000 of these companies collect and sell private information, but, stunningly—and I am not sure I have ever used that word in this Committee—stunningly, not one of them has been willing to show up and speak in front of this Committee today. Not one.

These companies expect to be trusted with the most personal and private information you could imagine about millions of Americans. They are not even willing to show up and explain how their industry works. Some define this as cowardice. It is hard to disagree with that. I think it tells you all you need to know about how much they want their own faces and names associated with that industry.

As Maciej Ceglowski told us at our last hearing, "the daily activities of most Americans are now tracked and permanently recorded by automated systems at Google or Facebook."

Most of that private activity is not useful without data that anchors it to the real world. Facebook, Google, and Amazon want to know where you are using your credit cards, where you buy your brand-name appliances, if you are recently divorced, and how big your life insurance policy is—the kind of data that big tech gets from data brokers. They then combine it with your social media activity to feed into their algorithms.

You might have noticed it seems like every product or service you buy comes with a survey or a warranty card that asks for strangely personal information. Why are all these nontech companies so interested in your data?

It is simple: Data brokers will pay these companies for any of your personal information they can get their hands on so they can turn around and sell it to Silicon Valley. It is hard for ordinary consumers to have any power when, unbeknownst to them, they are actually the product bought and sold.

It reminds me of a time when corporations that had no business being in the lending industry decided to start making loans and selling them off to Wall Street. We know what happened. Manufacturers or car companies decided that consumer credit would be a great way to boost their profits. When big banks and big tech are willing to pay for something, everyone else will find a way to sell it to them, often with devastating results.

For example, Amazon is undermining retailers and manufacturers across the country through anticompetitive practices. At the same time, it scoops up information from the very businesses it is pushing out of the market.

Then there is Facebook, almost single-handedly undermining the profitability of newspapers across the country. It also gobbles up personal information that the New York Times allows data brokers to collect from its readers.

Just like in the financial crisis, a group of shadowy players sits at the center of the market, exercising enormous influence over consumers and the economy while facing little or no rules at all. Then they do not show up.

Chairman Crapo and I are committed to shining a light on these companies and keeping an unregulated data economy from spiraling out of control. Yesterday it was reported that a Department of Homeland Security contractor allowed unauthorized access to photos of travelers and their license plates to be exposed to potential identity thieves.

One of the principal differences between the two political parties in this town is the suspicion that Democrats have of private power and suspicion Republicans typically have of Government power. I think you are seeing two parties come together on our suspicion of what these data brokers are doing.

The Chairman and I agree that protecting sensitive information like this is timely and important. I look forward to the witnesses' testimony.

Thanks.

Chairman CRAPO. Thank you, Senator Brown, and I appreciate our partnership on this issue.

We will go in the order I introduced you, and, Dr. Cackley, you may begin. But before you do, let me just remind both of you that we would like you to keep your initial remarks to 5 minutes so that we can have plenty of time for the Senators to engage with you.

Dr. Cackley.

**STATEMENT OF ALICIA PUENTE CACKLEY, Ph.D., DIRECTOR,
FINANCIAL MARKETS AND COMMUNITY INVESTMENT, GOVERNMENT ACCOUNTABILITY OFFICE**

Ms. CACKLEY. Thank you. Chairman Crapo, Ranking Member Brown, and Members of the Committee,

I am pleased to be here today to discuss GAO's work on consumer privacy and information resellers, also known as "data brokers."

My remarks are primarily based on our September 2013 report on privacy issues related to information resellers, as well as more recent work on internet privacy, data protection, facial recognition, and financial technology.

My statement will focus on two main issues: the lack of an overarching Federal privacy law and gaps that exist in the current consumer privacy framework.

No overarching Federal privacy law governs the collection, use, and sale of personal information among private sector companies, including information resellers. There are also no Federal laws designed specifically to address all the products sold and information maintained by information resellers. Instead, Federal privacy laws covering the private sector are narrowly tailored to specific purposes, situations, types of information, or entities, such as data related to financial transactions, personal health, and eligibility for credit.

For example, the Fair Credit Reporting Act requires that sensitive consumer information be protected and restricts how it is shared. But the law only applies to information used to determine

eligibility for things like credit, insurance, and employment. Similarly, the Gramm-Leach-Bliley Act restricts how certain financial information is shared, but it only applies to entities that fall under the law's specific definition of a "financial institution." Other privacy statutes address other specific circumstances, but there is no Federal statute that comprehensively addresses privacy issues in the private sector.

GAO has stated previously that gaps exist in the U.S. consumer privacy framework. We have reported that Federal law provides consumers with limited ability to access, control, and correct their personal data, particularly data used for marketing purposes. Similarly, individuals generally cannot prevent their personal information from being collected, used, and shared. Yet information that resellers collect and share for marketing purposes can be very personal or sensitive. For example, it can include information about physical and mental health, income and assets, political affiliations, and sexual habits and orientation.

Another area where there are gaps in the consumer privacy framework is with respect to new technologies. For example, Federal law does not address expressly when companies can use facial recognition technology to identify or track individuals, nor does it address when consumer knowledge or consent should be required for its use. Similarly, no Federal privacy law explicitly addresses the full range of practices for tracking or collecting data from consumers' online activity or the application software for mobile devices. And the rise of financial services technologies, known as "FinTech," raises new privacy concerns, for example, because new sources of personal data are being used to determine creditworthiness.

In summary, new markets and technologies have vastly changed the amount of personal information private companies collect and how they use it. But our current privacy framework does not fully address these changes. Laws protecting privacy interests are tailored to specific sectors and uses, and consumers have little control over how their information is collected, used, and shared with third parties for marketing purposes. As a result, the current privacy framework warrants reconsideration by Congress in relation to consumer interests, new technologies, and other issues.

Chairman Crapo, Ranking Member Brown, and Members of the Committee, this concludes my statement. I would be pleased to answer any questions you may have.

Chairman CRAPO. Thank you.

Ms. Dixon.

**STATEMENT OF PAM DIXON, EXECUTIVE DIRECTOR, WORLD
PRIVACY FORUM**

Ms. DIXON. Thank you. Chairman Crapo, Ranking Member Brown, and Members of the Committee, thank you for your invitation and for the opportunity to talk about something very, very meaningful today: the Fair Credit Reporting Act, data brokers, and privacy.

Fifty years ago, this Committee struck a blow for consumers for transparency and for fairness when it passed the Fair Credit Reporting Act. This Committee talked with stakeholders. They found

best practices. And before the famous HEW Report came out, the Committee report that defined what became fair information practices, this Committee created the Fair Credit Reporting Act. It was and still is the most important American privacy law that we have. But it is not as important as it was. There are three reasons why.

First, credit scores and other scores are being sold and used in consumers' lives, and these are unregulated.

Second, the technology of prediction, what can be called "predictive analytics," otherwise known as AI and machine learning, this technology and suite of technologies has advanced profoundly, and especially in the last 3 to 4 years, new kinds of predictive abilities have come forth, and we have new levels of accuracy in prediction, so that what used to be the accuracy of the credit score now is also the accuracy of an unregulated credit score, and this introduces new problems for consumers.

Third, these scores are created without due process for consumers. How on Earth do we deal with this? This is why Congress must expand the Fair Credit Reporting Act to regulate currently unregulated scores, especially in the financial sector, that are being used in meaningful ways in consumers' lives.

We have other solutions to discuss and other issues to discuss. I look forward to your questions. Thank you.

Chairman CRAPO. Thank you very much, Ms. Dixon.

I would like to ask each of you to answer my first three questions, and then I want to get into more discussion. But I would like you, if you possibly can, to limit your answers to yes or no answers to the first three. I know you will be tempted to elaborate, but I will give you that chance.

First, do you agree that data brokers collect and process vast amounts of personal information on nearly every American to the extent that they hold more information about individuals than the U.S. Government or traditional credit bureaus?

Ms. DIXON. Yes.

Ms. CACKLEY. Yes.

Chairman CRAPO. Second, do you both agree that most Americans have no knowledge of these activities and in most cases no rights to access, correct, or control the information collected about them?

Ms. DIXON. Yes.

Ms. CACKLEY. Yes.

Chairman CRAPO. And then, third, can certain processing and uses of this information have significant impact on their financial lives?

Ms. DIXON. Yes. Absolutely.

Ms. CACKLEY. Yes.

Chairman CRAPO. All right. Now we will get to where you can elaborate. You have both authored reports, as the FTC in 2014, that highlight the gaps in the Fair Credit Reporting Act and other privacy laws. You have both testified about that in your introductory remarks. These gaps allow data brokers to evade certain requirements that should be imposed on them.

What are the steps that we can take? You indicated, Ms. Dixon, that we need to expand the Fair Credit Reporting Act, and you es-

entially said the same thing, Dr. Cackley. But what specifically does this Committee need to do with regard to that?

Ms. DIXON. Thank you. In regards to the Fair Credit Reporting Act, I think very small changes would be very meaningful. Let me give you an example. Right now, as you know, as you well know, the Fair Credit Reporting Act in regards to credit scores applies to individuals. So when we are—you know, that is regulated at the individual level.

However, if you look at the new forms of credit scores that are available, they are scored at the household level where the Fair Credit Reporting Act does not apply. So you take a ZIP+4, and you score a household and give them, let us say, a score of 720. The household has a very accurate score of 720. Then that becomes an unregulated form of credit score. And, you know, 10 years ago, these scores were quasi-accurate. That has changed.

Chairman CRAPO. Thank you.

Dr. Cackley?

Ms. CACKLEY. So the Fair Credit Reporting Act has a certain number of elements to it that are very helpful. It gives consumers access, control, the ability to correct information, and safeguards privacy. But it only applies in certain situations for eligibility decisions. It would be possible to think about looking at a broader set of personal sensitive information that the Fair Credit Reporting Act could cover that would give consumers more of those things, access, control, ability to correct, over more personal sensitive information than is currently available.

Chairman CRAPO. All right. And I am going to use the term—well, Ms. Dixon, you used the term “unregulated credit scores.” There is a set of data that is collected about individuals and, as you indicate, households, and this data is turned into some kind of an analysis that allows those who use the data to influence and manipulate individuals in the marketplace.

Historically, as you have both indicated, the Fair Credit Reporting Act has focused primarily on credit bureaus, but the scope of who is collecting this data and how it is being used has exploded, as you both also discussed.

The question I have is: Isn't this unregulated score that we are talking about that is created for people and then managed by AI, isn't that impacting people's credit? Isn't it impacting their financial decisions? Isn't it significantly focused on that type of influence and manipulation of individuals?

Ms. CACKLEY. I think it certainly can be. The scores may not be credit scores, but they may apply to decisions that companies are making about what kinds of products they offer people, and at what price they offer things. This is based on a score that the consumer does not necessarily see, cannot tell is correct, or cannot make any attempt to improve if they do not even know it exists.

Chairman CRAPO. And to influence them to make such a transaction. I will let you go ahead, Ms. Dixon. I am running out of time here, but go ahead, please.

Ms. DIXON. Thank you. We call any score that is not regulated by the Fair Credit Reporting Act “consumer scores,” and we define that. It is in the written testimony. Consumer scores are quite dangerous when they are used in eligibility circumstances.

So, for example, the line between a lead generation, which is allowable—you do not have to pull a credit score to create a lead generation for a marketing product or a financial product. However, if you are just maybe marketing a financial product and you have something that is equivalent in accuracy to a credit score, all of a sudden this changes the equation. There is not even a micrometer in between, you know, what a regulation would be and a nonregulated score.

So if you have essentially something that looks like a credit score and that acts like the credit score and is being used like the credit score, well, it is the same thing as a duck. If it quacks, it is a duck.

So I think we have to look at the financial products that are being marketed with quasi-credit scores very closely. That is of high concern. But there are other categories. In “The Scoring of America,” we identified literally hundreds of types of scores: consumer lifetime value scores where consumers are segmented according to how valuable they are in terms of their purchasing power. There are frailty scores, which is more of a medical score. But the scores abound, and the concern I have is when people lose opportunities that are meaningful in their lives, for example, scores that are used in eligibility circumstances not described by the Fair Credit Reporting Act, such as admissions to colleges and what-not, imagine having a wonderful high school background and working very hard to achieve the American dream, and then all of a sudden some score says that you will not be as qualified a candidate, having nothing to do with your academic achievements but just somehow with maybe the neighborhood you grew up in. I find this disturbing.

Chairman CRAPO. Agreed.

Senator Brown.

Senator BROWN. Thank you, Mr. Chairman.

Ms. Dixon, you noted that tens of thousands of consumers’ scores affecting millions and millions of consumers are used to predict our behaviors, our secret, as you said. Are you surprised that Chairman Crapo was not able nor were we able to bring in data brokers to speak and testify? Are you surprised they were not willing to testify about their business practices before this Committee today?

Ms. DIXON. Actually, I am surprised, and I am actually—I wish they were here, and I wish the credit bureaus were here as well, because we need to have good industry step forward and to give us their best practices that they use. If there is no good industry to step forward with best practices, then this Committee cannot rearticulate what it did 50 years ago. And I do not understand why these industries are not willing to discuss what is happening, and I also do not understand why we cannot see our scores. Why?

Senator BROWN. I am not sure that they did not show up. I guess I would like to—I am not sure I have done this before either. I would like to ask anybody in the room that represents the data brokers to raise their hand. Lobbyists, lawyers, people paid by the data broker industry, any of you here? Any of you here that want to raise your hand? I guess is the question.

OK. And if you are, I mean, I will give you an opportunity of a lifetime. If you are, we will set up a different chair, and you can sit next to Ms. Dixon and Dr. Cackley. OK. All right. I guess no

surprise there, Mr. Chairman, and that does illustrate how—because I know they are watching. I mean, this is really important to their industry. It is very important to their bottom line, whether they are watching here or whether they are watching live stream. But we will move on.

Ms. Dixon, it seems that data predictions create a vicious cycle where the predictions end up often dictating the outcomes. For example, could people who have been systematically targeted by predatory lenders, having lower credit scores, therefore be likely only to see advertisements for other predatory financial products? I assume that happens. Are there other examples you can think of quickly?

Ms. DIXON. Yes, the predatory example is one we get phone calls about in our office from people who received advertisements for financial products, and they did not understand that they could have gone out on the market and affirmatively looked for the best offer. So these predatory marketing devices based on unregulated scores are very significant.

Other significant scores are scores that predict repayment of debt. So, for example, it is the poorest consumers who are targeted the most for debt repayment, all sorts of things like this. The consumer lifetime value scores impact how well you are treated by businesses, by how long you are standing in line, but the most meaningful circumstances that I can think of is when kids are applying to schools and they are getting scores that dictate whether or not they are going to be accepted to a school based not on their academics but based on all of these other things, like a pseudo credit score, like what neighborhood they grew up in. There are neighborhood risk scores which are the modern-day redlining, and I find them deeply objectionable because if we are going to be scored by where we live, how have we advanced and how have all the laws that have been meant to protect from such things, how are they operating if this is still happening today?

Senator BROWN. Thank you for that. So companies that—particularly your analogy to redlining, bank redlining, insurance redlining, now these companies redlining, are you worried that companies would offer discounts for products and services in exchange for sensitive data, which would lead—you sort of implied this—to a two-tiered system where the wealthy can afford privacy and everyone else will have to sacrifice sensitive information to get access to basic internet services?

Ms. DIXON. That is certainly part of it. I think it goes even more broadly than that. One of the big issues is that you get locked into a filter bubble of sorts, a marketing bubble, and it is not that people mean to get locked into these, but if you are receiving offers, especially for financial tools and services, and a consumer does not go outside of the offers they receive, they can pay more for autos; they can pay more for products; they can pay more for, for example, a TV. Simple things. But if you are a consumer on a fixed income, a television that costs \$2,000 instead of \$200 makes a meaningful difference in a person's life. That is what worries me the most.

Senator BROWN. There is one follow-up, not a question but a comment, Mr. Chairman. Thanks for your forbearance. The whole idea that people prey on people that are less able to fight back,

yesterday I was in Des Moines, not running for President but in Des Moines, and I was at a manufactured housing neighborhood, and a large hedge fund from Salt Lake City has begun to buy up manufactured housing neighborhoods. There are six of them in my State. There are a number of them in Iowa. They are in a half dozen States at least. They come and they buy these. People have paid \$50,000 or \$60,000 or \$70,000 for their manufactured home. They pay \$200 to \$300 a month for the rent on the land, and this hedge fund is raising rents over about a period of a year, a year and a half, up to 70 percent, and people have nowhere to turn. And it is like these companies out there are just looking: Where can we come in, extract the most money at the lowest cost against people that are the most—have the least ability to fight back without political connections? And it is just happening across our economy.

Thank you.

Chairman CRAPO. Thank you, Senator Brown.

Senator SCOTT.

Senator SCOTT. Thank you, Mr. Chairman. I will note that some people go to Des Moines not to run for President, but perhaps Vice President.

[Laughter.]

Senator SCOTT. I apologize. I meant—

Senator BROWN. Mr. Chairman, Senator Scott is a really smart guy, but that was not the smartest thing he ever said.

[Laughter.]

Senator BROWN. Go on.

Senator SCOTT. Senator Brown, I realize you do not actually run for Vice President by the number of votes you get, but I think there is a process by which people say they are qualified to do things—like ask Ms. Dixon a question.

So one of your comments that you made sounded—I spent about 25 years in the insurance industry, so one of the comments you made sounded a little bit like redlining, and I would love for you to unpack that a little bit, but just to make sure I heard you. So in unregulated ways, credit scores that consumers themselves do not know about, that consumers have not seen, heard, or contributed to, are being used in ways that will impact their financial well-being to include perhaps even the likelihood of jobs that they may or may not be qualified for, that to me sounds fairly nefarious, but it sounds a whole lot like redlining. Can you unpack—if that is not what you meant, please clarify what you did mean. And if it is what you meant, please drill down a little bit so that we can have a little more clarity to what you are talking about.

Ms. DIXON. Thank you. It is a really complex issue, and in “The Scoring of America” and in my written testimony, I have articulated it more fully with footnotes.

Senator SCOTT. We have that part.

Ms. DIXON. Yes. So thank you for your question, because it is complex and it is difficult to abstract into a few words. Let me try and make a big effort here. All right—

Senator SCOTT. I will give you 3 minutes if you need it.

Ms. DIXON. Let us go for it.

Senator SCOTT. OK.

Ms. DIXON. So there are amazing real-time analytic products. Actually, in our update to “The Scoring,” we have looked at this. So, for example, financial service companies, you can look across the United States and see pretty much real time the marketplace activity of people who are spending and buying and what that looks like in real time. You can drill down to the census block level and see how well a neighborhood is performing. There is, for example, a product that gives you what is called an “up-front score,” what the score of that neighborhood is. And I will send on follow-up a series of screen shots of this to you so you can see it.

Senator SCOTT. Thank you.

Ms. DIXON. But let us say that you are applying for a university position, and your neighborhood has a very poor score. Well, now that can be taken into consideration. We have the college board doing this. They have an adversity score that is doing exactly this. So I find this difficult. The lines are narrow—

Senator SCOTT. Just to interrupt you, Ms. Dixon. I read an article I guess a couple weeks ago, Mr. Chairman, about this new SAT score that would take into consideration challenges. Are you suggesting that that score could—the neighborhood score could have an impact on one’s SAT score and college admittance?

Ms. DIXON. I do not believe it will have an impact on a person’s SAT score. I do believe that it can have a much further and much larger impact—

Senator SCOTT. Ms. Dixon, are you familiar with the new iteration of the SAT score which takes into consideration the family challenges in—

Ms. DIXON. Yes.

Senator SCOTT. OK.

Ms. DIXON. Yes, I am, and that is what I am referring to. So while that score is meant to provide context, here is the problem. One of the factors that it uses is a neighborhood risk score, and that neighborhood risk score is a secret score. Consumers do not get to see it. Currently, the college board adversity score, the students’ score, they are not allowed to see it. It is a secret score.

Now, let us bring this score into transparency. Let us apply some of the principles of the Fair Credit Reporting Act. Let us give people access to the score. Let them know what factors went into the score. Let us make it fair. That is my point.

And right now this does not fall under the Fair Credit Reporting Act by all law. It does not fall into any eligibility circumstance, not yet. But that is what I am saying. We need to have fairness. Technology is going to advance, and it is important that it does. We need to stay competitive in the United States within machine learning and AI. It is very, very crucial for our economic future. But we need fairness and transparency, and we really need the Fair Credit Reporting Act to be guiding best practices and saying, look, technology, yes, uses need to be right. That is the deal.

Senator SCOTT. Thank you.

Mr. Chairman and Ranking Member, I would love for us to do all that we can to compel some of the companies in the industry to participate in a future hearing.

Chairman CRAPO. You have both of our agreement already on that, Senator Scott.

Senator SCOTT. Thank you, sir.

Chairman CRAPO. Thank you.

Senator REED.

Senator REED. Well, thank you, Mr. Chairman. And thank you to the witnesses for their testimony.

In previous hearings, echoing some of the comments of my colleagues, in particular Senator Kennedy, where a lot of the information should be viewed as being owned by the person, not by these data brokers. And we have to create real opportunities to protect your data. We have got some legal statutes in place like the Fair Credit Reporting Act, HIPAA, *et cetera*, where it is clear by statute. And then we have got some information that is very public. It is published, and it is linked notices in newspapers, *et cetera*. And then there is all the information that is just accumulated by being on a computer.

It comes back down to, I think, three principles. This is my view. One is that consumers, people, should have the ability to opt out of any information collection system. Then, second, this information should be at some point expunged, 6 months, a year, *et cetera*. And then if it is violated by anybody, a data broker or a collector or anyone else, then they should have the right to go to court and say, "You have ruined me."

So let us start with both your comments on how do we get sort of an effective opt-out. You know, my sense is that someone using or going to a website, it is hard to figure out where the opt-out is. Sometimes they do not even offer that. Should we in the U.S. Congress say you have to have a very prominent opt-out, do not collect my data? Let us start with Dr. Cackley and then Ms. Dixon.

Ms. CACKLEY. So an opt-out possibility is certainly something that is available and is used in certain circumstances. I think there are more circumstances where it could be helpful. I do not know that that as a solution alone would do the trick in terms of if you think about all of the times when you go online and you are supposed to read the disclosures and click on things.

Senator REED. No one reads the disclosures.

Ms. CACKLEY. Yeah, exactly, and so it may be that no one will read the opt-out either.

Senator REED. That is why the opt-out cannot be hidden in the disclosures. It has to pop right up here saying, "Click yes or no."

Ms. CACKLEY. Absolutely. Right. I think if someone knows that they do not want their data to be collected and they can opt out right away, that is a way to do it. In other circumstances, people may not understand what the opt-out is, really—

Senator REED. I think if you start with the major platforms, the Googles, and *et cetera*, if they cannot collect the data, then that data is not going to get down the road to the brokers because they do not have it.

Ms. CACKLEY. Absolutely.

Senator REED. And that is the first place, I think, to begin.

Ms. Dixon?

Ms. DIXON. Thank you. I was honored to serve at the OECD as part of their AI expert group. I just finished helping them write the global guidelines on AI, and something that I learned in that process even more so than I already had is that our data world, our

data ecosystems have become so profoundly complex that I am not at all persuaded anymore that opt-out is possible, because if you recall, you know, the Russian nesting dolls where you have the big doll and then all the—you open the doll and there is another doll. And then you open it up again and there is another doll. This is what data is like.

So let us say we do opt out of, you know, a platform. Well, what about all of the financial transactions. The financial transactions and our retail purchase histories are actually the basis of a lot of data broker analysis. And then it gets worse. As you get into the dolls, here is one that really is very, very challenging, and that is this. Data brokers right now, if they did not collect another piece of data on us—here is something really to think about—they could simply create data about us because that is the state of the technology. And I do not know how to create an opt-out that is that far removed from us.

However, that being the case, I do believe there are things we can do, especially if we focus on restricting negative uses that harm consumers and really look at the endpoints of that process, and also at the beginning and say, hey, what are the standards you are using? What can we do to make good standards? And at the end, what are the standards for use? How can we control these two points?

But I think there is a role for opt-out, for example, especially for human subject research, where there must be meaningful consent. As a tool, I think it has lost a lot of its power.

Senator REED. You have studied this longer than I, but I think it is a place to begin, and it is not a perfect solution, but, you know, you cannot make the perfect the enemy of the good. If it gives people a little more protection, I think it should be pursued.

The other aspects of this, too, as you pointed out, with this synthetic—they create the synthetic data. Sort of purging it periodically might also help this. Again, I think you have put your finger on this dilemma now. The complexity, the ability to gather indirectly, not directly, data is profound. But if we do not take some simple steps, it gets worse. It does not get better.

Thank you.

Chairman CRAPO. Thank you.

Senator Schatz.

Senator SCHATZ. Thank you, Mr. Chairman. Thank you to the testifiers.

Ms. Dixon, you know, we are talking about some reforms to the Fair Credit Reporting Act, and what worries me a bit is that, as important as I think it is to bring data brokers back into the fold in terms of how the statutes governs their behavior, the Fair Credit Reporting Act does not actually work as it relates to the credit bureaus. The credit bureaus put the onus on the consumer. The consumer has to pay to correct or monitor his or her own data, and so that statute is broken. And so to the extent that we are going to put all of these shadow data brokers under FCRA, I think we have to be clear-eyed about how imperfect that system is for millions and millions of Americans. I would like you to comment on that.

Ms. DIXON. Well, I agree with you. That is why I said that even our best American privacy law is not as important as it used to be. It does have cracks and fissures. However, it does something very important. It makes it so that things are not secret. You and I, we can look at our credit score. This is huge. This is a huge improvement from pre-2000 when it was illegal to do so. We can see our bureau report and correct it. We cannot see our other scores, and this is problematic.

Senator SCHATZ. Fair enough. Let me ask you a sort of technical question. What is the relationship between data brokers and credit bureaus? In other words, are some of these credit bureaus getting into the data broker business? Have some of them acquired data brokers? What is their relationship?

Ms. DIXON. Yes, so, for example, Equifax and Experian, a lot of times what they will do is they will have part of their business as a formal regulated credit reporting business, and then other aspects of their business are unregulated—

Senator SCHATZ. Which is what they would characterize as the “marketing side.”

Ms. DIXON. Yes, I am aware that they call it “marketing.” However, I call it the “consumer scoring side.” But, yes, your point is absolutely correct. And, additionally, you mentioned that there is, you know, also first party. One of the things that has been happening is there is a lot of data privacy concerns, and there is a real move now for a lot of different types of businesses to purchase data brokers and bring them in so that they are dealing with first-party data. So now we have a fracture in the data broker business model where you cannot just say, “Well, here are the data brokers. Let us regulate them.” That is not possible anymore. Maybe 25, 30 years ago, but not now. I think we really have to look at practices and say, hey, are you using the data for these purposes, especially in regards to eligibility.

Senator SCHATZ. But the challenge, to follow up on what Senator Scott talked about in terms of digital redlining, is that to the extent that they are using data sets that are essentially in combination a proxy for race, and to the extent that those algorithms are not transparent, it is incredibly difficult to imagine that even if we put them under FCRA and even if the FTC were authorized to go after—or CFPB were authorized to go after them, just to make the case would be incredibly difficult. Am I correct there?

Ms. DIXON. I believe you are correct, and that is why we proposed a standard bill that really looks at creating new standards to start to build a mesh network to fill in these gaps. Because you are correct, there are important gaps here.

Senator SCHATZ. And under FCRA and in the sort of old days, you used to have shadow shoppers to try to figure out whether there was discrimination in terms of impact as opposed to in terms of intent. And yet it seems to me that there could be a way where we could subject all of these data brokers to a regime where they had to—they did not have to provide the code for their algorithm, but they had to provide a regulator with the ability to utilize the algorithm and see if the—and run a bunch of reps and figure out if, statistically speaking, it was, in fact, a proxy for race or if there was a disparate impact on protected classes.

Ms. DIXON. I think that is right. And, you know, it is not that algorithms are bad. It is not that scoring is bad. It is how it is used—

Senator SCHATZ. And some of this could actually alleviate the problem of the credit bureaus in terms of the 3 or 4 million people who have bad credit scores that are incorrect. And so if you can come up with an alternative that is nondiscriminatory, it provides a real opportunity.

I will just offer one last thought, and I would like both of your comments for the record. We are working on legislation and I am working on legislation to establish a duty of care, because I think the problem is in a sectoral approach some of these companies are—I do not know if they are a FinTech company or a tech company or under the HIPAA regime, and they sort of evade the various regulations because it is not clear where they belong. And in any case, once the data has been collected, either voluntarily or not, either through the internet of Things or at one point you clicked “I agree” because you signed up for a social platform, the question is: What is the obligation of the company who is in possession of your data? And the duty of care is the most simple way to say cross-sectorally you may not intentionally harm any person whose data you are in possession of. And that is why the duty of care is such a clean way to address all of this because, otherwise, we are going to be always a decade behind whatever these new-fangled companies are attempting to do to us. But if I could take that for the record, please.

Ms. DIXON. Yes, I think that that is a potentially very good approach. I think Vermont did something like this at the State level where they said you cannot purchase data with the intent to defraud or discriminate. So I do think that ensuring that fairness is percolating throughout the system is a really good remedy.

Senator SCHATZ. Thank you.

Chairman CRAPO. Senator Cortez Masto—oh, did you want to have Dr. Cackley—

Senator SCHATZ. No. I was going to take those for the record.

Chairman CRAPO. So he will let you respond in writing, is what he is saying.

Senator SCHATZ. Thank you.

Chairman CRAPO. Senator Cortez Masto.

Senator CORTEZ MASTO. Thank you. I appreciate that. But I would like to hear what Dr. Cackley had to say as well.

Chairman CRAPO. All right.

Ms. CACKLEY. So in terms of, I think, a duty of care, a basic part of a comprehensive privacy law, that would be a good element to include. What we have reported is that given the gaps that the sectoral approach allows in terms of privacy, we have recommended that Congress really consider a more comprehensive approach and include within it several different elements, and a duty-of-care element should certainly be part of that consideration.

Senator CORTEZ MASTO. Yeah, I like that idea, too. I think it is very innovative. Along with that, transparency would be key, right? The consumer knows that whatever regulated credit score or unregulated credit score, whatever is being used that is based on an

algorithm that is identifying their factors, they should have access to that, correct?

Ms. CACKLEY. Access, control, ability to correct, all of those are important elements, yes.

Senator CORTEZ MASTO. OK. So, Ms. Dixon, I understand in 2015 Allstate insurance began selling consumer driving data, and Allstate Chairman and CEO Tom Wilson said that the property casualty insurance company hopes to profit from the sale of telematics data and then pass on savings to consumers by lowering premiums.

Is Allstate unusual in its plans to capture this information about people's driving data to earn additional profit? And, I am just curious, how many insurers have adopted telematics? And what has been the impact, if you know?

Ms. DIXON. So my understanding is that they are no longer the only insurance company doing this. There are now several insurance companies. And there are also health insurance companies who are saying, hey, give us access to a variety of your data and we will give you commensurate lower rates when applicable.

So I think that these are rather uncomfortable things, and, to put it mildly, I would really like to see guardrails on how these are used. I do not think we can stop what is happening in prediction. Prediction is getting cheaper, and it is getting more accurate. So we cannot stop it. However, I think we can take a multifactorial approach to the problems, the real problems that these situations impose. Do we want consumers giving away their data in order to, you know, have a better premium? And I think that you should be able to have protections without giving away your data. We need good rights here.

Senator CORTEZ MASTO. Right.

Ms. DIXON. And to do that, we are going to have to have good rules of the road that encompass new technology, but keep the values, let us make a decision, and not be financially penalized for it. And should an insurance company be able to sell this data? That is a question we need to have as a matter of public discussion. It should not just be decided just by industry. It needs to be a multi-stakeholder conversation about that.

Senator CORTEZ MASTO. And this type of data is what goes into what you have identified as the neighborhood risk scores that—

Ms. DIXON. That is part of it.

Senator CORTEZ MASTO.—companies could use, correct?

Ms. DIXON. Oh, there are so many scores, but, yes—

Senator CORTEZ MASTO. But that could be part of it, there is so much data.

Ms. DIXON. Absolutely.

Senator CORTEZ MASTO. And the other concern I am understanding is that because of the new technology and algorithms, the concern is that this information with respect to unregulated credit scores could end up providing higher accuracy levels than the regulated credit scores, such that the banks or other financial institutions would start using those unregulated credit scores more so than the regulated. Is that right?

Ms. DIXON. Well, I think that banks in particular are very, very careful about these kinds of uses. Of the people that we have interviewed, they have been very, very careful. Actually, some of the

people I worry about the most are the people who are not in banks and who want to pull a credit score product to do marketing. And instead of actually going through the regulation and making a firm offer of credit or insurance, they will just kind of skirt around the edges and pull the, you know, unregulated credit score and then make these offers. Someone discussed today especially if it is a predatory offer, this is where things get very problematic. If you have a consumer who is identified in the credit score 400 to 500 level and someone does not want to make a firm offer of credit or insurance but they want that number and they want to use that number to market a product maybe for bill consolidation or for payday loans, then I think we all need to be very interested in protections for that.

Senator CORTEZ MASTO. Thank you. And I notice my time is also up. I will also just submit this for the record, facial recognition and data that comes from that. It is topical right now, and the question would be: Should that information be shared with third parties like data brokers to be utilized? I am curious about your thoughts on companies in general—which I think it was just in the paper today, airlines were looking at using this type of facial recognition data. So I will submit that for the record.

Thank you so much for this conversation today. I appreciate it, Mr. Chairman.

Chairman CRAPO. Thank you.

Senator WARNER.

Senator WARNER. Thank you, Mr. Chairman. Let me, first of all, just associate myself with both your comments and the Ranking Member's comments. It is pretty remarkable that you invited the data industry, the data brokers to come, and they did not show up. I think that is a very telling statement.

I know folks have talked about the Fair Credit Reporting Act. I know we have talked about a variety of issues. I have been thinking a lot about this in terms of the social media companies. You know, the data brokers are really just one piece of the overall growing data economy, and we are talking a lot about third-party vendors. Obviously, I have got concerns as well about first-party vendors, the Amazons, the Facebooks, the Googles.

Would you both agree that, candidly, most Americans do not have the slightest idea of what kind of data is being collected about them and what that data is worth?

Ms. CACKLEY. I think it is definitely true that most Americans do not understand the breadth of data that is collected about them. They may be aware in certain instances where they have checked yes or provided something, but they do not know the true extent of it.

Senator WARNER. Ms. Dixon?

Ms. DIXON. Thank you. The complexity of data flows right now is extraordinary, and you are correct, first parties, third parties, everything is blending. And if you look at even just identity, you can have an identity that overlaps in 20 different data ecosystems. And as a result, it has become very difficult for anyone to map the data.

There is this amazing chart that was produced by the advertising industry for itself, actually, and it maps this extraordinarily. It looks like the Tokyo subway lines. I mean, it is incredibly complex.

And I do not know that it is possible to fully map our data anymore.

So if that is the case, how on Earth do we cabin practices so that there is almost like a set of routine uses where here are the acceptable uses for companies, end of discussion, boom; and then outside of this, not acceptable uses. We are going to have to find our way to something like that, and we might have to distinguish it by sector and by perhaps even individual companies. But I would like to see that very fairly adjudicated. I am really interested in seeing people talk with each other to figure this out. We need to have very meaningful discussions to figure out where the data is going and how we can best protect it. But I do not think people know about—

Senator WARNER. One of the things that you touched on briefly, one of the areas I have got some bipartisan legislation that would try to focus on some of the manipulative practices, so-called dark patterns use, where, you know, in layman's terms, you have six sets of arrows clicking on—you know, pointing you toward the "I agree" button and you can never find the "unsubscribe" button, and there are a host of practices that go on in the industry where people give up this information, oftentimes unwittingly, and through extraordinarily sophisticated psychological tools being used by the companies and others to get this information.

I know my time is getting down. I would just like your commentary. I believe consumers ought to have a right to know what data is being collected about them. I believe we need to take it a step further and also have some basic valuation in terms of how much that data is worth. And I am an old telcom guy. For a long time, it used to be really hard to bring competition in the telco market until we instituted, by Government regulation, number portability. I believe that same concept, data portability ought to be brought into the data economy so that if you are not liking how you are being treated—I think about it mostly in the social media context, but there are a variety of areas, in the credit-scoring areas as well, where, you know, if we had that knowledge of what data was being collected, what it is worth, and then if you did not like the way Facebook was treating you or some other enterprise, you were easily able to move all of your data in one swipe to a new company or a new platform. I think you could bring some additional competitive practices to the area.

In these last couple seconds, data valuation, data knowledge, and data portability, ideas? Comments? Suggestions?

Ms. DIXON. I really like the idea of data interoperability so there is more freedom—

Senator WARNER. With portability, you have got to have interoperability or it does not work.

Ms. DIXON. Yes. But I think that it is going to be something that will end up working out in time, but it should be a good priority.

Ms. CACKLEY. So this is not something that we have looked at specifically, but I think to the degree that you are talking about comprehensive legislation that really covers all of the different platforms and parties, then that kind of interoperability would be—

Senator WARNER. We would like to share with both of you some of the work we have been doing, and I think there could be broad-based bipartisan support.

Thank you, Mr. Chairman.

Chairman CRAPO. Thank you.

Senator Kennedy.

Senator KENNEDY. Thank you, Mr. Chairman.

If I go on the internet and I search and I look at social media and I buy something on Amazon, let us say, who—I mean, my actions, my behavior is recorded. We call that “data.” Who owns it?

Ms. DIXON. I have a white paper I am going to send to you. We spent a lot of time thinking about this issue. So the issue of data ownership is quite difficult to parse, but let me give you my best shot and let us have a discussion.

Senator KENNEDY. Well, I would like to have a discussion, but first I would like to have an answer.

Ms. DIXON. Here is the answer: I view data in our current data ecosystems as a common pool resource. I think a lot of different entities can lay claim to that data. However, no one gets to own it, and—well, in some cases they can.

Senator KENNEDY. You do not think that I own my data?

Ms. DIXON. It depends on where you have used it and where it is. I think there are some—

Senator KENNEDY. How about you?

Ms. CACKLEY. I do not think there is an answer to who owns your data once you have taken an action, especially in some ways interacted with another company.

Senator KENNEDY. Well, let us suppose that Congress passed a law that said the consumer owns his data and he or she can knowingly license it. What would be wrong with that?

Ms. CACKLEY. I do not think there would be anything wrong with it. I think it would have impact on who could then collect your data or whether data could be collected.

Senator KENNEDY. No, I could license my data knowingly.

Ms. CACKLEY. Right.

Senator KENNEDY. Now in terms of knowingly licensing my data that I own, what sort of disclosures should a social media company, for example, make to me in terms of how it is going to use my data? Right now they make disclosures, but they do not inform the consumer. I have said before some of those things are 7, 8, 9, 10 pages, written by lawyers, you could hide a dead body in them, and nobody would find the body. I mean, nobody reads them. That is not knowing consent. What would a social media company have to tell me in order for me to know what they are doing?

Ms. DIXON. May I offer an example from the medical field? So under HIPAA, there are very meaningful mechanisms prior to a consumer agreeing to release their information outside of the protection of HIPAA. However, one of the concerns that has come up with this is that it has become very, very easy for consumers, patients, to “donate” their data. And what has happened is that people have donated their data and taken it out of the protections of HIPAA without meaningful consent.

Senator KENNEDY. Ms. Dixon, I am not trying to be rude. I am trying to get answers. Here is my question: If I own my data and

I license it, I need to understand what licensing it means. What needs to be disclosed to me?

Ms. DIXON. My understanding, looking at other fields—because this is not something I have studied at length. My understanding is that is a serious agreement, and it would require massive disclosures. I think you could almost put a graveyard in that disclosure, you know, compared to—

Senator KENNEDY. And you do not think it is possible to write a disclosure that the consumer would understand? Is that what you are saying?

Ms. DIXON. In this area, I would have to really look at that. Again, this is not an area of research for me, but I—

Senator KENNEDY. What do you think, Doc?

Ms. CACKLEY. I think it would be very complicated. It is not an area that we have looked at either, but if Congress were to pass a law that allowed consumers to license their own data, that would require a large amount of regulations to go along—

Senator KENNEDY. So you both think that we should just allow companies to do what they want with our data, that this problem is impossible to solve?

Ms. CACKLEY. No, no. I do not think I meant that at all. I just meant that it would have to be worked through. It is not an easy fix.

Senator KENNEDY. No, I do not think there are any easy fixes around here.

Ms. DIXON. And I do not mean that either. I believe that we should have rules of the road, and we should have agreed-upon rules on what—

Senator KENNEDY. I agree with that, too, and everybody—we have had a lot of interesting discussions about this, but no offense to you, two, but the experts never offer a solution. To me the solution is the consumer owns his data. You can license it. Licensing has to be knowing and intentional. You can move your data. Portability should be an option. I can change my mind about licensing it. And companies will adapt to that. They will have no choice.

Thank you, Mr. Chairman.

Chairman CRAPO. Senator Menendez.

Senator MENENDEZ. Thank you, Mr. Chairman.

I have the same concerns as Senator Kennedy because we seem to be living in an age of data breaches. Just last week, we learned of a breach concerning a medical billing company, American Medical Collection Agency, that may have exposed the personal, financial, and even medical data of 20 million patients who were customers of Quest Diagnostics and LabCorp.

So let me ask you, Ms. Dixon, people are rightly concerned that some of their personal data is now exposed and could be used against them. Can data brokers legally compile, aggregate, or sell data that has been acquired through an illegal hack?

Ms. DIXON. I am not an attorney, so I think that is a question an attorney could better answer you. But my first best guess is I do not think you can use improperly information that has been disclosed in an unauthorized manner for your own business purposes. That seems like that would be really out of bounds.

Senator MENENDEZ. Dr. Cackley, do you have any idea?

Ms. CACKLEY. I do not know the answer, but I can certainly find out.

Senator MENENDEZ. Yeah, well, I would appreciate that.

Should people be concerned that data not otherwise covered by HIPAA is ending up in the hands of data brokers even in the absence of a hack? Are billing companies like American Medical Collection Agency selling non-HIPAA data to brokers?

Ms. DIXON. This is an ongoing area of grave concern for us. There are actually scores of health data. There is a frailty score that can predict very closely how sick you are and when you might possibly die. I think that there are all sorts of scores and products related to—

Senator MENENDEZ. I am not sure I want to check on that data myself.

Ms. DIXON. Yeah. Me either. But—

Senator MENENDEZ. But that is pretty frightening, isn't it?

Ms. DIXON. It is. You know, health data that is not covered under HIPAA has become an increasing area, so—

Senator MENENDEZ. Well, let me ask you this: When hackers gain access to non-HIPAA data like in the Quest data breach, can data brokers apply machine learning to these data points to infer or reconstruct sensitive HIPAA-protected medical data?

Ms. DIXON. I actually do not think that they need to acquire unauthorized data to do that. They can just look at our purchase histories and get an awful lot of data about us. But in terms of what is happening with this entire area, the data breaches of medical data actually can lead to forms of identity theft and medical identity theft that are very, very difficult to cure and can have extremely meaningful consequences in people's lives.

Senator MENENDEZ. Well, let me ask you, then, HIPAA is nearly 25 years old, and the 2009 HITECH Act provided updates which were concerning health information technology. But I am still concerned that we are playing catch-up when it comes to protecting patients. You know, of all the information that should be private and privileged to you, your health standing should be extraordinary—there are all types of consequences in that, in employment and discrimination, in a whole host of things. Are there gaps in HIPAA and other data security laws that need to be addressed to better protect people today in this 21st century threat? What coordination is missing between existing legal protections?

Ms. DIXON. I do think there are gaps, and the biggest gaps that exist right now are the gaps that exist between the sectoral protections, and I do not think the answer is to just rip out the sector protections that exist, such as the Fair Credit Reporting Act or HIPAA or Sarbanes-Oxley, *et cetera*, but to find a way to fill those gaps in. For example, victims of medical identity theft can use their Fair Credit Reporting Act rights to get their financial information corrected. But under HIPAA it is not possible for them because it does not exist in the statute. It is not possible for them to get a deletion similar to the FCRA in their health file, so they can actually carry around inaccurate information which can really have an impact on their treatment and insurance costs. And there is not a solution yet. So this is the kind of gap we need to address.

Senator MENENDEZ. All right. Last, there was one breach that compromised the personal information of 20 million patients. That is pretty troubling. One data broker has data on 300 million consumers. We are still reeling from the Equifax breach which affected 145.5 million consumers. If the information of 300 million consumers were to be compromised, we might start calling private information public information because at the end of the day that is the result of it.

What are the ramifications for a consumer if a data broker is breached? And should we hold them to a higher standard of security, especially because their volume is so consequential?

Ms. DIXON. Data broker breaches are very significant. So my assessment of this is that the various State data breach laws are doing a pretty good job, especially in some cases where the data breach law is quite strong, in forcing disclosures and notices. But I think we need to do more to ensure that all of the information held that is sensitive and health related, *et cetera*, is duly notified to the consumer.

The problem with the data brokers is what they will say is, oh, wait, wait, we do not have a direct relationship with the consumers; we cannot notify them. And I think that is a gap that needs to be resolved. Now, the State of Vermont has resolved that gap.

Senator MENENDEZ. Well, they could reach back to the entity that provided them the data in the first place, and they could notify, could they not?

Ms. DIXON. I believe that that could happen. And it has happened in some—

Senator MENENDEZ. I just think they should be held to a higher standard of security because the consequences of incredible numbers of Americans that are subject to having their privacy breached and their health care breached is just beyond acceptance.

Thank you, Mr. Chairman.

Chairman CRAPO. Thank you.

Senator ROUNDS.

Senator ROUNDS. Thank you, Mr. Chairman, and thank you for holding this hearing today.

I would like to see, in listening to this, if I have picked up the grasp of some of the challenges we have here. It would appear to me that we are talking about, first of all, the question of the security of the data that is actually being collected. Second of all, it appears that we are questioning whether or not there is an appropriate way for individual consumers or individuals to actually find out and to have access to what these organizations, these nonregulated organizations actually have. And, finally, this appears that it may very well be a work-around with regard to the information that is being collected and then disseminated from what a regulated entity would have.

In a nutshell, are those the three areas? And would there be other areas that you would also identify? I would ask each of you for your thoughts.

Ms. CACKLEY. Those are certainly three of the main points that have come up today. I think the other piece that we have not touched on maybe as much is outside of the data brokers them-

selves. There are other technologies with privacy issues, you know, mobile devices, facial recognition technology—we did mention that—with financial technology. All of these are areas of concern that fall outside potentially the protections of FCRA in particular.

Senator ROUNDS. The use of machine learning and artificial intelligence in this process. OK.

Ms. Dixon?

Ms. DIXON. So my focus has really never been on the technologies as an endpoint. My focus has always been on, OK, so we have technological processes that are going to continue through time, but what does that actually mean in practice. I have always looked at the practice. So your assessment of where the sticking points are is accurate. The thing I would add is this: I think it is going to become, as we move forward and prediction gets cheaper, I think prediction is going to be coming to a mobile phone near us, like ours. And I think we have to be very cautious about looking at categories of technologies and labeling them as bad. Similarly, in industry, I think we have to be very careful and say, OK, what are the practices that we want to go after here and want to address because they are harming consumers. And if we can do that in a truly multifactorial way, I think that will be helpful. Wherever these practices exist, wherever they are, we need to be addressing them because they are meaningful and have impacts.

Senator ROUNDS. There is a difference between the way that we have looked at data and data collection and privacy in the United States versus the way that it has been done in some other parts of the world. Here we follow and we use Gramm-Leach-Bliley within the United States, but in Europe they take a different approach—the GDPR, which seeks to really achieve a different and more comprehensive approach, but would be rather challenging.

Can you share with me the thought process or your analysis of the differences or the advantages, one versus the other, between the way that we handle it today in the United States versus what they are doing in Europe with the GDPR in its current form?

Ms. CACKLEY. So we have not looked at GDPR directly yet, but I can say that there are definitely some elements of GDPR that embody the Fair Information Practices Principles, which are the basis of some of our privacy regulation already. There are other pieces of GDPR that are not in the U.S. privacy framework, and one of the main ones, I would say, is the right to be forgotten. The right to be forgotten is a part of GDPR that really is not encompassed in the U.S. privacy framework.

Senator ROUNDS. Ms. Dixon?

Ms. DIXON. The GDPR, as you know, it was built on the EU 95/46, so it has a lot of bureaucratic history behind it. If you look at what they were trying to do and all the derogations and what-not, it is a really complex and thought structure.

I think that it does provide for baseline privacy protections, but they do not have the sectoral system and they do not have government privacy. So I think there is one thing I will say. In our country, the Privacy Act is very effective in regulating certain aspects of government information collection. They do not have anything like that.

Senator ROUNDS. Thank you. I see my time has expired, Mr. Chairman. Thank you to both of you for your answers today. And, Mr. Chairman, once again thank you for the opportunity here today with this hearing on this very important topic.

Chairman CRAPO. Thank you, Senator Rounds.

Senator SINEMA.

Senator SINEMA. Thank you, Mr. Chairman. And thank you to our witnesses for being here today.

At the Committee's last hearing on privacy, I spoke about the importance of privacy to Arizonans. We are practical people who want the modern conveniences that technology brings, but we value our privacy. So I am committed to making sure that Arizonans know how our data is being used so that we can make informed decisions.

Arizonans also do not like assumptions being made about us or how we choose to live our lives, particularly if some of those assumptions are wrong, which is why current privacy and consumer scoring laws concern both me and many Arizonans.

In 2013, the FTC completed and published a 10-year congressionally mandated study on the accuracy of credit reports. The FTC found that one in five consumers had an error on at least one of their three credit reports. So, Ms. Dixon, first, thank you for being here. I want to talk quickly about credit scores as a starting point and what happens if you or I were one of those consumers.

How drastically could an error in a credit report negatively affect an Arizonan's credit score?

Ms. DIXON. Yes, that effect would be profound. So, for example, for victims of identity theft, if someone has run up your credit and it is not actually your error, you could be seen as not making your payments, *et cetera*, and you can literally move from a 780 score to a 620 in very short order. It only takes about a month. And then what you have is a situation where, if you are about to buy a home—and these are from the calls we get. This is not just a hypothesis here. The home you are about to buy, all of a sudden you cannot qualify for a mortgage because of identity theft.

So, yes, any error from any source that is in your credit report, it is a piece of serious business.

Senator SINEMA. So, Ms. Dixon, you said this could potentially prevent an Arizonan from buying a home. Would it also get in the way of financing an education or starting a small business or expanding one's business?

Ms. DIXON. Absolutely.

Senator SINEMA. Wow, that is really troubling.

Under the Fair Credit Reporting Act, if an Arizonan thinks his or her credit report or score is inaccurate, they can appeal it with the bureau. Is that correct?

Ms. DIXON. That is correct.

Senator SINEMA. And if so, how?

Ms. DIXON. Yes, there is a very specific procedure outlined in law where the bureaus must respond, and there is a series of steps that they can take, and both the Federal Trade Commission and the CFPB have numerous help- and hot-lines to help everyone through, and the State AGs also do as well. But there are very well documented recourses for consumers in this situation.

Senator SINEMA. Well, that is good. So we have established it is important to have an accurate credit score and there is a process to appeal it and fix it. But, increasingly, businesses are using so-called consumer scores that rank, rate, and segment consumers based on public-private and government data that is packaged and sold by data brokers and others. So sometimes this public data is inaccurate. It is often outdated or it could be incomplete.

So are all consumer scores made available to consumers just like credit scores are?

Ms. DIXON. Actually, almost none of them are. In fact, I have had almost no success. Despite trying to get consumer scores and asking companies for my consumer score, it is almost impossible to get them.

Senator SINEMA. But then how would an Arizonan know if his or her consumer score was inaccurate if they cannot get access to it?

Ms. DIXON. That is the same question I have. They would not know.

Senator SINEMA. Wow. So let us say that an Arizonan were able to find out that his or her consumer score is inaccurate. Are all consumer scores covered under the FCRA so that there is a similar appeals process to resolve inaccuracies?

Ms. DIXON. No consumer scores that are unregulated are currently covered under the FCRA. Unless it is a formal credit score as articulated by the FCRA and used in an eligibility circumstance, it is not covered.

Senator SINEMA. Well, that is very concerning, but thank you for sharing that information with us.

Mr. Chairman and Ranking Member Brown, it is clear that we have a lot of work to do here. We have got to update our privacy laws to reflect new trends that are occurring in both business and technology to make sure that Americans have the right to correct their record, whether it is their credit score or their consumer score, on who they are, how they have lived their lives, and what mistakes or inaccuracies that might be occurring in their lives.

So I thank you for being here, our witnesses, and I look forward to working with the Committee on this. And, Mr. Chairman, I yield back.

Chairman CRAPO. Thank you.

That concludes the first round, but Senator Brown and I would like to do a second round, and you are welcome to join in with us, Senator, if you would like.

There are so many questions. One of them I want to get back to which has been brought up by several Senators is this notion of the tension between doing a comprehensive bill like the GDPR in Europe or a sectoral approach like we do in the United States. And I think we all can understand there is sort of a push and a pull on both sides of that question.

It seems to me, though, that we do not have a choice, at least at a basic level, to deal with all data collection in the same way. I think one of you mentioned earlier that it is all blending. It used to be that we could clearly distinguish what a credit bureau did and the credit report that a credit bureau prepared. Now we have massive amounts—I think Senator Brown referenced the 4,000 number, but I do not even know what the number is—of entities

that are collecting data. My understanding is that the apps on my iPhone, many of them collect data even when I am not using them to report further to others about whatever it is, data that is not even often related to the app. And it seems to me that all of that data is in one way or another not just blending but being utilized for many, many different purposes, one of which is credit, one of which is retail sales, one of which is college applications, one of which is mortgages. I mean, the list can go on and on and on.

So I guess I would like to have each of you just briefly—because I have got some more questions, but briefly indicate do you believe that at some basic point the United States needs to have a comprehensive set of standards and requirements that would cover some basics, like when data is being collected, who is collecting it, whether there is an opt-in or an opt-out, what rights to manage or even remove one's data exist?

Ms. CACKLEY. Yes, I think that is where we are right now, that the sectoral approach leaves too many gaps. You may not need to completely change to a comprehensive framework, you could merge elements of a comprehensive and sectoral approach in some ways. But a comprehensive framework that gives basic privacy rights and abilities for consumers to know what their data is and how to correct it, how to control it, is definitely something that needs to be addressed.

Chairman CRAPO. Thank you.

Ms. Dixon?

Ms. DIXON. Let me share with you that I have been seeking an answer to the question you just asked for about 27 years, so here is what I have come up with, and it is just—it is my opinion. What if the sectoral system was a feature, not a bug, born from thoughtful deliberation about very focused issue areas with a lot of buy-in? What if we have not been able to pass comprehensive legislation because our system requires more buy-in than other systems? These are just the hypotheses that I am working with.

So if that is the case—and, also I have to tell you, I am quite concerned about the deep disruption to privacy law that would occur if there was massive preemption. But be that the case, what if there was a way to do a surgical strike and to provide guardrails in the areas that need it the most, that would fill in the sectoral gaps? That is what I am very interested in.

So I think that something that had really important principles, fair information practices, principles, and then the adaptation of those principles for the gaps that exist. So I do think that standards have been a neglected part of the privacy conversation. I have no idea why we do not have more standards in privacy.

This mobile phone has loads of standards that attach to it, but for our privacy and for data brokers, where are the standards? Well, let us create some. Let us start there. I am all for starting cautiously and working with best practices, but to give things teeth and to abide by the larger principles.

So a nice amalgamation of all of the above, something that is multifactorial. I do not think we have silver bullets available to us anymore.

Chairman CRAPO. Well, thank you. And just one other quick question, and then I will turn to Senator Brown. We have talked

a lot about the problems we are trying to address here, whether harm is caused by the use of data, whether credit is impacted, whether people are redlined or denied access to products or opportunities. It seems to me that when you approach the issue from that perspective, which is a very legitimate approach, that there is another issue that is—I do not know if I would call it a “harm.” Maybe it is. But there is simply a privacy issue. A lot of Americans, I believe, do not want to have to prove that they were harmed. They do not want people collecting data on them, or they do not want certain data collected. It is sort of the right to be forgotten or the right to opt out of certain segments of data collection.

Is that a legitimate right that we should try to protect?

Ms. DIXON. It is a legitimate option that we need to be able to have. The adversity score, I think that any child who is applying for college should be able to say, hey, wait, I do not want my neighborhood being part of that. Do they have to prove harm? I do not think they should have to. They should be able to say, hey, no, this is not something I want. It is legitimate.

Chairman CRAPO. Dr. Cackley?

Ms. CACKLEY. I think that is right, that it is important for people to be able to make a choice about what data they share and what data they do not.

Chairman CRAPO. Senator Brown.

Senator BROWN. Thank you.

Ms. Dixon, this is the last round of questions, blessedly, for both of you. And please be really brief on these because I have several questions.

Should Federal regulators and supervisors have full access to every company’s predictive models so they can evaluate them for bias and other legal compliance?

Ms. DIXON. I believe they would have to hire about a million people if they did that. I am not sure of the answer to that question, but I have a lot of thoughts on this, and I will send you written follow-up.

Senator BROWN. OK. That would be good, including if there is a list of companies whose models you believe should be available to regulators for review.

Ms. DIXON. I will send that to you.

Senator BROWN. OK.

Senator BROWN. A technology expert at our last hearing stated, “While our online economy depends on collection and permanent storage of highly personal data, we do not have the capacity to keep such large collections of user data safe over time.” Do you agree with that statement?

Ms. DIXON. I think it is very difficult to keep user data safe 100 percent of the time.

Senator BROWN. Should companies be required to expunge certain types of user data after, say, 60 or 90 days?

Ms. DIXON. You know, I think there are very good arguments for that, and there is a continuum for that. And I will respond to that in writing.

Senator BROWN. OK. Thanks.

Senator BROWN. Do companies who currently use personal data for profit see existing penalties as little more than the cost of doing

business? That is often the case in this town, that a few-million-dollars fine on a multi-billion-dollar company is the cost of doing business. How strong do penalties and other enforcement mechanisms need to be in order to hold these companies accountable?

Ms. DIXON. I do not know the answer to that question. However, I do think that having very good enforcement is an important stick, and I think we need carrots and sticks to make things right.

Senator BROWN. Is holding executives personally accountable one way?

Ms. DIXON. I do not know about that.

Senator BROWN. Does that mean no or you just do not know?

Ms. DIXON. It means that I literally do not know the answer to that.

Senator BROWN. A technology expert at our last hearing stated, "While it is possible in principle to throw one's laptop into the sea and renounce all technology, it is no longer possible to opt out of a surveillance society." Do you agree with that statement?

Ms. DIXON. Absolutely. I do not believe that an opt-out village exists.

Senator BROWN. So what would a meaningful consent contract between users and tech companies or users and data brokers or a meaningful opt-out policy look like?

Ms. DIXON. So it needs to be multifactorial and not just rely on consent, because consent is a really difficult vehicle for that. I have a lot of very complete thoughts on that, and I will follow up in writing.

Senator BROWN. OK. You are going to be busy in the next few days.

Ms. DIXON. That is all right. I have a lot on this.

Senator BROWN. And the last question. As you point out in your testimony, household data can serve as a proxy for an individual credit score. Some data that seems innocuous, like Instagram posts, can actually yield predictive data about a user's mental health. How do we know what data is inherently sensitive and what data is innocuous but can become sensitive when it is used to make predictions?

Ms. DIXON. Right. One of the most difficult things that I have had to grapple with as a privacy expert and someone who cares so much about privacy is that it is so difficult to say, here, this is sensitive data, here, this is sensitive data. It is all becoming sensitive depending on how it is analyzed, and that is why privacy protections have had to become much more multifactorial and much more subtle in responding to this new issue.

Senator BROWN. In part, that movement, if you will, from it is initially not sensitive but becomes that is a result of just the power of—the quantity and quality of computing power, correct?

Ms. DIXON. We were in a digital era. We are really moving into the predictive era, and it changes everything.

Senator BROWN. OK. Very good.

Thank you, Mr. Chairman.

Chairman CRAPO. Thank you, and that does conclude the questioning for today's hearing.

For Senators who wish to submit questions for the record, those questions are due to the Committee by Tuesday, June 18th, and we

ask the witnesses to respond to those questions as quickly as you can once you receive them.

Again, we thank you both for not only your time here today but the attention and analysis that you have given to this issue and will give to the issue as we proceed.

With that, this hearing is adjourned.

[Whereupon, at 11:32 a.m., the hearing was adjourned.]

[Prepared statements, responses to written questions, and additional material supplied for the record follow:]

PREPARED STATEMENT OF CHAIRMAN MIKE CRAPO

Providing testimony to the Committee today are experts who have researched and written extensively on big data: Dr. Alicia Cackley, Director of Financial Markets and Community Investment at the Government Accountability Office; and Ms. Pam Dixon, Executive Director of the World Privacy Forum.

As a result of an increasingly digital economy, more personal information is available to companies than ever before.

I have been troubled by government agencies and private companies' collection of personally identifiable information for a long time.

There have been many questions about how individuals' or groups of individuals' information is collected, with whom it is shared or sold, how it is used and how it is secured.

Private companies are collecting, processing, analyzing and sharing considerable data on individuals for all kinds of purposes.

Even more troubling is that the vast majority of Americans do not even know what data is being collected, by whom and for what purpose.

In particular, data brokers and technology companies, including large social media platforms and search engines, play a central role in gathering vast amounts of personal information, and often without interacting with individuals, specifically in the case of data brokers.

In 2013, the GAO issued a report on information resellers, which includes data brokers, and the need for the consumer privacy framework to reflect changes in technology and the marketplace.

The report noted that the current statutory consumer privacy framework fails to address fully new technologies and the growing marketplace for personal information.

The GAO also provided several recommendations to Congress on how to approach the issue to provide consumers with more control over their data.

In 2018—five years later—GAO published a blog summarizing its 2013 report, highlighting the continued relevance of the report's findings.

The Federal Trade Commission also released a report in 2014 that emphasized the big role of data brokers in the economy.

The FTC observed in the report that “data brokers collect and store billions of data elements covering nearly every U.S. consumer,” and that “data brokers collect data from numerous sources, largely without consumers' knowledge.”

In her report “The Scoring of America,” Pam Dixon discusses predictive consumer scoring across the economy, including the big role that data brokers play.

She stresses that today, no protections exist for most consumer scores similar to those that apply to credit scores under the Fair Credit Reporting Act.

Dixon says, “Consumer scores are today where credit scores were in the 1950s. Data brokers, merchants, government entities and others can create or use a consumer score without notice to consumers.”

Dr. Cackley has also issued several reports on consumer privacy and technology, including a report in September 2013 on information resellers, which includes data brokers.

She says in her report that the current consumer privacy framework does not fully address new technologies and the vastly increased marketplace for personal information.

She also discusses potential gaps in current Federal law, including the Fair Credit Reporting Act.

The Banking Committee has been examining the data privacy issue in both the private and public sectors, from regulators to financial companies to other companies who gather vast amount of personal information on individuals or groups of individuals, to see what can be done through legislation, regulation or by instituting best practices.

Enacted in 1970, the Fair Credit Reporting Act is a law in the Banking Committee's jurisdiction which aims to promote the accuracy, fairness and privacy of consumer information contained in the files of consumer reporting agencies.

Given the exponential growth and use of data since that time, and the rise of entities that appear to serve a similar function as the original credit reporting agencies, it is worth examining how the Fair Credit Reporting Act should work in a digital economy.

During today's hearing, I look forward to learning more about the structure and practices of the data broker industry and technology companies, such as large social media platforms; how the data broker industry has evolved with the development of new technologies, and their interaction with technology companies; what information these entities collect, and with whom it is shared and for what purposes; what

gaps exist in Federal privacy law; and what changes to Federal law, including the Fair Credit Reporting Act, should be considered to give individuals real control over their data.

I appreciate each of you joining us today to discuss this important issue.

PREPARED STATEMENT OF SENATOR SHERROD BROWN

I appreciate Chairman Crapo continuing these important, bipartisan efforts to protect Americans' sensitive personal information.

Today, we're looking at a shadowy industry known as "data brokers." Most of you probably haven't heard of these companies. The biggest ones include names like Acxiom, CoreLogic, Spokeo, ZoomInfo, and Oracle. According to some estimates, 4,000 of these companies are collecting and selling our private information, but not one of them was willing to show up and speak in front of the committee today. Not one.

These companies expect to be trusted with the most personal and private information you could imagine about millions of Americans, but they're not even willing to show up and explain how their industry works. I think that tells you all you need to know about how much they want their own faces and names associated with their industry.

As Maciej Ceglowski told us at our last hearing, "the daily activities of most Americans are now tracked and permanently recorded by automated systems at Google or Facebook"

But most of that private activity isn't useful without data that anchors it to the real world. Facebook, Google, and Amazon want to know where you're using your credit cards, whether you buy name-brand appliances, if you're recently divorced, and how big your life insurance policy is. That's the kind of data that big tech gets from data brokers, and they then combine it with your social media activity to feed into their algorithms.

You might have noticed it seems like every product or service you buy comes with a survey or a warranty card that asks for strangely personal information. Why are all these nontech companies so interested in your data?

It's simple—data brokers will pay those companies for any of your personal information they can get their hands on, so they can turn around and sell it to Silicon Valley. It's hard for ordinary consumers to have any power when unbeknownst to them, they're actually the product being bought and sold.

It reminds me of a time when corporations that had no business being in the lending industry decided to start making loans and selling them off to Wall Street. Manufacturers or car companies decided that consumer credit would be a great way to boost their profits. When big banks and big tech companies are willing to pay for something, everyone else will find a way to sell it to them, often with devastating results.

For example, Amazon is undermining retailers and manufacturers across the country through anti-competitive practices, and at the same time, it's scooping up data from the very businesses it's pushing out of the market.

Then there's Facebook—it has almost single-handedly undermined the profitability of newspapers across the country. It's also gobbling up personal information that The New York Times allows data brokers to collect from its readers.

Just like in the financial crisis, a group of shadowy players sits at the center of the market, exercising enormous influence over consumers and the economy while facing little or no rules at all.

Chairman Crapo and I are committed to shining a light on these companies, and to keeping an unregulated data economy from spiraling out of control. I look forward to the witnesses' testimony, and to continuing to work with Chairman Crapo in a bipartisan manner.

United States Government Accountability Office



Testimony
Before the Committee on Banking,
Housing, and Urban Affairs, U.S. Senate

For Release on Delivery
Expected at 10:00 a.m. ET
Tuesday, June 11, 2019

CONSUMER PRIVACY

Changes to Legal Framework Needed to Address Gaps

Statement of Alicia Puente Cackley, Director
Financial Markets and Community Investment

Chairman Crapo, Ranking Member Brown, and Members of the Committee:

I am pleased to be here today to discuss our prior work on privacy, personal information, and information resellers. Information resellers (also known as data brokers) are companies that collect and resell information on individuals. Privacy concerns about resellers stem, in part, from consumers not always knowing what personal information is collected and how it is used. Moreover, growing use of the internet, social media, and mobile applications has intensified privacy concerns because these media make it much easier to gather personal information, track online behavior, and monitor individuals' locations and activities.

My remarks today are primarily based on our September 2013 report on privacy issues related to the consumer data that information resellers collect, use, and sell, and on our 2015 and 2019 High Risk Reports.¹ In 2013, we found that the framework of federal laws relating to the privacy of consumer information had gaps. We recommended that Congress consider strengthening the consumer privacy framework to reflect changes in technology and the marketplace. In our 2015 High Risk Report, we expanded an area of concern—cybersecurity—to include protecting the privacy of personally identifiable information.² We also conducted more recent work in the consumer privacy area on facial recognition technology, financial technology, internet privacy, and consumer data protection.³ In our 2019 High Risk Report, we reiterated our recommendation that Congress consider what additional actions are needed to protect consumer privacy.⁴ My statement will focus on the (1)

¹GAO, *Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace*, GAO-13-663 (Washington, D.C.: Sept. 25, 2013).

²Every 2 years, we report on federal programs and operations that are vulnerable to waste, fraud, abuse, and mismanagement, or that need broad reform—our High Risk List. See GAO, *High-Risk Series: An Update*, GAO-15-290 (Washington, D.C.: Feb. 11, 2015).

³GAO, *Facial Recognition Technology: Commercial Uses, Privacy Issues, and Applicable Federal Law*, GAO-15-621 (July 30, 2015); *Financial Technology: Additional Steps by Regulators Could Better Protect Consumers and Aid Regulatory Oversight*, GAO-18-254 (Mar. 22, 2018); *Internet Privacy: Additional Federal Authority Could Enhance Consumer Protection and Provide Flexibility*, GAO-19-52 (Jan. 15, 2019); and *Consumer Data Protection: Actions Needed to Strengthen Oversight of Consumer Reporting Agencies*, GAO-19-196 (Washington, D.C.: Feb. 21, 2019).

⁴GAO, *High-Risk Series: Substantial Efforts Needed to Achieve Greater Progress on High-Risk Areas*, GAO-19-157SP (Washington, D.C.: Mar. 6, 2019).

existing federal laws and regulations related to the privacy of consumer information held by information resellers and (2) any gaps that may exist in this legal framework.

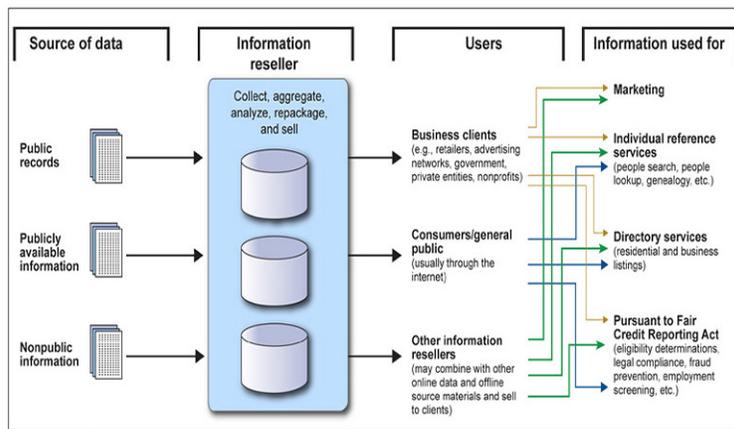
For our September 2013 report ([GAO-13-663](#)), we reviewed and analyzed relevant laws, regulations, and enforcement actions. We interviewed representatives of federal agencies, trade associations, consumer and privacy groups, and resellers to obtain their views on data privacy laws related to resellers. The work for our 2015 report on facial recognition technology ([GAO-15-621](#)), 2018 report on financial technology ([GAO-18-254](#)), and January and February 2019 reports on internet privacy and consumer data protection ([GAO-19-52](#) and [GAO-19-196](#)) included analyzing laws and regulations and interviewing representatives of federal agencies, regulators in other countries, market participants, consumer advocacy groups, and academia. For this statement, we verified that findings of our previous reports about gaps in the statutory framework for consumer information privacy remain relevant. More details about our scope and methodology can be found in our published reports.

We conducted the performance audit on which the majority of this statement is based from August 2012 through September 2013, in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

Resellers maintain large, sophisticated databases with consumer information that can include credit histories, insurance claims, criminal records, employment histories, incomes, ethnicities, purchase histories, and interests. As shown in figure 1, resellers largely obtain their information from public records, publicly available information (such as directories and newspapers), and nonpublic information (such as from retail loyalty cards, warranty registrations, contests, and web browsing).

Figure 1: Typical Flow of Consumer Data through Resellers to Third-Party Users



Source: GAO. | GAO-19-621T

Consumer information can be derived from mobile networks, devices (including smartphones and tablets), operating systems, and applications. Resellers also may obtain personal information from the profile or public information areas of websites, including social media sites, or from information on blogs or discussion forums. Depending on the context, information from these sources may be publicly available or nonpublic.

In 1973, a U.S. government advisory committee first proposed the Fair Information Practice Principles for protecting the privacy and security of personal information. While these principles are not legal requirements, they provide a framework for balancing privacy with other interests. In 2013, the Organisation for Economic Co-operation and Development (OECD) developed a revised version of the principles (see table 1).⁵

⁵Organisation for Economic Co-operation and Development, *Guidelines on the Protection of Privacy and Transborder Flow of Personal Data* (Paris, France: Sept. 23, 1980). OECD's 30 member countries include the United States. OECD has been considering whether to revise or update its privacy guidelines to account for changes in the role of personal data in the economy and society.

Principle	Description
Collection limitation	The collection of personal information should be limited, obtained by lawful and fair means, and, where appropriate, with the knowledge or consent of the individual.
Data quality	Personal information should be relevant to the purpose for which it is collected, and should be accurate, complete, and current as needed for that purpose.
Purpose specification	The purposes for the collection of personal information should be disclosed before collection and upon any change to those purposes, and the use of the information should be limited to those purposes and compatible purposes.
Use limitation	Personal information should not be disclosed or otherwise used for purposes other than a specified purpose without consent of the individual or legal authority.
Security safeguards	Personal information should be protected with reasonable security safeguards against risks such as loss or unauthorized access, destruction, use, modification, or disclosure.
Openness	The public should be informed about privacy policies and practices, and individuals should have ready means of learning about the use of personal information.
Individual participation	Individuals should have the following rights: to know about the collection of personal information, to access that information, to request correction, and to challenge the denial of those rights.
Accountability	Individuals controlling the collection or use of personal information should be accountable for taking steps to ensure the implementation of these principles.

Source: Organisation for Economic Co-operation and Development. [GAO-19-621T]

The Fair Information Practice Principles served as the basis for the Privacy Act of 1974—which governs the collection, maintenance, use, and dissemination of personal information by federal agencies.⁶ The principles also were the basis for many Federal Trade Commission (FTC) and Department of Commerce privacy recommendations and for a framework for consumer data privacy the White House issued in 2012.⁷

⁶See Pub. L. No. 93-579, 88 Stat. 1896 (1974) (codified as amended at 5 U.S.C. § 552a). The act generally prohibits (with a number of exceptions) the disclosure by federal entities of records about an individual without the individual's written consent and provides U.S. persons with a means to seek access to and amend their records.

⁷The framework included a consumer privacy bill of rights and encouraged Congress to provide FTC with enforcement authorities for the bill of rights. The White House, *Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy* (Washington, D.C.: Feb. 23, 2012).

Several Laws Apply in Specific Circumstances to Consumer Data That Resellers Hold

As we reported in 2013 and as continues to be the case, no overarching federal privacy law governs the collection, use, and sale of personal information among private-sector companies, including information resellers. There are also no federal laws designed specifically to address all the products sold and information maintained by information resellers. Federal laws addressing privacy issues in the private sector are generally narrowly tailored to specific purposes, situations, types of information, or sectors or entities—such as data related to financial transactions, personal health, and eligibility for credit. These laws include provisions that limit the disclosure of certain types of information to a third party without an individual's consent, or prohibit certain types of data collection. The primary laws include the following:

Fair Credit Reporting Act (FCRA).⁸ FCRA protects the security and confidentiality of personal information collected or used to help make decisions about individuals' eligibility for credit, insurance, or employment.⁹ It applies to consumer reporting agencies that provide consumer reports.¹⁰ Accordingly, FCRA applies to the three nationwide consumer reporting agencies (commonly called credit bureaus) and to any other information resellers that resell consumer reports for use by others. FCRA limits resellers' use and distribution of personal data—for example, by allowing consumers to opt out of allowing consumer reporting agencies to share their personal information with third parties for prescreened marketing offers.

Gramm-Leach-Bliley Act (GLBA).¹¹ GLBA protects nonpublic personal information that individuals provide to financial institutions or that such institutions maintain.¹² GLBA sharing and disclosure restrictions apply to financial institutions or entities that receive nonpublic personal information

⁸Pub. L. No. 91-508, Tit. VI, 84 Stat. 1114, 1128 (1970) (codified as amended at 15 U.S.C. §§ 1681-1681x).

⁹See 15 U.S.C. §§ 1681-1681x.

¹⁰For the definition of "consumer reporting agency," see 15 U.S.C. § 1681a(f). For the definition of "consumer report," see 15 U.S.C. § 1681a(d).

¹¹Pub. L. No. 106-102, 113 Stat. 1338 (1999) (codified as amended in scattered sections of 12 and 15 U.S.C.).

¹²See 15 U.S.C. §§ 6801-6802. Subtitle A of Title V of the act contains the privacy provisions relating to the disclosure of nonpublic personal information. 15 U.S.C. §§ 6801-6809.

from such institutions.¹³ For example, a third party that receives nonpublic personal information from a financial institution to process consumers' account transactions may not use the information or resell it for marketing purposes.

Health Insurance Portability and Accountability Act of 1996 (HIPAA).¹⁴ HIPAA establishes a set of national standards to protect certain health information. The HIPAA privacy rule governs the use and disclosure of an individual's health information for purposes including marketing.¹⁵ With some exceptions, the rule requires an individual's written authorization before a covered entity—a health care provider that transmits health information electronically in connection with covered transactions, health care clearinghouse, or health plan—may use or disclose the information for marketing.¹⁶ The rule does not directly restrict the use, disclosure, or resale of protected health information by resellers or others not considered covered entities under the rule.

Children's Online Privacy Protection Act of 1998 (COPPA).¹⁷ COPPA and its implementing regulations apply to the collection of information—such as name, email, or location—that would allow someone to identify or contact a child under 13.¹⁸ Covered website and online service operators must obtain verifiable parental consent before collecting such information. COPPA may not directly affect information resellers, but the covered entities are potential sources of information for resellers.

¹³15 U.S.C. § 6802. A "financial institution" is any institution the business of which is engaging in financial activities as described in section 4(k) of the Bank Holding Company Act (12 U.S.C. § 1843(k)). 15 U.S.C. § 6809(3)(a).

¹⁴Pub. L. No. 104-191, 110 Stat. 1936 (1996) (codified as amended in scattered sections of 18, 26, 29, and 42 U.S.C.).

¹⁵45 C.F.R. Parts 160, 164.

¹⁶For the definition of "marketing," including exceptions, see 45 C.F.R. § 164.501.

¹⁷Pub. L. No. 105-277, Div. C, Tit. XIII, 112 Stat. 2681-728 (1998) (codified at 15 U.S.C. §§ 6501-6506).

¹⁸FTC issued regulations implementing COPPA at 16 C.F.R. Part 312.

Electronic Communications Privacy Act of 1986 (ECPA).¹⁹ ECPA prohibits the interception and disclosure of electronic communications by third parties unless an exception applies (such as one party to the communication consenting to disclosure). For example, the act would prevent an internet service provider from selling the content of its customers' emails to a reseller for marketing purposes, unless the customers had consented to disclosure. However, ECPA provides more limited protection for information considered to be "non-content," such as a customer's name and address.

Federal Trade Commission Act (FTC Act), Section 5.²⁰ The FTC Act prohibits unfair or deceptive acts or practices in or affecting commerce. Although the act does not explicitly grant FTC the specific authority to protect privacy, FTC has interpreted it to apply to deceptions or violations of written privacy policies. For example, if a retailer's written privacy policy stated customers' personal information would not be shared with resellers and the retailer later sold information to such parties, FTC could bring an enforcement action against the retailer for unfair and deceptive practices.

Some states also have enacted laws designed to regulate resellers' sharing of personal information about consumers. For example, in 2018, Vermont passed a law that contains, among other requirements, consumer protection provisions related to data brokers.²¹ Among other things, the law requires data brokers to register annually and prohibits the acquisition and use of brokered personal information through certain means and for certain uses.

¹⁹Pub. L. No. 99-508, 100 Stat. 1848 (1986) (codified as amended in scattered sections of 18 U.S.C.).

²⁰15 U.S.C. § 45. Section 5 of the FTC Act, as originally enacted, only related to "unfair methods of competition." The Wheeler-Lea Act, passed in 1938, expanded the Commission's jurisdiction to include "unfair or deceptive acts or practices." Wheeler-Lea Amendments of 1938, Pub. L. No. 75-447, 52 Stat. 1111 (1938).

²¹Vt. STAT. ANN. tit. 9, §§ 2430, 2433, 2446 and 2447. Data broker means 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).

Gaps Exist in the Consumer Privacy Framework

The scope of consumer privacy protections provided under federal law has remained narrow in relation to (1) individuals' ability to access, control, and correct their personal data; (2) collection methods and sources and types of consumer information collected; (3) new technologies; and (4) some regulatory authorities. The examples in the following sections are drawn from our earlier reports and remain pertinent today.

Federal Law Provides Individuals Limited Ability to Access, Control, and Correct Their Personal Data

In our 2013 report, we found that no federal statute that we examined generally requires resellers to allow individuals to review personal information (intended for marketing purposes), control its use, or correct it. The Fair Information Practice Principles state that individuals should be able to know about and consent to the collection of their information and have the right to access the information, request correction, and challenge the denial of those rights.

We also reported in 2013 that no federal statute provides consumers the right to learn what information is held about them and who holds it for marketing or look-up purposes. FCRA provides individuals with certain access rights, but only when information is used for credit eligibility purposes. And GLBA's provisions allowing consumers to opt out of having their personal information shared with third parties apply only in specific circumstances. Otherwise, under federal law, individuals generally cannot require that their personal information not be collected, used, and shared. Also, no federal law we examined provides correction rights (the ability to have resellers and others correct or delete inaccurate, incomplete, or unverifiable information) for marketing or look-up purposes.

Laws Largely Do Not Address Data Collection Methods, Sources, and Types

Our 2013 report also found that federal privacy laws are limited in addressing the methods by which, or the sources from which, resellers collect and aggregate personal information, or the types of information collected for marketing or look-up purposes. The Fair Information Practice Principles state that personal information should be relevant, limited to the purpose for which it was collected, and collected with the individual's knowledge or consent.

Federal laws generally do not govern the methods resellers may use to collect personal information. For instance, resellers, advertisers, and others use software to search the web for information about individuals and extract and download bulk information from websites with consumer

information. Resellers or retailers also may collect information indirectly (by combining information from transactions).

Current federal law generally allows resellers to collect personal information from sources such as warranty registration cards and surveys and from online sources such as discussion boards, social media sites, blogs, and web browsing histories and searches. Current federal law generally does not require disclosure to consumers when their information is collected from these sources.

The federal laws that address the types of consumer information that can be collected and shared are not comprehensive. Under most circumstances, information that many people may consider very personal or sensitive can be collected, shared, and used for marketing. This can include information about physical and mental health, income and assets, political affiliations, and sexual habits and orientation. For health information, HIPAA rule provisions generally apply only to covered entities, such as health care providers.

Privacy Framework Largely Has Not Kept Pace with Changes in Technology

The current privacy framework does not fully address new technologies such as facial recognition technology, privacy issues raised by online tracking and mobile devices, and activities by financial technology firms. The original enactment of several federal privacy laws predates these trends and technologies. But in some instances existing laws have been interpreted to apply to new technologies. For example, FTC has taken enforcement actions under COPPA and revised the statute's implementing regulations to account for smartphones and mobile applications.

Facial Recognition Technology

One example of how privacy law has not kept pace with changes in technology is the use of facial recognition technology, which involves the collection of facial images and may be employed in a wide range of commercial applications. In our 2015 report we concluded that the future trajectory of this technology raised questions about consumer privacy.²² We found that federal law does not expressly address the circumstances under which commercial entities can use facial recognition technology to identify or track individuals, or when consumer knowledge or consent should be required for the technology's use. Furthermore, in most

²²GAO-15-621.

Activities by Financial
Technology Firms

contexts federal law does not address how personal data derived from the technology may be used or shared. The privacy issues stakeholders raised about facial recognition technology and other biometric technologies in use at the time of our 2015 report served as yet another example of the need to adapt federal privacy law to reflect new technologies. As such, we reiterated our 2013 recommendation that Congress strengthen the current consumer privacy framework to reflect the effects of changes in technology and the marketplace.

The rise of financial services provided by nonfinancial firms—often referred to as fintech—is another example of how new technology may create privacy concerns. For example, fintech lenders offer a variety of loans such as consumer and small business loans and operate almost exclusively online. In our 2018 report, we noted that while these lenders may still assess borrowers' creditworthiness with credit scores, they also may analyze large amounts of additional or alternative sources of data to determine creditworthiness.²³ We also found that some fintech firms may collect more consumer data than traditional lenders. For example, fintech lenders may have sensitive information such as consumers' educational background or utility payment information, and according to certain stakeholders, these data may contain errors that cannot be disputed by consumers under FCRA.

Furthermore, some data aggregators may hold consumer data without disclosing what rights consumers have to delete the data or prevent the data from being shared with other parties. A leak of these or other data held by fintech firms may expose characteristics that people view as sensitive. GLBA generally requires fintech firms and traditional financial institutions to safeguard nonpublic personal information about customers.²⁴ Our 2018 report discussed that some fintech firms use new technologies or mobile device features to mitigate data privacy risks and that some regulators have issued guidance to consumers publicizing practices that help maintain privacy when using online products and services, including those provided by fintech firms. Regulators also have issued GLBA guidance to businesses, including fintech firms, recommending that they adopt policies and procedures to prevent, detect, and address privacy threats.

²³GAO-18-254.

²⁴GLBA restricts, with some exceptions, the disclosure of nonpublic information by companies defined as financial institutions. See 15 U.S.C. §§ 6801-6802.

Internet Privacy Issues

Online tracking. In our 2013 report, we found that no federal privacy law explicitly addresses the full range of practices to track or collect data from consumers' online activity. Cookies allow website operators to recall information such as user name and address, credit card number, and purchases in a shopping cart. Resellers can match information in cookies and their databases to augment consumer profiles. Third parties also can synchronize their cookie files with resellers' files. Advertisers can use third-party cookies—placed on a computer by a domain other than the site being visited—to track visits to the websites on which they advertise. While current federal law does not, with some exceptions, explicitly address web tracking, FTC has taken enforcement actions related to web tracking under its authority to enforce the prohibition on unfair or deceptive acts. For example, in 2011, FTC settled charges with Google for \$22.5 million after alleging that Google violated an earlier privacy settlement with FTC when it misrepresented to users of Apple's Safari web browser that it would not track and serve targeted advertisements to Safari users.²⁵ Google agreed to disable its advertising tracking cookies.

Mobile devices. In 2013, we also explained that no federal law comprehensively governs applications software for mobile devices. Application developers, mobile carriers, advertisers, and others may collect an individual's information through services provided on a mobile device. However, FTC has taken enforcement action against companies for use of mobile applications that violate COPPA and FCRA.²⁶ The agency also has taken action under the FTC Act.²⁷ We and others have reported that the capability of mobile devices to provide consumer's location engenders privacy risks, particularly if companies use or share

²⁵*United States v. Google Inc.*, No. CV 12-04177-SI, 2012 WL 5833994 (N.D. Cal. Nov. 16, 2012).

²⁶FTC settled charges that a social networking service deceived consumers when it collected information from children under 13 through its mobile application in violation of COPPA. See *United States v. Path, Inc.*, No. C13-0448 (N.D. Cal. Jan. 31, 2013). FTC also settled charges that a company compiled and sold criminal record reports through its mobile application and operated as a consumer reporting agency, in violation of FCRA. See *In the Matter of Filiquarian Publishing, LLC*, FTC File No. 112 3195 (Apr. 30, 2013).

²⁷In addition to the alleged COPPA violation, Path allegedly deceived users by collecting personal information from their mobile address books without their knowledge and consent. See *United States v. Path, Inc.*, No. C13-0448 (N.D. Cal. Jan. 31, 2013).

location data without consumers' knowledge.²⁸ ECPA might not apply if location data were not deemed content and would not govern entities that are not covered by ECPA. But FTC could pursue enforcement action if a company's collection or use of the information violated COPPA.

More recently, in January of this year, we issued a report on internet privacy that reinforces what we reported in 2013.²⁹ To varying extents, internet content providers and internet service providers collect, use, and share information from their customers to enable their services, support advertising, and for other purposes. Consumers access such services through mobile phones and tablets, computers, and other internet-connected devices. However, there is no comprehensive federal privacy statute with specific standards. FTC has been addressing internet privacy through its unfair and deceptive practices authority, among other statutes, and other agencies have been addressing this issue using industry-specific statutes. We concluded that recent developments regarding internet privacy suggest that this is an appropriate time for Congress to consider comprehensive internet privacy legislation. To address such privacy concerns, states and other countries have adopted privacy rules. For example, the European Union's General Data Protection Regulation, which came into force in May 2018, is a set of privacy rules that give consumers control over the collection, use, and sharing of their personal information, and California passed its own privacy law in June 2018 that becomes effective in 2020.³⁰

Regulatory Authorities
under Current Law May
Be Limited

In February of this year, we reported that FTC does not have civil penalty authority for initial violations of GLBA's privacy and safeguarding requirements, which, unlike FCRA, includes a provision directing federal regulators and FTC to establish standards for financial institutions to

²⁸Risks included disclosure to third parties for unspecified uses, tracking of consumer behavior, and identity theft. See GAO, *Mobile Device Location ID: Additional Federal Actions Could Help Protect Consumer Privacy*, [GAO-12-903](#) (Washington, D.C.: Sept. 11, 2012). A Federal Communications Commission report also noted privacy risks. See Federal Communications Commission, *Location-Based Services: An Overview of Opportunities and Other Considerations* (Washington, D.C.: May 2012).

²⁹[GAO-19-52](#).

³⁰California's law generally will require companies to report to customers, upon their request, the categories of personal information they collected about the customer, the business or commercial purpose for collecting and selling such personal information, and what categories of third parties received it.

protect against any anticipated threats or hazards to the security of customer records.³¹ To obtain monetary redress for these violations, FTC must identify affected consumers and any monetary harm they may have experienced. However, harm resulting from privacy and security violations (such as a data breach) can be difficult to measure and can occur years in the future, making it difficult to trace a particular harm to a specific breach. As a result, FTC lacks a practical enforcement tool for imposing civil money penalties that could help to deter companies from violating data security provisions of GLBA and its implementing regulations. We recommended that Congress consider giving FTC civil penalty authority to enforce GLBA's safeguarding provisions.

Additionally, in our January 2019 report, we found that FTC had not yet issued regulations for internet privacy other than those protecting financial privacy and the internet privacy of children, which were required by law. FTC uses its statutory authority under the FTC Act to protect consumers from unfair and deceptive trade practices. For FTC Act violations, FTC may promulgate regulations but is required to use procedures that differ from traditional notice-and-comment processes and that FTC staff said add time and complexity. In addition, under this authority, FTC can generally only levy civil money penalties after a company has violated an FTC final consent order. In our recommendation that Congress consider developing comprehensive internet privacy legislation, we also suggested that such legislation consider providing rulemaking and civil money penalty authorities to the proper agency or agencies.

In summary, new technologies have vastly changed the amount of personal information private companies collect and how they use it. But our current privacy framework does not fully address these changes. Laws protecting privacy interests are tailored to specific sectors and uses. And, consumers have little control over how their information is collected, used, and shared with third parties for marketing purposes. As a result, current privacy law is not always aligned with the Fair Information Practice Principles, which the Department of Commerce and others have said should serve as the foundation for commercial data privacy. Thus, the privacy framework warrants reconsideration by Congress in relation to consumer interests, new technologies, and other issues.

³¹GAO-19-196.

Chairman Crapo, Ranking Member Brown, and Members of the Committee, this concludes my statement. I would be pleased to respond to any questions you may have.

GAO Contacts

For further information on this statement, please contact Alicia Puente Cackley at 202-512-8678 or cackleya@gao.gov. Contact points for our offices of Congressional Relations and Public Affairs may be found on the last page of this statement. In addition to the contact above, Jason Bromberg (Assistant Director), William R. Chatlos, Rachel DeMarcus, Kay Kuhlman (Assistant Director), Christine McGinty (Analyst in Charge), Barbara Roesmann, and Tyler Spunaugle contributed to this statement. Other staff who made key contributions to the reports cited in the testimony are identified in the source products.



This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's website (<https://www.gao.gov>). Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. To have GAO e-mail you a list of newly posted products, go to <https://www.gao.gov> and select "E-mail Updates."

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, <https://www.gao.gov/ordering.htm>.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on [Facebook](#), [Flickr](#), [Twitter](#), and [YouTube](#).
Subscribe to our [RSS Feeds](#) or [E-mail Updates](#). Listen to our [Podcasts](#).
Visit GAO on the web at <https://www.gao.gov>.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact FraudNet:

Website: <https://www.gao.gov/fraudnet/fraudnet.htm>

Automated answering system: (800) 424-5454 or (202) 512-7700

Congressional Relations

Orice Williams Brown, Managing Director, WilliamsO@gao.gov, (202) 512-4400,
U.S. Government Accountability Office, 441 G Street NW, Room 7125,
Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, DC 20548

Strategic Planning and External Liaison

James-Christian Blockwood, Managing Director, spel@gao.gov, (202) 512-4707
U.S. Government Accountability Office, 441 G Street NW, Room 7814,
Washington, DC 20548



Please Print on Recycled Paper.

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 Personix Customer Segmentation. See: Acxiom, Personix Home Page. Available at: <https://www.acxiom.com/what-we-do/consumer-segmentation-personix/>.

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.*, Springer 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-119¹⁰). 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-ai-go-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.

¹⁰OMB 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/revise_d_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.

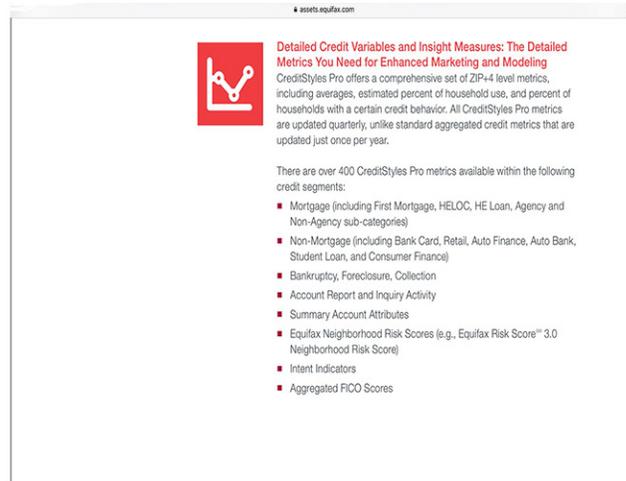
Aggregated FICO scores

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

CrediStyles 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 CrediStyles Pro metrics are updated quarterly, unlike standard aggregated credit metrics that are updated just once per year.

There are over 400 CrediStyles 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.

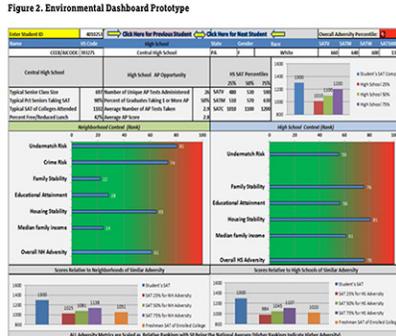
In an unregulated prediction marketplace, abuses can occur. Regulators should rightly be concerned about the uses of unregulated credit scores in the razor-thin line between the marketing of credit opportunity, and firm offers of credit.

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



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

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*.



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/revise/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/revise/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 agencies³⁷ 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/revise/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=.4172E3ec00cf.

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)

World Privacy Forum

**The Scoring of America: How
Secret Consumer Scores
Threaten Your Privacy and
Your Future**

By Pam Dixon and Robert Gellman
April 2, 2014



Brief Summary of Report

This report highlights the unexpected problems that arise from new types of predictive consumer scoring, which this report terms *consumer scoring*. Largely unregulated either by the Fair Credit Reporting Act or the Equal Credit Opportunity Act, new consumer scores use thousands of pieces of information about consumers' pasts to predict how they will behave in the future. Issues of secrecy, fairness of underlying factors, use of consumer information such as race and ethnicity in predictive scores, accuracy, and the uptake in both use and ubiquity of these scores are key areas of focus.

The report includes a roster of the types of consumer data used in predictive consumer scores today, as well as a roster of the consumer scores such as health risk scores, consumer prominence scores, identity and fraud scores, summarized credit statistics, among others. The report reviews the history of the credit score – which was secret for decades until legislation mandated consumer access -- and urges close examination of new consumer scores for fairness and transparency in their factors, methods, and accessibility to consumers.

About the Authors

Pam Dixon is the founder and Executive Director of the World Privacy Forum. She is the author of eight books, hundreds of articles, and numerous privacy studies, including her landmark Medical Identity Theft study and One Way Mirror Society study. She has testified before Congress on consumer privacy issues as well as before federal agencies. Robert Gellman Robert Gellman is a privacy and information policy consultant in Washington DC. (www.bobgellman.com.) He has written extensively on health, de-identification, Fair Information Practices, and other privacy topics. Dixon and Gellman's writing collaborations include a reference book on privacy *Online Privacy: A Reference Handbook*, as well as numerous and well-regarded privacy-focused research, articles, and policy analysis.

About the World Privacy Forum

The World Privacy Forum is a non-profit public interest research and consumer education group focused on the research and analysis of privacy-related issues. The Forum was founded in 2003 and has published significant privacy research and policy studies in the area of health, online and technical, privacy, self-regulation, financial, identity, and data brokers among other many areas. www.worldprivacyforum.org.

Contents

THE SCORING OF AMERICA:	6
HOW SECRET CONSUMER SCORES THREATEN YOUR PRIVACY AND YOUR FUTURE.....	6
INTRODUCTION	6
PART I: SUMMARY AND BACKGROUND	7
WHAT IS A CONSUMER SCORE?	8
WHO HAS A SCORE?	8
GAPS IN CONSUMER PRIVACY RIGHTS AND PROTECTIONS AROUND CONSUMER SCORING (AND WHY EXISTING LAWS DON'T ALWAYS APPLY).....	9
<i>Key Issue: Score Secrecy</i>	11
<i>Key Issue: Score Accuracy</i>	12
<i>Key Issue: Identity Theft and Consumer Scoring</i>	13
<i>Key Issue: Unfairness and Discrimination</i>	13
<i>Key Issue: Sensitive Health and Lifestyle Information and Consumer Scoring</i>	14
<i>Key Issue: Consent and Use of Consumer Data in Predictive Scores</i>	15
SCORES THEN: A HANDFUL OF FACTORS.SCORES NOW: THOUSANDS OF FACTORS	15
SCORING METHODS AND MODELS ARE OPAQUE.....	18
EXAMPLES AND NUMBERS OF CONSUMER SCORES.....	19
USES OF CONSUMER SCORES, REGULATION, AND MODERN ELIGIBILITY	19
DEJA VU: WHY THE HISTORY OF THE CREDIT SCORE IS IMPORTANT.....	21
SUMMARY OF FINDINGS AND RECOMMENDATIONS.....	23
<i>Findings:</i>	23
<i>Recommendations:</i>	24
ADVICE FOR CONSUMERS:	26
PART II. CONSUMER SCORES: WHAT GOES INTO THEM, AND HOW THEY ARE MADE. 27	27
DEFINING CONSUMER SCORES MORE DEEPLY	27
MAKING A CONSUMER SCORE, STEP 1: WHAT CONSUMER INFORMATION GETS PUT INTO A CONSUMER SCORE?	30
TRADITIONAL SCORE INGREDIENTS: CREDIT SCORES	30
MODERN SCORE-MAKING INGREDIENTS: RAW CONSUMER DATA IN THE DIGITAL AGE.....	32
CONSUMER DATA AVAILABLE FOR PURCHASE AND USE IN ANALYTICS	33
<i>Demographic Information:</i>	33
<i>Contact Information:</i>	34
<i>Vehicles:</i>	34
<i>Lifestyle, Interests and Activities data (including medical):</i>	34
<i>Financial and Economic – Property and Assets data:</i>	36
<i>Financial and Credit data:</i>	37
SCORING MODELS: HOW THE CONSUMER SCORES ARE MADE, PART TWO	38
ALGORITHM, INC.	38
LAYERING SCORING MODELS	40
HOW MANY VARIABLES?	41
POLICY QUESTIONS.....	41
PART III. THE CONSUMER SCORES	42

CATEGORY: FINANCIAL AND RISK SCORES	43
<i>ChoiceScore</i>	43
<i>Median Equivalency Score (Summarized credit statistics)</i>	44
<i>Risk IQ Score</i>	45
<i>Consumer Profitability Score</i>	45
<i>Job Security Score</i>	47
<i>Consumer Prominence Indicator Score</i>	47
<i>Discretionary Spending Index Score</i>	48
<i>Invitation to Apply Score</i>	48
<i>Charitable Donor Score</i>	49
<i>Household Segmentation Scoring Systems (Personicx, Mosaic, etc.)</i>	50
<i>Collection and Recovery Scores</i>	50
<i>Churn Scores</i>	51
CATEGORY: FRAUD SCORES.....	52
<i>FICO Falcon Fraud Manager</i>	53
<i>Other Fraud Scores</i>	54
CATEGORY: CUSTOM SCORES	55
<i>The Emergence of Custom Scores and the Pregnancy Predictor Score Example</i>	55
CATEGORY: REGULATED CREDIT AND FINANCIAL SCORES.....	57
<i>FICO Score</i>	58
<i>Vantage Score</i>	58
<i>Beacon Score</i>	59
<i>Small Business Intelliscore</i>	59
<i>Tenant Scores</i>	60
CATEGORY: IDENTITY AND AUTHENTICATION SCORES	60
<i>ID Analytics ID Score</i>	60
<i>Insurance Scores</i>	61
<i>Category -- Health Scores</i>	61
<i>Affordable Care Act Individual Health Risk Score</i>	62
<i>FICO Medication Adherence Score (MAS)</i>	63
<i>Frailty Scores: General</i>	65
<i>CMS Frailty Adjustment Score</i>	66
<i>Hopkins Frailty Score</i>	66
<i>Other Health Scores</i>	66
<i>Personal Health Scores: WebMD, others</i>	67
<i>Resource Utilization Group Scores</i>	68
<i>SF-36 Form</i>	69
<i>Complexity Scores</i>	69
CATEGORY – SMART GRID AND ENERGY SCORES.....	70
<i>Peer-to-Peer Energy People Meter Score (EPM)</i>	70
CATEGORY - SOCIAL SCORING.....	72
<i>Klout Score</i>	73
<i>Employment Success Score</i>	75
TAX RETURN SCORES	75
CATEGORY – LAW ENFORCEMENT SCORES, INCLUDING POLICE, TRANSPORTATION, SAFETY, AND	
OTHER.....	76
<i>Automated Targeting System Score</i>	76
<i>Richard Berk Algorithm</i>	77
<i>Youth Delinquency Scores</i>	77
<i>Predictive Anti-Fraud Scores: US Postal Service Office of Inspector General</i>	77

CATEGORY -- ENVIRONMENTAL SCORES 78
 EPA Health Risk Score 78
 AIQ Green 79
CATEGORY - OTHER CONSUMER SCORES..... 79
 Potential Scores..... 80
 Non-included Scores:..... 80

PART IV: THE HISTORY OF SCORING: HOW THE CREDIT SCORE AND CONSUMER SCORES BEGAN, AND WHY IT IS RELEVANT TODAY..... 80
 THE BEGINNINGS OF CONSUMER SCORING..... 80
 CREDIT SCORING BECOMES ENTRENCHED 81
 HOW THE FORMERLY SECRET CREDIT SCORE BECAME AVAILABLE TO THE PUBLIC..... 82
 ONGOING DISCLOSURE CHALLENGES AND OTHER ISSUES WITH CONSUMER CREDIT SCORES..... 83

CONCLUSION..... 84

ABOUT THIS REPORT AND CREDITS 86

APPENDIX A 87
 TIMELINE: HIGHLIGHTS IN SCORING 87

APPENDIX B 89
 SCORE TAXONOMY 89



April 2, 2014 | Pam Dixon and Robert Gellman

The Scoring of America: How secret consumer scores threaten your privacy and your future

Introduction

To score is human. Ranking individuals by grades and other performance numbers is as old as human society. Consumer scores — numbers given to individuals to describe or predict their characteristics, habits, or predilections — are a modern day numeric shorthand that ranks, separates, sifts, and otherwise categorizes individuals and also predicts their potential future actions.

Consumer scores abound today. Credit scores based on credit files receive much public attention, but many more types of consumer scores exist. They are used widely to predict behaviors like, spending, health, fraud, profitability, and much more. These scores rely on petabytes of information coming from newly available data streams. The information can be derived from many data sources and can contain financial, demographic, ethnic, racial, health, social, and other data.

The Consumer Profitability Score, Individual Health Risk Score, Summarized Credit Statistics that score a neighborhood for financial risk, fraud scores, and many others seek to predict how consumers will behave based on their past behavior and characteristics.

Predictive scores bring varying benefits and drawbacks. Scores can be correct, or they can be wrong or misleading. Consumer scores — created by either the government or the private sector — threaten privacy, fairness, and due process because scores, particularly opaque scores with unknown ingredients or factors, can too easily evade the rules established to protect consumers.

The most salient feature of modern consumer scores is the scores are typically secret in some way. The existence of the score itself, its uses, the underlying factors, data sources, or even the score range may be hidden. Consumer scores with secret factors, secret

sources, and secret algorithms can be obnoxious, unaccountable, untrustworthy, and unauditible. Secret scores can be wrong, but no one may be able to find out that they are wrong or what the truth is. Secret scores can hide discrimination, unfairness, and bias. Trade secrets have a place, but secrecy that hides racism, denies due process, undermines privacy rights, or prevents justice does not belong anywhere.

Broader transparency for consumer scores with limited secrecy may offer a middle ground. Knowing the elements but not necessarily the weights of a scoring system provides a partial degree of openness and reassurance. Knowing that there is a scoring system and how and when it is used helps. Knowing the source and reliability of the information used to make a score helps. Being able to challenge a score and correct the data on which it is based helps. Knowing that some types of information will not be used for scoring helps. Knowing that data collected for one purpose will not be used for another or in violation of law helps. Knowing that the person running the scoring system is accountable in a meaningful way helps.

The history of the credit score provides a useful model for the new batch of predictive consumer scores. Developed in the 1950s, the credit score became part of consumer credit granting. The credit score was largely secret to the consumers that it scored and affected until 2000, when a long and well-documented history of unfair uses and abuses finally culminated in the credit score being made available to consumers. Eventually, public pressure caused the credit score's use and even its underlying factors to become public. The use of factors such as race, gender, and religion were prohibited and this was spelled out in detail in law.

No similar protections exist for most consumer scores today. Consumer scores are today where credit scores were in the 1950s. Data brokers, merchants, government entities, and others can create or use a consumer score without notice to consumers. For various reasons laws governing credit scores do not typically extend protection to the new consumer scores. We need rules that will make consumer scores fair, accountable, accurate, transparent, and non-discriminatory.

This report discusses and explores consumer scores, what goes into them and how they are made, how they are used, the regulations in place that control some but not most new consumer scores, and how scores affect broader privacy and fairness issues. The discussion of findings and recommendations points toward solutions and reforms that are needed.

Part I: Summary and Background

As the numbers of predictive consumer scores increase and their usage expands, Americans face a future that may be shaped in significant ways by consumer scores. By itself, consumer scoring is not necessarily good or bad. Scoring orders consumers along a

mathematically defined scale. However, scoring has the prospect of being used to affect individuals in significant ways that may not always be fair or even legal.

If a predictive score unknown to a consumer determines how that consumer is treated, the results may not be acceptable to the American public. The quality and relevance of the data used, the transparency of the methodology of how the score was created, plus the reasonableness of the application of the consumer score are the major factors that determine the fairness of any scoring activity. These issues should be the central focus of any policy debate about consumer scoring. These issues also suggest the elements of best practices that should apply to consumer scoring.

What is a Consumer Score?

With this report, the World Privacy Forum introduces a term: *consumer scores*. Consumer scores – the ones we discuss in this report – are built using predictive modeling. Predictive modeling uses copious amounts of information fed through analytical methods to predict the future, based on past information.

Predictive consumer scores are important because they affect the lives, privacy, and wellbeing of individuals. Many people know about credit scores, but few know about the broader range of new consumer scores. Consumer scores are already abundant and are in active use. Consumer scores are not just an online phenomenon. Consumer scores are found in a wide array of “offline” arenas, including businesses, health care providers, financial institutions, law enforcement, retail stores, federal and state government, and many other locations. Some social consumer scores may have online applications, but mostly, consumer scores are not solely focused on just online activities. And unlike credit scores, consumer scores remain largely secret and unregulated.

The World Privacy Forum defines a consumer score 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.

Who has a Score?

Consumer scoring is already more widespread than most people realize. Many hundreds of consumer scores exist, perhaps thousands. How many Americans have them? Almost

all do. Minors are less likely to be scored than adults, although they, too can have or influence some consumer scores. For example, household scores often reflect interests and activities of minors.¹

Among American adults, each individual with a credit or debit card or a bank account is likely to be the subject of one or more scores.² Many individuals signed up under the Affordable Care Act have a score.³ Individuals who buy airline tickets have a score.⁴ Individuals who make non-cash purchases at large retail stores likely have a score.⁵

Scores such as the medication adherence score, the health risk score, the consumer profitability score, the job security score, collection and recovery scores, frailty scores, energy people meter scores, modeled credit scores, youth delinquency score, fraud scores, casino gaming propensity score, and brand name medicine propensity scores are among the consumer scores that score, rank, describe, and predict the actions of consumers.

In short, almost every American over the age of 18 has at least one score, and most adult Americans have many scores. An individual could easily be the subject of dozens or even hundreds of secret consumer scores. We can safely predict that there will be many more consumer scores in the future. Fed by the masses of consumer data now available, consumer scoring is quickly becoming a form of shorthand to make sense of a sea of information.

Gaps in Consumer Privacy Rights and Protections around Consumer Scoring (And why existing laws don't always apply)

This report's analysis is that many new consumer scores exist, and many of these new scores do not appear to fall under the narrow protections offered by the Fair Credit Reporting Act⁶ or the Equal Credit Opportunity Act⁷ for a variety of reasons. Scores built from factors outside a formal credit bureau file, scores designed to predict the behavior of groups of people instead of individuals, and new scores in emerging and unregulated areas may all fall outside of existing protections.

For example:

¹ A good example of this would be an *aggregate credit score*, which scores neighborhoods versus individual consumers.

² Likely scores for an adult with a credit or debit card would be real-time or near real-time fraud and/or identity scores.

³ This is the ACA Health Risk Score. Specific scores are discussed in detail in Part III of the report.

⁴ This score is generated by the federal Transportation Security Agency. This score is discussed in Part III of the report.

⁵ See the Custom Scores section of Part III.

⁶ 15 U.S.C. § 1681 et seq.

⁷ 15 U.S.C. 1691 et seq.

- Energy consumption scores, churn scores, and identity scores are not likely to fall under the FCRA and other laws as currently written. This is because those scores do not meet the layers of qualifications that would bring them under the FCRA.
- Scores that identify the approximate credit capacity of neighborhoods instead of individuals also appear to be unregulated. This is because the FCRA applies to individuals, not neighborhoods. Formal credit scores may only be used in certain circumstances, for example, for extending a firm offer of credit or insurance. Credit scores cannot be used for general marketing purposes, but aggregate credit statistics tied to a neighborhood do not appear to be subject to the same restrictions for the reasons mentioned. Lead generation is not the same thing as a formal offer of credit under the FCRA.
- Risk scores -- like health risk scores -- that use broad demographic information and aggregate financial statistics about consumers to assess financial or other risks (credit bureau files are not typically used) also don't appear to fall under the layers of requirements that would bring them under current regulation.
- The Equal Credit Opportunity Act requires credit scoring systems to not use race, sex, marital status, religion, or national origin as factors comprising the credit score. But this law applies only to what is today a narrowly defined credit scoring system. Other scores which fall outside of the narrow definitions -- like identity, fraud, churn, and other predictive scores can incorporate factors that would in other situations be considered prohibited factors to use.

As a result, consumers may have scant rights to find out what their non-FCRA consumer scores are, how the scores apply to them and with what impact, what information goes into a score, or how fair, valid, or accurate the score is. Even if the input to a score is accurate, consumers do not know or have any way to know what information derived from their lifestyle, health status, and/or demographic patterns is used to infer patterns of behavior and make decisions that affect their lives.

Further, consumers can have difficulty exercising basic Fair Information Principles for many if not most new consumer scores.⁸ Fair Information Principles form the base for most global privacy law today, including some US privacy laws. However, those who create unregulated scores have no legal obligation to provide Fair Information Practices or due process to consumers.

⁸ In this report we refer to Fair Information Practices as a baseline and standard by which to judge consumer scoring. FIPS are an established set of eight principles guiding privacy. The U.S. has ratified the FIPs twice since the 1970s. The FIPs include the principles of collection limitation, data quality, purpose specification, use limitation, security safeguards, openness, individual participation, and accountability. See Robert Gellman, *Fair Information Practices: A Basic History*. <<http://bobgellman.com/rg-docs/rg-FIPShistory.pdf>>. A brief introduction is here <<http://www.worldprivacyforum.org/2008/01/report-a-brief-introduction-to-fair-information-practices/>>.

These significant gaps in consumer protections mean that consumer scores may include or use discriminatory factors in their composition, or uncorrected or otherwise inaccurate information could be included. Scores developed to characterize individuals or predict their behaviors need to provide fairness and due process. The credit score is already subject to some regulation, but that is not to say that consumers would not benefit from better rules for credit scores.

There is a great need to examine the effects and fairness for all consumer scores now in use. Intriguing possibilities exist that a certain stratification of consumer experience based on opportunities offered to each consumer could become commonplace. Victims of identity theft, for example, may consistently receive different and less desirable marketing treatment than individuals with clean credit scores, even if most other demographic factors are similar.

Disparate treatment, even in the area of marketing opportunities granted to consumers, raises many questions, questions that the general field of risk-based pricing has raised. Oddly, direct marketing lists and activities have the potential to strike deeply into the lives of individuals in quirky ways that can have an impact on consumer lifestyle. Much remains to be learned about the impact of consumer scoring in the direct marketing arena, as well as eligibility issues and edge-eligibility issues like scores for identity and authentication.

Some of the specific issue areas around gaps in protection and information fairness in scoring including the following.

Key Issue: Score Secrecy

There are good reasons why credit scores are not secret anymore, nor are the foundational factors that comprise the score. By law, consumers have the right to see credit scores now. This report finds that with the exception of the credit score and a handful of other consumer scores, at this time, secrecy is the hallmark of many consumer scores.

The factors that go into most scores are usually secret, the models used are usually secret, and in many cases, the score itself is also secret. This report's analysis is that consumer scores that are risk scores bear many similarities to scores regulated under the FCRA. Yet industry treats these risk scores as falling outside the FCRA so that consumers have none of the rights guaranteed by the FCRA.

Consumers have no formal rights to find out what their non-FCRA consumer scores are, or how these scores affect their lives. Victims of identity theft and other individuals may have errors or omissions affecting their scores, but they do not necessarily have a right to see or correct the scores. Even if information is accurate, consumers do not know or have any way to know how companies use information derived from their lifestyle, health status, and/or demographic patterns to infer patterns of behavior and make decisions that affect their lives. Unseen scores can affect consumers' marketplace experiences and much more.

Key Issue: Score Accuracy

Because consumers do not have the right to correct or control what personal information goes into a consumer score as an attribute or factor, the accuracy of the scores is suspect. Consumers also do not have the right to see the scoring models used to make the score, nor do they typically have information about the model validity. Because of the lack of transparency, consumers cannot be assured of the reliability, fairness, or legality of scoring models. Inaccurate, incomplete, and illegal factors may be used today to make decisions about consumers without any oversight or redress.

Credit scores offer a useful model here. Credit scores are based on credit reports. Credit report accuracy has been the subject of substantial, meaningful scrutiny over decades. The CFPB, in its 2012 study on credit reports, noted the significant problems with inaccuracy that occur:

Given the volume of data handled, the challenges of matching tradelines to the correct consumer files, and the number and variety of furnishers, inaccuracies in some credit files inevitably occur. Inaccuracies in credit files and credit reports can occur where information that does not belong to a consumer is attached to his or her file, where information belonging to a consumer is omitted from the file, or where there are factual inaccuracies in trade line or other information in the consumer's file. Some of these inaccuracies can be attributed to matching challenges in assigning a trade line to a consumer's file. Other causes of inaccuracies include data and data entry errors, NCRA system or process inaccuracies, furnisher system or process inaccuracies, identity fraud, or time lags.⁹

In a ten-year, Congressionally-mandated study published in 2013, the FTC found that overall “one in five consumers had an error on at least one of their three credit reports.”¹⁰ The FTC found that these credit report errors did impact the credit score. The FTC found that, specifically:

- “Slightly more than one in 10 consumers saw a change in their credit score after the CRAs modified errors on their credit report; and;
- Approximately one in 20 consumers had a maximum score change of more than 25 points and only one in 250 consumers had a maximum score change of more than 100 points.”¹¹

⁹ Consumer Financial Protection Bureau, *Key Dimensions and Processes in the U.S. Credit Reporting System* at 23 (Dec. 2012).

¹⁰ Federal Trade Commission, *Section 319 of the Fair and Accurate Credit Transactions Act of 2003: Fifth Interim Federal Trade Commission Report to Congress Concerning the Accuracy of Information in Credit Reports*, <<http://www.ftc.gov/reports/section-319-fair-accurate-credit-transactions-act-2003-fifth-interim-federal-trade>>.

¹¹ *Id.*

Errors in credit scores abound, and credit scores are based on credit reports, which also are subject to significant errors. If a transparent score with few factors has these kinds of errors, what about consumer scores? Consumer scoring relies on dozens, hundreds, or thousands of data elements that have no standards for accuracy, timeliness, or completeness. The quality of data matters: errors in data used to make a score create a score that is not predictive. With thousands of factors, error rates and false readings become a big issue.

Key Issue: Identity Theft and Consumer Scoring

Victims of identity theft – both financial and medical forms of the crime -- may have significant and stubbornly ongoing errors or omissions affecting their scores. ID theft victims can be seriously affected by identity scoring because their identity scores and fraud scores may be incorrect as a direct result of criminal activity. This can cause a range of problems from being denied services to being tagged as a potential fraudster. Yet even this vulnerable group has no right to see or correct many consumer scores.

Key Issue: Unfairness and Discrimination

One of the fundamental policy issues regarding scoring activities is the question of what characteristics it is appropriate to use in scoring consumers. In the world of home loans, ECOA has answered that question. But in the world of direct marketing, this area is nearly without boundaries. In a prescient early critique of scoring policy, Columbia University professor Noel Capon wrote in 1982:

Since prediction is the sole criterion for acceptability, any individual characteristic that can be scored, other than obviously illegal characteristics, has potential for inclusion in a credit scoring system.¹²

As a bewildering plethora of new databases of consumer information become available, these databases may be scored in various ways by being run through one or more scoring models. More databases of consumer information fundamentally can mean more potential scores, and more potential characteristics to score.

The Equal Credit Opportunity Act protects consumers from invidious discrimination in formal credit granting situations. Notably, the ECOA requires that credit scoring systems may not use race, sex, marital status, religion, or national origin as factors comprising the score. The law allows creditors to use age, but it requires that seniors be treated equally.¹³

¹² Noel Capon. *Credit scoring systems: a critical analysis*, 46 *Journal of Marketing* 82-91(1982).

¹³ For more information, see Federal Trade Commission, *How Credit Scores Affect the Price of Credit and Insurance*, <<http://www.consumer.ftc.gov/articles/0152-how-credit-scores-affect-price-credit-and-insurance>>.

But in the modern consumer scores, marital status – a protected factor under ECOA – is commonly used as a consumer score factor. Consumer scores may also contain underlying factors of race, sex, and religion without disclosure to consumers. In some cases, health factors may also be included in scores, for example, if a person smokes, or has a chronic illness. (See the section on Factors below for an example of a score that incorporates smoking and ethnicity).

As discussed in Part II of this report, a single score is often created from the admixture of more than 600 to 1,000 to even 8,000 individual factors or data streams. These factors can include race, religion, age, gender, household income, zip code, presence of medical conditions, zip code + 4, transactional purchase information from retailers, and hundreds more data points about individual consumers.¹⁴ Therefore, one individual score can have the potential to contain hidden factors that range from bland – like mail order buyer of sports goods -- to quite sensitive – like ethnicity.

A score designed to assess or assign consumer value to a business could easily include factors that would be entirely unacceptable or that, in the context of either the Equal Credit Opportunity Act (ECOA) or the Fair Credit Reporting Act, would be flatly illegal. If ECOA factors are present in consumer scores, in most cases it would be difficult or impossible for consumers to find out if the scoring system or its factors were secret.

While carefully directed and controlled use of credit scoring and credit automation has reduced some discriminatory practices, new consumer scoring that uses elements that correlate with prohibited factors such as race can reintroduce discrimination and hide the effects behind a secret or proprietary screen that falls entirely outside of current consumer protection regulations. This is not acceptable.

Key Issue: Sensitive Health and Lifestyle Information and Consumer Scoring

Health scores already exist. This category of score deserves special attention and scrutiny. Some health scores are used in the HIPAA context, some are used outside the HIPAA context. The health scores used outside the HIPAA context are of most concern. Actuaries already use some new consumer scores to underwrite risk, for example, the Brand Name Medicine Propensity Score from a health category and the Underbanked Indicator from the financial category.¹⁵ Scores can contain health information as hidden information within the score, and used for health purposes, or used for non-health related purposes such as marketing, or risk scoring. Many consumers with chronic health conditions would object strenuously to having their financial risk be determined by their health status. While health risk may be very predictive in a score, is it fair to use without consumer knowledge?

¹⁴ Part II contains a substantive list of scoring factors.

¹⁵ For more details about this issue, see the Health Score discussion in Part III of the report. See also Tim Hill, *Predictive Modeling Topic*, Future of Preferred Underwriting, Society of Actuaries, August 25-27, 2013.

Just because a score contains information about a consumer's health status, it does not mean the score will be subject to the federal health privacy rule (HIPAA). In fact, much of health information available for commercial use outside of the healthcare environment falls outside the scope of HIPAA. HIPAA, for example, provides no consumer privacy rights over health data held by list and data brokers.

Health information often leaks outside of HIPAA protections when it is revealed by consumers through surveys, website registrations, and other online activities. After a consumer reveals his or her health information to a non-HIPAA third party, that information is considered out of HIPAA's bounds. It is in this way that consumers' most sensitive health information can wind up used as fodder for a consumer score, with unknown consequences.

Consumer scores that use health or other sensitive information such as sexual orientation as factors need close examination for fairness, and consumers need rights over whether their health information is used in predictive scores, whether for marketing or any other purpose.

Key Issue: Consent and Use of Consumer Data in Predictive Scores

If a consumer fills out a registration on a health-related web site or a consumer warranty card that accompanied a purchase, the consumer did not give *informed consent* that the information can be used in downstream consumer scoring in ways that affect the consumer's marketplace opportunities. A buried statement in an unread privacy policy that "we may share your information for marketing purposes with third parties" is not informed consent to allow unfettered use information for predictive scoring.

Does making a purchase with a credit or debit card at a retailer grant consent for use of a consumer's purchases and other information to be used in a score? Part II of this report contains a detailed discussion of what kinds of information go into consumer scores. Many individuals would be quite surprised to learn just how the details of their lives are fodder for scores they may never see or have access to – and did not knowingly consent to. The issue of consent becomes increasingly important for scores that affect any kind of eligibility, such as jobs, credit, insurance, identity verification, or other significant opportunities.

Scores Then: A Handful of Factors.Scores Now: Thousands of Factors

The research for this report found that consumer scores may rely on hundreds or thousands of pieces of consumer information coming from many different data sources. This report identifies a large roster of raw consumer data that includes demographic information like age, race, gender, ethnicity, and home address as well as religion, mobile phone number, online and offline purchase history, health conditions like Alzheimer's, diabetes, and multiple sclerosis, as well as intimate financial details such as net worth,

card holder information, low or high-end credit scores, money market funds, ages of children, and a great deal more.

Statistical scoring methods rely on the increasing availability of large amounts of new source data from social media, the web in general, and elsewhere. The input for consumer scores can include information that is mostly unobjectionable or public. But, as discussed, consumer scores also can incorporate highly sensitive information that in other contexts could be used in a prejudicial, unfair, or unethical way in making decisions about consumers. Some data, such as social media data, can be unobjectionable in one context, but inappropriate as a factor, for example, in credit decisioning models.

An example, and a fairly common one, is of a predictive model that a major US health insurer worked with an analytics company to create. The idea was to determine whether or not publicly available consumer data could enhance the quality and effectiveness of their predictive risk models. They tested approximately 1,500 factors at the household level and found that the consumer information that showed the most value in predicting individual level risk included:

- Age of the Individual
- Gender
- Frequency of purchase of general apparel
- Total amount from inpatient claims
- Consumer prominence indicator
- Primetime television usage
- Smoking
- Propensity to buy general merchandise
- Ethnicity
- Geography – district and region
- Mail order buyer - female apparel
- Mail order buyer - sports goods¹⁶

Those unfamiliar with predictive models can find it surprising to learn that information about purchasing sporting goods can become a part of a predictive score for a health insurer. But the factors used in this example are not surprising factors to find in a modern predictive consumer score model. This is actually a fairly short list compared to some models with thousands of factors.

The raw source material for the factors fed into consumer scores comes from sources such as:

- Retailers and merchants via Cooperative Databases and Transactional data sales & customer lists.
- Financial sector non-credit information (PayDay loan, etc.)

¹⁶ Satish Garla, Albert Hopping, Rick Monaco, & Sarah Rittman, *What Do Your Consumer Habits Say About Your Health? Using Third-Party Data to Predict Individual Health Risk and Costs. Proceedings*, SAS Global Forum 2013. <<http://support.sas.com/resources/papers/proceedings13/170-2013.pdf>>.

- Commercial data brokers
- MultiChannel direct response
- Survey data, especially online
- Catalog/phone order/Online order
- Warranty card registrations
- Internet sweepstakes
- Kiosks
- Social media interactions
- Loyalty card data (retailers)
- Public record information
- Web site interactions, including specialty or knowledge-based web sites
- Lifestyle information: Fitness, health, wellness centers, etc.
- Non-profit organizations' member or donor lists
- Subscriptions (online or offline content)

(Part II of the report contains a more complete list.)

Traditionally, much of this data came from data brokers or mailing list sellers. That is still the case, but now many new data streams are now available. So-called big data (large data sets) is one source. Other new data, particularly mobile and social data streams, comes via application programming interfaces (API).

Data sets that used to be too large for all but the largest of companies to handle computationally can now be replicated and massaged by smaller firms and dedicated analytics teams within companies. Small analytics companies now compete with large data brokers to offer predictive analytics as well as data. One company states they use 300 billion data attributes in compiling their predictive scores, compiled from 8,000 data files.¹⁷ This is no longer an extraordinary feat, it is competitive and to be expected in a world with large data flows.

Analytics tools will continue to come down in price, just as consumer data has become a commodity item. Widespread and inexpensive data and analytics have the potential to allow broader use of predictive analytics. Consumer scores may proliferate, especially in the absence of any need for accuracy, fairness, or transparency. Consumer scoring may expand just because it is cheap and fashionable. Merchants themselves may have little ability to judge the accuracy of consumer scoring.

In short, given abundant data and more data tools, factors used to create consumer scores could continue to increase. With each new unverified factor comes the risk of extra errors or unfairness due to sensitive or prejudicial or irrelevant factors.

¹⁷ See Part III, Churn and Fraud scores for further discussion.

Scoring Methods and Models are Opaque

To create a consumer score, the score modeler feeds raw information (factors about consumers) into an algorithm designed to trawl through reams of data to detect consumer behavior patterns and to eventually sift consumers into a ranking by their scores. Each score generally has a name and predictive or descriptive function.

Credit scores are the best-known example of this. With credit scores, information culled from a consumer's credit bureau file becomes the raw input into a formal credit scoring model. **Credit** scores are built on **credit file** data. There is a nexus between the score and the data. The data is intrinsic to the score. The Fair Credit Reporting Act lays out, in concert with the Equal Credit Opportunity Act, a variety of responsibilities and restrictions in the uses of credit report data to use in credit scores. It is a balanced approach, and the Fair Credit Reporting Act remains a strong privacy law that enables Fair Information Practices for consumers.

Today, though, scoring models are easily built from data that is **extrinsic** to the final score. No nexus may exist between the **input** to a score and the **output**. In the financial scoring area, companies can now build financial scores from social media, demographic, geographic, retail purchase history, and other non-traditional information that may not be included in the formal credit file. In the health arena, analysts can now build health risk scores from mere wisps of demographic data, without any actual patient records.

In this is a new world of scoring, where analysts use factors extrinsic to the purpose of the score to build scores, that a person has red hair can be used as a factor. And the more factors, the better. Instead of using 30 factors, why not 3,000?

The use of credit information for pricing insurance risk is an example of this.¹⁸ Statisticians and actuaries predict the cost of providing car or homeowners insurance using selections of credit report factors about a driver or homeowner. These insurance scores reportedly have a predictive capability. Yet there is no overt reason why credit worthiness correlates with the risk involved with driving safety. There is much controversy about the use of a statistical correlation that does not appear to be causal. Some states restrict the use of credit information for insurance, but the practice remains common.

Many of the consumer scores discussed in Part III are new classes of scores. When scores have hundreds and thousands of factors, it stands to reason that a causal link becomes much more tenuous. The more factors, the less casual the link may be. Risks associated with models are discussed in Part II of the report.

¹⁸ See generally, Federal Trade Commission, *How Credit Scores Affect the Price of Credit and Insurance*, <<http://www.consumer.ftc.gov/articles/0152-how-credit-scores-affect-price-credit-and-insurance>>.

Examples and Numbers of Consumer Scores

Consumer scoring is growing. In 2007, the research for this report uncovered less than 25 scores. In 2014, the research uncovered hundreds of scores, with the strong likelihood that thousands of custom scores exist beyond our ability to confirm them.

Here are some examples of consumer scores:

Consumer profitability scores predict, identify, and target marketing prospects in households likely to be profitable and pay debt.

The Job Security Score claims to predict future income and capacity to pay.

Churn scores seek to predict when a customer will move his or her business or account to another merchant (e.g., bank, cell phone, cable TV, etc.)

The Affordable Care Act (ACA) health risk score creates a relative measure of predicted health care costs for a particular enrollee. In effect, it is a proxy score for how sick a person is.

The Medication Adherence Score predicts if you are likely to take your medication according to your doctor's orders.

Brand Name Medicine Propensity Score – will you be purchasing generics or brand name medications?

Fraud Scores indicate that a consumer may be masquerading as another, or that some other mischief is afoot. These scores are used everywhere from the Post Office at point of sale to retailers at point of sale to behind-the-scenes credit card transactions. This is a very widely used score, and a number of companies compete in the fraud score area.

Part III of the report discusses these and other specific scores in detail.

Uses of Consumer Scores, Regulation, and Modern Eligibility

After a consumer is scored, ranked, described, or classified, companies, governments, private enterprises, health care entities, and others including law enforcement, can then use the resulting score to make decisions about an individual or group.¹⁹ This is why scores impact consumers every day.

¹⁹ We note in passing that consumer scores are typically created by analytics companies or professionals, and then the company or individual can either sell the score or sell their abilities to create custom scores for third parties. This is neither good nor bad, it is simply the basic business model, and it is quite old, stretching back many decades now. See discussion in Part IV.

Scores are gaining footholds as part of routine business processes for an expanding number of purposes for everything from marketing to assessing a person's identity to predicting a person's likelihood to commit fraud, and more.²⁰ The consumer score acts as a form of predictive evaluation to measure, predict, and generally facilitate making a decision about things such as an individual's:

- Credit worthiness,
- Popularity,
- Reputation,
- Wealth,
- Propensity to purchase something or default on a loan,
- Measure health,
- Measure/predict likelihood to commit fraud,
- Measure/predict identity
- Measure/predict energy consumption
- Job success probability
- Etc.

In the traditional credit score realm, some have argued that scoring is a foundational activity in the credit market, as well as a wholly positive factor. Others have said that the sub-prime meltdown of the late 2000's was fueled by overreliance on scoring products.²¹ Some of the reasons for using credit scores have the potential to be helpful directly to consumers. Better and faster credit decisions help consumers, for example.

In new consumer scoring, some have argued that the scores are mainly just for marketing and are largely beneficial.²² There can be potential benefits for consumers. For example, consumer risk scores that prevent fraud are helpful up to a point. But any potential benefits are real only if the scoring models are correct and non-discriminatory, the data is timely, and the scores are something that consumers want. Credit score regulation provides transparency and imposes some limits on use and construction. That offers some assurance to consumers. But when other consumer scores enter the marketplace without transparency or the limits that apply to credit scoring,²³ consumer benefits are much more uncertain, and unfairness is more likely.

²⁰ A trend in the data business is that consumer data itself has become a commodity due to the ease with which much consumer data can be acquired. Predictive analytics are becoming the key drivers of the data business. Instead of just lists of consumer information, a predictive score is a "value add" to data offerings.

²¹ Federal Reserve Board Chairman Alan Greenspan, remarks at the annual convention of the American bankers association, (October 7, 2002), page 4. The extent to which the same credit scoring technologies touted by Chairman Greenspan may have been responsible for the mortgage meltdown and financial crisis that started in 2008 is beyond the scope of this report.

²² FTC – Alternate Scoring Workshop, March 19, 2014. <<http://www.ftc.gov/news-events/press-releases/2014/03/ftc-announces-agenda-panelists-alternative-scoring-seminar>>.

²³ We do not mean to suggest that consumer scores have flaws and lack a full range of consumer protections, only that some limits and rights exist.

There is continuum of concern regarding consumer scores. Some scores are used for straightforward marketing purposes. These scores may be of less concern (however the fairness of factors and secrecy and validity are still a concern). Of greater concern are the consumer scores that are used for what we call “modern eligibility.” This includes identity verification and fraud assessment scores, as well as credit decisioning scores and scores that are used to predict job success or decide between job applicants. These scores are especially worrisome because errors in these scores could lead to significant deleterious consumer impacts.

Whether a consumer receives a coupon for a free soda is not a big deal. In comparison, whether a consumer can complete a transaction is of significant consequence. Any score used for eligibility – like being approved for credit or a job -- becomes important. The most casual social scores meant just to measure social reach have on occasion been used as a criterion for judging applicant hiring qualifications, so all scores need to be explored and assessed.

Some scores –for example, aggregate credit scores not subject to the FCRA – can determine a neighborhood’s general credit score or range. Opportunities for individuals living in that neighborhood will be affected in ways that they cannot anticipate and in ways that bear no relationship to their personal situation. Forms of redlining – the practice of turning someone down for a loan or insurance because they live in an area deemed to be high risk – is a threat in these situations.

By all appearances, consumer scoring has sped beyond the old constraints that were imagined in a largely analog era, and real consumer harms can be the result.

Deja Vu: Why the History of the Credit Score is Important

History is repeating itself with consumer scoring. Before secret predictive consumer score issues, there were secret credit score issues. Credit scores had many of the same problems: secrecy, unfairness, inaccuracy, and opacity. Part IV of the report contains a detailed history of the credit score, including how that score became public and how consumers got important rights regarding credit scores and reports.

In brief here, credit scores were unknown to most consumers through the 50s, 60s, 70s, and 80s. Trickle of a score that was not disclosed to consumers but that could be used to deny a person credit began to leak out slowly to some policymakers, particularly around the time ECOA passed. In May 1990, the Federal Trade Commission failed to protect consumers when it wrote commentary indicating that risk scores (credit scores) did not have to be made available to consumers. But when scoring began to be used for mortgage lending in the mid 90s,²⁴ many consumers finally began hearing about a “credit score,”

²⁴ In 1995 Freddie Mac and Fannie Mae endorsed the use of credit scores as part of the mortgage underwriting process. This had a substantial impact on the use of credit scores in the mortgage loan industry. See for example Kenneth Harney, *The Nation's Housing Lenders might rely more on credit scores*, *The Patriot Ledger*, July 21 1995.

many of them for the first time, and mostly when they were being turned down for a loan.²⁵ A slow roar over the secrecy and opacity of the credit score began to build.

By the late 90s, the secrecy of credit scores, the underlying methodology or factors that went into the score, and the scoring range became a full-blown policy issue. Beginning in 2000, a rapid-fire series of events – particularly the passage of legislation in California that required disclosure of credit scores to consumers – eventually ended credit score secrecy. Now, credit scores **must** be disclosed to consumers, and the context, range, and key factors are now known.²⁶ This is an example of how brave State privacy legislation serves as a model for state and federal policy makers. In this case, the US “laboratory of democracy” took state legislation and turned it into a federal rule that protects consumers everywhere.

Credit scores are no longer secret anywhere in the United States, and this was and still is the right policy decision. Why are other scores used for important decisions about consumers still secret? Why do score factors and numeric ranges remain secret, when the risk of the data comprising the score of a factor used in modern eligibility practices such as identity verification or fraud identification is very high?

Consumer scores stand today where credit scores stood in the 1950s: in the shadows. While there are some happy exceptions to this, such as most social scores and a few other consumer scores, most consumer scores are not available for consumers to see. As a result, consumers have little to no ability to learn when their lives are affected in a major or minor way by a consumer score that they never heard about. Credit scores are not perfect and still present some issues, but we have learned much from the credit score.²⁷ What we have learned most of all is that there should be no secret consumer scores and no secret scoring factors. If a score is being used in any meaningful way in a consumer’s life, he or she needs to know about it and have some choices regarding that score.

²⁵ See for example, comments of Peter L. McCorkell, Senior Counsel to Wells Fargo, to the Federal Trade Commission, August 16, 2004 in response to FACT Act Scores Study.

²⁶ As of December 2004, the Fair Credit Reporting Act as modified by the Fair and Accurate Credit Transactions Act, or FACTA, ended score secrecy formally, and required consumer reporting agencies to provide consumers with more extensive credit score information, upon request. Also made available to the public was the context of the score (its numeric range), the date the score was created, some of the key factors that adversely affected the score, and some other items.

²⁷Historically, some known consumer issues with the credit score include the following:

- Credit scores reflect inaccuracies in the credit reports they are based on, and credit reports have repeatedly been found to contain errors.
- Victims of ID theft can experience changed credit scores.
- Consumers who experience major life events such as medical events or divorce can pay a long price in the scoring world.
- The FTC has brought cases around “mission creep” in the use of credit score outside of its regulated uses. (Credit scores may only be used for firm offers of credit or insurance, not for general marketing use.)

Summary of Findings and Recommendations

Key Findings:

Consumer scores are expanding in type, number, and use because of the growth of predictive analytics and the ready access to hundreds and thousands of factors as raw material. Just as credit scores were secret for decades until state and federal legislation mandated that consumers could see their credit scores, today consumer scores are largely secret.

While new scores multiply, consumers remain in the dark about many of their consumer scores and about the information included in scores they typically don't have the rights to see, correct, or opt out of. A primary concern is how these scores affect individuals and meaningful opportunities available to them. Another area of concern is the factors used in new consumer scores, which may include readily commercially available information about race, ethnicity, religion, gender, marital status, and consumer-reported health information. This report's other key findings are:

- Unregulated consumer scores – as well as regulated credit scores – are both abundant and increasing in use today.
- The information used in consumer scores comes from a large variety of sources. Some scores use thousands of factors or consumer attributes.
- Many consumer scores, the ranges of the scores, and the factors used in them are secret.
- A consumer score may, without any public notice, rely on an underlying factor or attribute that has discriminatory implications (e.g., race or gender) or that most consumers consider sensitive (e.g., health or financial).
- Consumer scores in use today affect a consumer's marketplace opportunities. Some of these opportunities are major (e.g., financial, employment, health), some are minor (e.g., receiving a coupon, spam, or junk mail), and many are in between. Consumers are adversely affected by scores that are kept secret, and consumers are adversely affected when they do not have rights to correct scores.
- Consumer scores are found in a wide array of "offline" arenas, including businesses, health care providers, financial institutions, law enforcement, retail stores, federal and state government, and many other locations. Some of the more social consumer scores may be online, but mostly consumer scores are not solely focused on just online activities.
- Consumers usually have no way to know what the scores predict or how the scores are used.

- Consumers typically have no notice or knowledge about the data sources used in scores predicting their behavior or characterizing them. Consumers typically have no rights over the data about themselves, and consumers usually have little to no ability to control use of the data.
- Consumers typically do not have the right to opt out of being the subject of a consumer score or to prevent use of a consumer score.
- Except where the Fair Credit Reporting Act applies to a consumer score, most consumer scores are not subject to any regulation for privacy, fairness, or due process. A lack of transparency makes it difficult or impossible to determine if creation or use of the scores violates a law that prohibits discrimination.
- Consumers who are victims of identity theft can have their credit or consumer scores affected thereby and may have little recourse even though errors may have major consequences for their ability to function in the economic marketplace can be major. Other consumers can also have their lives affected by the use of consumer scores to determine eligibility for important opportunities in the marketplace. Some consequences may be less significant.
- Consumers have remedies under state and federal law with respect to correcting and seeing their credit reports, but not necessarily with respect to the many records that contribute to consumer scores. Secret consumer scores do not provide consumers with correction rights of underlying information.

Key Recommendations:

Consumer scoring is not inherently evil. When properly used, consumer scoring offers benefits to users of the scores and, in some cases, to consumers as well. Some uses are neutral with respect to consumers. Consumer scores can also be used in ways that are unfair or discriminatory. The goal of these recommendations is to protect the benefits of consumer scoring, guarantee consumer rights, and prevent consumer harms.

- No secret consumer scores. No secret factors in consumer scores. Anyone who develops or uses a consumer score must make the score name, its purpose, its scale, and the interpretation of the meaning of the scale public. All factors used in a consumer score must also be public, along with the nature and source of all information used in the score.
- The creator of a consumer score should state the purpose, composition, and uses of a consumer in a public way that makes the creator subject to Section 5 of the Federal Trade Commission Act. Section 5 prohibits unfair or deceptive trade practices, and the FTC can take legal action against those who engage in unfair or deceptive activities.

- Any consumer who is the subject of a consumer score should have the right to see his or her score and to ask for a correction of the score and of the information used in the score.
- There are so many consumer scores in existence that consumers should have access to their scores at no cost in the same way that the law mandates credit reports be available at no cost, as mandated by Congress. Otherwise, if a consumer had to pay only one dollar for each meaningful score, a family could easily spend hundreds or thousands of dollars to see the scores of all family members.
- Those who create or use consumer scores must be able to show that the scores are not and cannot be used in a way that supports invidious discrimination prohibited by law.
- Those who create or use scores may only use information collected by fair and lawful means. Information used in consumer scores must be appropriately accurate, complete, and timely for the purpose.
- Anyone using a consumer score in a way that adversely affects an individual's employment, credit, insurance, or any significant marketplace opportunity must affirmatively inform the individual about the score, how it is used, how to learn more about the score, and how to exercise any rights that the individual has.
- A consumer score creator has a legitimate interest in the confidentiality of some aspects of its methodology. However, that interest does not outweigh requirements to comply with legal standards or with the need to protect consumer privacy and due process interests. All relevant interests must be balanced in ways that are fair to users and subjects of consumer scoring.
- The FTC should continue to examine consumer scores and most especially should collect and make public more facts about consumer scoring. The FTC should establish (or require the scoring industry to establish) a mandatory public registry of consumer scores because secret consumer scoring is inherently an unfair and deceptive trade practice that harms consumers.
- The FTC should investigate the use of health information in consumer scoring and issue a report with appropriate legislative recommendations.
- The FTC should investigate the use of statistical scoring methods and expand public debate on the proprietary and legality of these methods as applied to consumers.
- The Consumer Financial Protection Bureau should examine use of consumer scoring for any eligibility (including identity verification and authentication) purpose or any financial purpose. CFPB should cast a particular eye on risk

scoring that evades or appears to evade the restrictions of the FCRA and on the use and misuse of fraud scores. If existing lines allow unfair or discriminatory scoring without effective consumer rights, the CFPB should change the FCRA regulations or propose new legislation.

- The CFPB should investigate the selling of consumer scores to consumers and determine if the scores sold are in actual use, if the representations to consumers are accurate, and if the sales should be regulated so that consumers do not spend money buying worthless scores or scores that they have no opportunity to change in a timely or meaningful way.
- Because good predictions require good data, the CFPB and FTC should examine the quality of data factors used in scores developed for financial decisioning and other decisioning, including fraud and identity scores. In particular, the use of observational social media data as factors in decisioning or predictive products should be specifically examined.
- The use of consumer scores by any level of government, and especially by any agency using scores for a law enforcement purpose, should only occur after complete public disclosure, appropriate hearings, and robust public debate. A government does not have a commercial interest in scoring methodology, and it cannot use any consumer score that is not fully transparent or that does not include a full range of Fair Information Practices. Government should not use any commercial consumer score that is not fully transparent and that does not provide consumers with a full range of Fair Information Practices.
- Victims of identity theft may be at particular risk for harm because of inaccurate consumer scores. This is a deeply under-researched area. The FTC should study this aspect of consumer scoring and try to identify others who may be victimized by inaccurate consumer scoring.

Advice For Consumers:

- Consumers can take several steps to help reduce some but not all of the data flows regarding scores. If a consumer opts out of pre-screened offers of credit and insurance (Opt Out Prescreen)²⁸, this can help reduce the overall volume of credit information circulating about them. Opting out of affiliate information sharing (as allowed) under the Graham Leach Bliley opt out²⁹ can also help reduce information flows. For scores regulated under the Fair Credit Reporting Act, consumers can get one free credit report each year from the major bureaus at

²⁸ <<https://www.optoutprescreen.com>>. *World Privacy Forum Top Ten Opt Outs*, <<http://www.worldprivacyforum.org/2013/08/consumer-tips-top-ten-opt-outs/>>.

²⁹ <<http://www.fdic.gov/consumers/privacy/privacychoices/index.html#yourright>> and *World Privacy Forum Top Ten Opt Outs* <<http://www.worldprivacyforum.org/2013/08/consumer-tips-top-ten-opt-outs/>>.

www.annualcreditreport.com. For more information about consumer rights under the FCRA, ECOA, and other laws, see FTC's consumer resources³⁰ and CFPB's consumer resources.³¹

Part II. Consumer Scores: What Goes Into Them, and How They Are Made

Part II discusses how consumer scores are constructed in more depth. This part includes three main segments: 1) a technical discussion of how scores are made (which may be skipped for those not interested in the technical details of predictive analytics); 2) a list of data sources used in scores, and policy questions; and a more detailed definition of consumer scores.

Defining Consumer Scores More Deeply

We repeat here the definition we introduced earlier in this report:

Consumer scores are scores that describe an individual or sometimes a group of individuals (like a household), and have a demonstrated ability to predict one or more consumer behaviors or outcomes. Consumer scores use information about consumer characteristics, behaviors, and other attributes in different amounts and combinations in statistical models that produce a numeric score, a range of scores, or a yes/no. Consumer scores are used to rate, rank, segment, and make decisions and predictions about individual consumers and groups of consumers. Those decisions can range from innocuous to important.

Generally, there are three elements in scoring, scores, factors, and models (or algorithms). The score is a metric, often but not always a number (e.g., categorical), that measures some quality of an individual (or group) or a transaction. A score is often used to determine a course of action regarding an individual or a transaction. *Consumer* scores are a class of scores used to make a determination about a consumer or a transaction related to or affecting the consumer either directly or indirectly.³² Scores can be

³⁰ <<http://www.consumer.ftc.gov/topics/credit-and-loans>> and <<http://www.consumer.ftc.gov/articles/0347-your-equal-credit-opportunity-rights>>.

³¹ <<http://www.consumerfinance.gov/fair-lending/>> and <http://www.consumerfinance.gov/askcfpb/search?selected_facets=category_exact:credit-reporting>.

³² In the scoring literature, consumer score as a term does not appear frequently, but its occurrences generally concur with the definition the World Privacy Forum has used in this report. See for example Dan Meder, *Blended scores are better scores*, 109 Business Credit 48-49 (2007).

generated by a variety of means, including fully automated algorithms³³ and hybrid models.

Factors can be thought of as pieces of information that describe or relate in some way to the consumer, or to consumer behavior that the creator of the model feels or has determined is important to increase the predictive power of the model. Predictive mathematical models that generate scores can use many inputs, including but not limited to heuristics, demographics, transaction behavior, user history, comparative and/or existing profiles and results from other scores. Other factors can be payment history, number of late payments, and length of credit are all factors about an individual's transactions related to credit that may be used in assigning a credit score. Age, income, race, geographic location, education level, and patterns of behavior (for example, how many times a person has returned merchandise to a particular store, or how many times an individual has bounced a check) can be relevant data, depending on the goals of the score.

The model or algorithm takes the personal factors associated with an individual or a class that the individual belongs to (e.g., household or neighborhood), measures the factors against the model, and assigns a score to that individual. A scoring model can be simple or complex. It can use vast quantities of personal data, or just a little. A big data approach to consumer scoring typically requires analysis of large amounts of data to forecast future behavior, outcomes, or qualities.³⁴

Scores, and the models and factors that produce the scores, can be controversial. A score is only as predictive or as fair as the score model and the factors used in that model. One score can use a factor that cannot or is not used in another score. In some situations, for example, home mortgages, score models cannot use factors that would discriminate, such as age or race. But other score models, such as auto insurance, can use some factors prohibited in home insurance. Many consumer scores are completely unrestricted in the choice of factors.

Another controversy that comes up frequently in generating predictive models is the problem of over-fitting. Over-fitting arises when an algorithm is trained to perform very well on an existing set of data, but has been tailored so well to that data set that it can behave erratically or incorrectly outside of the specific scenario it has trained for.

³³ Algorithms are procedures for solving a problem, often a mathematical problem, in a series of finite steps. In scoring models, algorithms also refer generally to the processes by which data are analyzed within a predictive model in order to generate a score. See [Fair Isaac Corporation Overview, Decisions Made Simple](http://www.fairisaac.com/NR/rdonlyres/A609962F-371C-4FCA-BB2A-85DE8FD936F5/0/FairIsaacCorpOverview06.pdf), Glossary of EDM, <<http://www.fairisaac.com/NR/rdonlyres/A609962F-371C-4FCA-BB2A-85DE8FD936F5/0/FairIsaacCorpOverview06.pdf>>.

³⁴ This definition was culled from a wide variety of literature sources and from author interviews with experts in the scoring field. See bibliography for source list, see especially D.J. Hand & W.E. Henley, *Statistical Classification Methods in Consumer Credit Scoring: A Review*, 160 *Journal of the Royal Statistical Society*, 523-541 (1997). See also Allen N. Berger, W. Scott Frame, *Small Business Credit Scoring and Credit Availability*, 45 *Journal of Small Business Management* (2007). See also [Fair Isaac Corporation Overview, Decisions Made Simple](http://www.fairisaac.com/NR/rdonlyres/A609962F-371C-4FCA-BB2A-85DE8FD936F5/0/FairIsaacCorpOverview06.pdf), Glossary of EDM, <<http://www.fairisaac.com/NR/rdonlyres/A609962F-371C-4FCA-BB2A-85DE8FD936F5/0/FairIsaacCorpOverview06.pdf>>.

To understand overfitting, consider an algorithm training a robot to go to the kitchen and bring its master a drink. If trained specifically in its master's house, the robot may in time learn how to get the master's drink from the kitchen without fail. However, it does so by creating a set of rules specific to the master's house. For example, go past the pink chair 10 feet, turn right at the blue doorway two feet from the refrigerator, and then lift the arm exactly 20 inches to open the door and retrieve the diet soda on the third shelf. The problem with this scenario is that it doesn't generalize well across all housing, all masters, and all drinks.

What the developer of the algorithm wants is a model that works in all circumstances. That is a much harder task, and use of shortcuts risks overfitting. Overfitting is a constant danger for people who create algorithms, and models must be constantly tested in a variety of settings to ensure that they are not overfitting a specific scenario. Consequently, scores that have not been broadly tested may be inaccurate for some uses.

All consumer scores use some kind of consumer information – characteristics, behaviors, transactions, and/or attributes³⁵ – that describe an individual or sometimes a group of individuals and have demonstrative ability to predict some kind of behavior. Huge quantities of data may be collected and organized for this purpose. Each type of consumer score uses different factors, or consumer information, in different amounts and combinations.

Some scoring models, especially in the testing phases, can use as many as 1,000 variables or more to create a single score. Some use only one or two factors. But most use many factors. In the past few years, scoring has matured rapidly as predictive analytics teams and expertise have become part and parcel of how companies seek to monetize and understand their customer base, among other activities.

In this discussion, it is important to state what a consumer score is not.³⁶ For the purposes of this report, a consumer score does not measure a consumer's skill or abilities. For example, an SAT score is not a consumer score in and of itself. Similarly, a health score that uses clinical health factors in a clinical setting solely for health diagnosis or treatment by a health care provider is not a consumer score.³⁷ However, if an SAT score or a health score is used for a different purpose as part of a predictive consumer score, then those scores will become factors of a consumer score.

³⁵ Characteristics are, for example, the questions asked on the credit application. Characteristics can also be performance categories of the credit bureau report. Attributes are, for example, the answers given to questions on the application, or entries in the credit bureau report. Education is a characteristic, college degree or highest level of education achieved is the attribute.

³⁶ Some media reports have called alternative non-credit scores "e-scores." This term is actually a proprietary name for a product from a company called eBureau. <<http://www.ebureau.com>>. Even if the term were available for general use, the term e-score is too narrow and limiting for the broad range of health, energy, social and other consumer scores now in use.

³⁷ We assume without investigation that scores used for health care diagnostic or treatment purposes are fully transparent.

Also, if a score uses health factors for decisions outside of a strictly clinical setting, or if a score predicts a health outcome using extrinsic, non-clinical factors, then those scores are consumer scores. Health scores are discussed more in the discussion of individual consumer scores in Part III.

An appendix to this report includes a score taxonomy that describes consumer scores and score types, along with a decision tree to assist in determining when a score is a consumer score or not.

Making a Consumer Score, Step 1: What Consumer Information Gets Put Into a Consumer Score?

The underlying factors that go into a consumer score are important indicators of the fairness, accuracy, and non-discrimination of the score. If the factors selected to create a score are inaccurate, unfair, or discriminatory, then the score itself will be susceptible to the same biases.

Traditional Score Ingredients: Credit Scores

Much is known about what goes into credit scores and about credit score models. As with all scores, a key to a good credit score is the quality (e.g., accuracy, currency, and completeness) of the factors used in the scoring model. This is scoring 101.

David T. Kresge, formerly senior vice president of analytic services at Dun & Bradstreet noted the “data hunger” of scoring models in congressional testimony:

“One of the keys to the implementation of a knowledge-based decision system is to incorporate and make effective use of the widest possible range of information. The data should be gathered from all available sources and it should be as wide-ranging as possible. Experience clearly demonstrates that good credit decisions cannot be based on just a small number of factors.”³⁸

Examples of some factors that go into consumer scores in the United Kingdom include:

- Time at present address
- Home status (owner, tenant, other)
- Credit card information
- Type of bank account
- Telephone (yes, no)
- Age

³⁸ Prepared testimony of David T. Kresge before the House Committee on Small Business, Subcommittee on Government Programs and Oversight (July 17, 1998).

- County Court judgments (number)
- Purpose of loan
- Type of occupation (coded)
- Marital status (married, divorced, single, widow, other)
- Time with bank (years)
- Time with employer (years)³⁹

In the U.S., commonly used predictive variables for traditional financial scoring include:

- Payment history
- Public record and collection items
- Delinquencies
- Prior credit performance
- Outstanding debts
- Relationship between total balances of credit and total limit
- Age of oldest trade line
- Pursuit of new credit (applications to obtain additional credit)
- Time at present address
- Time with current employer
- Type of residence
- Occupation.⁴⁰

From these two examples, we see that characteristics included in a financial score model vary from country to country. They may vary from state to state, depending whether laws restrict the use of some characteristics or variables. The data available in different countries may differ, and that may explain in part the construction of the model. It is not unusual for *missing information* to be an actual characteristic and included as a factor in the scoring model. This is a much-debated area of scoring.⁴¹

Until 2000, the factors that went into credit scores were not public.⁴² However, these factors are now known. Fair Isaac reveals that for its FICO score,

- 35% is based on borrower's history payment history
- 30% is based on how much a borrower had drawn on available credit (amounts owed)

³⁹ These factors are culled from Table 1, Characteristics typical of certain credit scoring domains, D.J. Hand, *Statistical Classification Methods in Consumer Credit Scoring: A Review*, 160 *Journal of the Royal Statistical Society*, 527 (1997).

⁴⁰ These factors are culled from two sources. See Margaret Howard, *Shifting risk and fixing blame: The vexing problem of credit card obligations in bankruptcy*, *The American Bankruptcy Law Journal*, Winter 2001, 76 *Am. Bankr. L.J.* 63. See also *Scoring and Modeling*, FDIC – Risk Management Examination Manual for Credit Card Activities, Chapter VIII. Division of Supervision and Consumer Protection, March 2007. <http://www.fdic.gov/regulations/examinations/credit_card/ch8.pdf>.

⁴¹ See, e.g., Hand, DJ & Henley WE, *Can reject inference ever work?* 5 *IMA J Maths Appl Bus Ind* 5-55 (1993).

⁴² See the heading in this report “A Brief History of Consumer Scores” for a discussion of how credit score factors became public.

- 15 % is based on the length of the credit history
- 10% is based on the types of credit used
- 10 % is based on new credit⁴³

We know the factors and their weights, but we do not know FICO turns a particular consumers payment history into a number that becomes a score.

Modern Score-Making Ingredients: Raw Consumer Data in the Digital Age

The carefully selected scoring factors and the much-debated weights for regulated credit scores as discussed above are like a landscaped garden in a well-tended public park compared to the untamed jungles of the data factors available and used in the new consumer scores. The new kinds of consumer scores use a much wider array of data sources, to the point that the new data sources make traditional credit scores look under-sourced by comparison. Whether 500 factors result in a better algorithm than 5 factors is unknown, and the answer may vary from score to score. More may be better sometimes, but not all the time.

Data for consumer scores can come from many sources, including data broker lists, retail purchases, social scores, census tract data, purchasing patterns, health conditions, ethnicity, book purchasing patterns, exercise patterns, and many other factors. Data used may be individual to a consumer or modeled (e.g., all consumers in a census tract). As described above, the effectiveness of a model depends on its ability to predict accurately from a variety of real world datasets and designated factors. Understanding what a model is trying to predict, what data is used for testing, and how the elements mesh to achieve a result are important to assessing the value, impact, and potential pitfalls of the scoring model. In some cases, it may be that a model is equally effective with less information, negating the need for collection and storage of vast quantities of information and data that could have privacy implications. More data may not be better or necessary.

Below is a list of the most common elements of consumer data available and in circulation today. Most consumers would be stunned to learn the number of data elements available in the commercial marketplace. Not every consumer score and not every data broker file for sale includes each item on this list. Different scores use combinations of different elements and plug those into differing score algorithms and models.

This list includes independent data sets with both structured and unstructured data. This list is sourced in part from 2013 Government Accountability Report on information resellers. Other information came from a WPF review and analysis of data broker data cards viewed through NextMark ⁴⁴over the course of a year (primarily 2013), and also from WPF review and analysis of reliable data broker web sites that list data sources. For

⁴³ MyFico, "Credit Education," <<http://www.myfico.com/CreditEducation/>>.

⁴⁴ Nextmark List Finder <<http://lists.nextmark.com/>>.

example, the Acxiom *About the Data* portal⁴⁵ lists many categories of information collected and used for consumer marketing.

The data sets available for purchase today listed here – along with others we did not identify – can create multiple layers of predictive analysis of how consumer behavior, finance, demographics, geography, and the other factors listed here interact. That does not necessarily mean that the results are better.

Consumer Data Available for Purchase and Use in Analytics

The range of consumer data available for use in data analytics is broad and deep. The categories listed here is not exhaustive, but it offers an idea of the range of consumer information that goes into consumer scores.

Demographic Information:

- Age
- Age range
- Date of birth
- Education
- Exact date of birth
- Gender
- Marital status
- Home ownership
- Own or rent
- Estimated income
- Exact income
- Ethnicity
- Presence of children
- Number of children
- Age range of children
- Age of children
- Gender of children
- Language preference
- Religion
- Occupation - category of occupation
- Examples: Beauty (cosmetologists, barbers, manicurists) civil servants, clergy, clerical/office workers, doctors/physicians/surgeons, executives/administrators, farming/agriculture, health services, middle management, nurses, professional/technical, retail service, retired, sales, marketing, self-employed, skilled/trade/machine operator/laborer, teacher/educator.
- Occupation - title of occupation
- Military history

⁴⁵ Acxiom About the Data Portal <<https://aboutthedata.com>>.

- Veteran in household
- Voter party
- Professional certificates (teacher, etc.)
- Education level reached or median education

Contact Information:

- Full name
- Email address
- City
- State
- ZIP
- ZIP + 4
- Home Address
- Land-line phone
- Social IDs / social media handles and aliases
- Mobile phone number
- Carrier
- Device type
- Email address

Vehicles:

- Vehicle make, model and year
- VIN
- Estimated vehicle value
- Vehicle lifestyle indicator
- Model and brand affinity
- Used vehicle preference indicator

Lifestyle, Interests and Activities data (including medical):

- Antiques
- Apparel (women, men & child)
- Art
- Average direct mail purchase amounts
- Museums
- Audio books
- Auto parts, auto accessories
- Beauty and cosmetics
- Bible purchaser
- Bird owner
- Books
- Book purchases - plus types. (Mystery, romance, religious, etc.)

- Book clubs
- Career
- Career improvement
- Cat owner
- Charitable giving indicators:
- Charitable donor by type of donation (religious, health, social justice)
- Charitable donor by ethnicity or religion (Jewish donors, Christian donors, Hispanic donors)
- Charitable donor by financial status (wealthy donors)
- Children or teen interests
- Fashion and clothing (Multiple: sports, high fashion, shoes, accessories, etc.)
- Collectibles
- Collector
- Christian families
- Computer games
- Computers
- Consumer electronics (Many categories, including electronic fitness devices)
- Dieting and weight loss
- Telecommunications and mobile
- Dog owner
- Investing
- DVD purchasers
- Electronics - home, computing, office, other
- Empty nester
- Expectant parents
- Frequent mail order buyer
- Frequency of purchase indicator
- Getting married
- Getting divorced
- Gun ownership
- Health and beauty
- Health and medical: for example, Allergies, Alzheimer's disease, angina, arthritis/rheumatism, asthma, back pain, cancer, clinical depression, diabetes, emphysema, erectile dysfunction, epilepsy, frequent heartburn, gum problems, hearing difficulty, high blood pressure, high cholesterol, irritable bowel syndrome, lactose intolerant, ulcer, menopause, migraines/frequent headaches, multiple sclerosis, osteoporosis, Parkinson's disease, prostate problems, psoriasis/eczema, sinusitis/sinuses.
- High-end appliances
- Home improvement
- Household consumer expenditures — many categories.
- Jewelry
- Magazine subscriptions
- Mail order buyer
- Mobile location data (some analytics companies)
- Movies - attendance / collector

- Musical instruments
- Music
- New mover
- New parent
- Online and continuing education
- Online purchasing - many categories
- Parenting
- Pets - other
- Plus size clothing purchase
- Political affiliation
- Recent home buyer
- Recent mortgage borrower
- Retail purchasing - many categories.
- Science-related
- Sexual orientation
- Social media sites likely to be used by an individual or household, heavy or light users
- Spa
- Sports interests: (large category, these are examples)
- Birdwatching
- Equestrian
- Exercise and fitness
- Gardening
- Golf
- Fishing
- Outdoor interests - hiking, camping, climbing
- Swimming, diving, snorkeling
- Spectator Sports
- Stamps/coins
- Yoga
- Television, Cable, Satellite/Dish
- Travel: Vacations, domestic and/or international
- Purchase of international hotel or air flights
- Frequent flyer
- Types of purchases indicator
- Veteran in household
- Vitamins
- Volunteering

Financial and Economic – Property and Assets data:

- Estimated income
- Estimated household income
- Home value
- Length of residence
- Payment data: 30, 60, 90-day mortgage lates

- Purchase date
- Purchase price
- Purchase amount
- Most recent interest rate type
- Most recent loan type code
- Sales transaction code
- Most recent lender code
- Purchase lender code
- Most recent lender name
- Purchase lender name
- Fuel source
- Loan to value
- Purchase interest rate type
- Most recent interest rate
- Purchase interest rate
- Pool or spa
- Home - year built
- Air conditioning
- Boat ownership
- Plane ownership
- Motorcycle ownership
- Commercial assets or business ownership

Financial and Credit data:

- Bankruptcy
- Beacon score
- Credit score - actual
- Certificates of deposit/ money market funds
- Estimated household income ranges
- Income producing assets indicator
- Estimated net worth ranges
- IRAs
- Life insurance
- Low-end credit scores
- Mutual funds/annuities
- Summarized credit score or modeled credit score by neighborhood
- Payday loan purchaser
- Number of credit lines
- Tax liens
- Card data:
 - Card holder - single card holder
 - Range of new credit
 - Debit or credit card present in household
 - Card holder - brand (Discover, Visa, Mastercard, etc.)
 - Card holder - type (Gas, bank, premium, luxury, prepaid, etc.)

- Frequent credit card user
- New retail card holders
- Underbanked or “thin file”
- Stocks or bonds
- Average online purchase
- Average offline purchase

In addition, a business may use enterprise data (historic data from its own customer files) to create proprietary or custom scores for its own use.

Scoring Models: How the Consumer Scores are Made, Part Two

Just as underlying factors going into a score should be fair and accurate, the algorithms that analyze the information should be of a high quality, should generate an accurate prediction, and should be validated against real-world data. As models are ultimately judged on their ability to make useful predictions from data, understanding how they actually perform and against what data sets is key. Without constant validation, scores might have no actual predictive value in reality. A bad or ineffective model ultimately means that the score does not offer accurate predictions. Bad scores based on a faulty or overfit model can still affect the treatment of individual consumers, the most important being eligibility and health care availability decisions.

Score creators have a good reason to get their models right. The marketplace is likely to weed out bad models, although it may take considerable time before this happens. The effects on individuals of poorly predictive consumer scores are uncertain. It would be useful for model makers to disclose their assumptions, predictive accuracy and model limitations. If a model is inherently a bad predictor of something important, then model users and model data subjects will want to know. Is it the data, the assumptions, overfitting, or other issues? A good faith, robust, public dialogue here could be helpful to all parties.

Algorithm, Inc.

Consumer scores generated by sophisticated mathematical models that detect patterns in information are often predictive, involve one or more algorithms, and rely on factors that describe individuals in some way.⁴⁶ The historical databases and raw consumer information that supply information to a score model can be both wide and large. Credit

⁴⁶ In a formal statistical model process, for example, data characteristics and attributes that describe a consumer are compared with a scoring table, or scorecard, and can be awarded points according to where they fall within the table. The points can be tallied to arrive at the overall score. Whether a high score means low or high risk depends on the model’s construction. For more information about scoring models, see the discussion of this topic in this report. See also D.J. Hand, *Statistical Classification Methods in Consumer Credit Scoring: A Review*, Journal of the Royal Statistical Society, 523-541. (1997).

scoring databases used to build score models, for example, may contain records of well over 100,000 individuals, and the model may measure over 100 factors, or variables. Behavioral scoring databases that store transactional data on for example, retail or other purchases, repayment, or other activities can be even larger.⁴⁷

Score models have advanced rapidly from their inception to the present time. The first widely used scoring model dates back to 1941. A paper published by Durand proved that discriminant analysis could produce reliable predictions of how individuals would repay credit extended to them. For detecting fraud, some of the older methods (pre-neural networks) of detecting fraud used lists of risky transactions and thresholds. It is worth noting that today's *score modeling* is an established field of study⁴⁸ as well as a competitive business.

In the world of scoring, the model used to generate the score is as important as the score itself. Two different score models, using identical factors, will almost always come out with a different score or metric for the same individual. A small error in the original calculations or inputs can magnify errors in the outcome.

The Federal Deposit Insurance Corporation, discussing credit score models, describes score models succinctly:

Scoring models summarize available, relevant information about consumers and reduce the information into a set of ordered categories (scores) that foretell an outcome. A consumer's score is a numerical snapshot of his or her estimated risk profile at that point in time. Scoring models can offer a fast, cost-efficient, and objective way to make sound lending decisions based on bank and/or industry experience. But, as with any modeling approach, scores are simplifications of complex real-world phenomena and, at best, only approximate risk.⁴⁹

The type of model a company uses to score consumers depends on the information analyzed, the purpose of the score, how much data is available, and a complex maze of other issues.⁵⁰ Ultimately, it is about how well the score model predicts what the company wants it to predict. Currently, widely used scoring models include discriminant analysis, linear regression, logistic regression, and decision trees or recursive partitioning, among other classification techniques.⁵¹ Credit scoring is a well-established methodology at this point, and models have been fine-tuned for decades. The most typical approach in

⁴⁷ Id. at 526.

⁴⁸ See for example the Journal of Operational Research Society, Oxford, England; The Royal Statistical Society, UK; Applied Statistics, (Journal).

⁴⁹ *Scoring and Modeling*, FDIC – Division of Supervision and Consumer Protection, March 2007. <http://www.fdic.gov/regulations/examinations/credit_card/ch8.pdf>.

⁵⁰ See Taylor James, *Predictive Analytics: Making Little Decisions with Big Data*, Information Management. (September 26, 2012).

⁵¹ D.J. Hand, *Statistical Classification Methods in Consumer Credit Scoring: A Review*, 160 Journal of the Royal Statistical Society, 524-531-35 (1997).

a credit model is a logistic regression model.⁵² Some consumer scores, such as fraud scores, employ completely different models, such as probabilistic or neural networks models. Banks experiment with neural network models,⁵³ and the credit card services industry uses neural networks for fraud detection.

Newer modeling techniques include the use of genetic algorithms, and an antibody approach. Ted Crooks, a leading model developer, described how antibody systems currently in development work:

The next step after neural networks, the things we're working on these days are systems that use antibody approaches. It [an antibody system] looks at hundreds of millions of possible combinations of transactions, and recognizes those that individual fraudsters or fraud rings have found popular lately.⁵⁴

Layering scoring models

Models can be layered on top of one another as well, and hybrid models that blend together results from various models are common⁵⁵. Consumer information is often fed into not just one, but a variety of scoring models, compared with a variety of test or control data, and then the “best” or most accurate model is then validated and chosen for final deployment in the business setting. Because varying models can change the quality and meaning of the final score so much, which scoring model is chosen makes a difference. The design is a complex balancing:

“Each model may employ a different subset of observations, consider different variables, make different assumptions about the relationship among the variables, and use different design concepts.”⁵⁶

When just one model is chosen, some information diversity is inevitably lost. To combat this, some experiment with combining two separate scores together to produce a *blended score* or a *combined score*. For example, a blending of a credit bureau score plus an application credit score would be a blended score. The idea is that the two scores together make a stronger predictive whole value.⁵⁷

⁵² LC Thomas, RW Oliver, DJ Hand, *A Survey of Issues in Consumer Credit Modeling Research*, 56 *The Journal of the Operational Research Society* (2005).

⁵³ *Id.* at 536.

⁵⁴ Interview, American Public Media, Marketplace Money, *Define Suspicious Activity*, (March 30, 2007).

⁵⁵ The model that won the Netflix prize consisted of blending 107 results.

⁵⁶ <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.142.9009>>.

⁵⁶ H. Zhu, P.A. Beling, & G.A. Overstreet, 52 *Journal of the Operational Research Society* 974-980 (2001) (Special Issue: Credit Scoring and Data Mining).

⁵⁷ *Scoring and Modeling*, FDIC – Division of Supervision and Consumer Protection, March 2007.

<http://www.fdic.gov/regulations/examinations/credit_card/ch8.pdf>. “While most scores and models are generally established as distinct devices, a movement to integrate models and scores across an account’s life cycle has become evident.”

Not only can a consumer characteristic such as payment history be used in a scoring model, a score can be used as a characteristic, too. As such, it is not unusual to learn of combined uses of consumer scores in scoring models, and even entire score networks.⁵⁸ So what goes into a consumer score can be characteristics about individuals, or can be a score that has already been created about the consumer, or both.

How many variables?

Some see the use of large numbers of variables in modern models as a positive.

In general, the ability to effectively use many different variables increases the strength of predictive models as it incorporates all available customer data that may be predictive as compared to traditional systems that are unable to scale since each variable must be tested by hand. This is particularly relevant in the insurance industry where there is a vast amount of customer information and a high correlation between the data and predictive outcomes.⁵⁹

The type of variables used in a consumer scoring model have been controversial over the years. One of the complaints with consumer scores is that the scores derive from a black box of big data without a lot of thoughtful selection. One analyst noted that in the past, the factors fed into a model were in some way tied to the model. Credit file factors would go into credit scoring models. But in today's world, numerous unrelated factors may go into a model. For example, one financial company uses information about whether a person types in all caps or all lowercase letters as a predictive factor for loan repayment.⁶⁰

Some models produce scores that give explicit estimates, others are numerical scales that reflect increasing levels of risk.⁶¹ Scoring models can score and rank individual consumers, the characteristics of portfolios of loans,⁶² or can even score neighborhoods. "Neighborhood scoring" has led to aggregate credit scores and other proxies for consumer credit classifications that are discussed elsewhere in this report.

Policy Questions

⁵⁸ H. Zhu, P.A. Beling, & G.A. Overstreet, 52 *Journal of the Operational Research Society* 974-980 (2001) (Special Issue: Credit Scoring and Data Mining).

⁵⁹ MMA Deploys KXEN's Infinite Insight to Boost Up-Sell, Cross-Sell and Customer Retention. ENP Newswire, 30 May 2013.

⁶⁰ Kenneth Cukier, Co-author, "[Big Data: A Revolution That Will Transform How We Live, Work and Think](#)," Council on Foreign Relations Federal News Service Media Call Subject: Big Data (May 9, 2013). The company Cukier mentioned on the call was *Just Finance*, according to the transcript.

⁶¹ John Copas, *The Effectiveness of Risk Scores: The Logit Rank Plot*, 48 *Applied Statistics* 166 (1999).

⁶² The Basel Capital Accord, which regulates banks' lending from 2007 forward, has pushed portfolio scoring forward in the financial sector. See LC Thomas, RW Oliver, & DJ Hand, *A Survey of Issues in Consumer Credit Modeling Research*, 56 *Journal of the Operational Research Society* (2005).

A key policy issue for consumer scoring models is who gets to see the underlying information fed into the model. Another important issue is whether the factors are discriminatory or prejudicial in any way. The use of health factors in a non-medical consumer score and the use of other sensitive factors in a scoring model are also problematic.

Another issue focuses on the type of model used for the analysis, and if the model was appropriately and fully validated and kept up to date. Credit and consumer score models must work, be accurate, and be updated regularly. If they are not, even small deviations can lead to inaccuracies. An inaccuracy in a credit report can make a consumer the target of predatory lenders.

Another issue arises with the use of a consumer score as an underlying factor in a new or different consumer score. Error upon error can accumulate in a way that even transparency will not enable a highly-educated consumer to untangle. Again, prejudicial, inappropriate, or unfair factors could be in the mix, but a consumer wouldn't know it.

When a predictive model assigns a value or a range to a consumer, the model used to create that value must be transparent, accurate, reliable, and kept up to date. The numeric range should be well-quantified, and the results validated.

Part III. The Consumer Scores

This section describes and documents major consumer scores and categories. We found the scores included here through a lengthy and diligent literature search. We also conducted interviews with scoring experts and others knowledgeable about scores. The list here is not comprehensive because much information about consumer scoring is not public.

In 2007, when the initial stages of pre-research for this report began, many fewer consumer scores existed and the documentation was sparse. The first iteration of this report was only a handful of pages long. Now seven years later, an entire business sector has grown around predictive analytics and consumer scoring. The growth in consumer scoring has been rapid, suggesting that we are at the beginning of a significant growth period, probably closer to the beginning of the predictive analysis bell curve than at the middle.

Category: Financial and Risk Scores

Consumer scores that measure risk are a large category comprising several major subtypes. These scores are not usually subject to FCRA rules, but the law's application depends on the structure, use, and factors of each score.⁶³

The most typical risk scores measure forms of consumer financial risk using non-credit factors or measure consumers for various types of fraud. In this report, we include authentication and identity scores in this category, as these scores ultimately seek to reduce or mitigate risk.

ChoiceScore

Experian produces **ChoiceScore**, a type of financial risk score.⁶⁴ Experian creates the score from consumer demographic, behavioral, and geo-demographic information. This score is used to segment consumers, as described here by a reseller of the data:

ChoiceScore by Experian UnderBanked and Emerging Consumers

ChoiceScore helps marketers identify and effectively target under-banked and emerging consumers. Using the most comprehensive array of non-credit data available from Experian. A financial risk score (indicating the potential risk of future nonpayment) provides marketers with an additional tool for more precise targeting.⁶⁵ The data card also indicated that the ChoiceScore could be used to suppress some consumers from getting information.⁶⁶

Experian's web site indicates that the ChoiceScore is not likely accessible by consumers. The score appears to be available for non-FCRA uses.⁶⁷ The score's factors are not published, so it is difficult to know what kind of underlying data is included in the score. It is also difficult if not impossible to determine what businesses are buying or using the score.

Modern data analytics have made child's play of unearthing people who are in various credit score brackets without revealing the actual credit score. Congress acted to protect

⁶³ Sending an advertisement to a consumer is not the same thing as sending a formal, pre-approved offer of credit as described in the FCRA. This risk score category includes risk scores that may well be used to generate leads, but the advertisements themselves are not formal pre-approved offers of credit. This difference was discussed at length at the FTC Alternative Credit Scoring Workshop (March 19, 2014). <<http://www.ftc.gov/news-events/events-calendar/2014/03/spring-privacy-series-alternative-scoring-products>>.

⁶⁴ Experian ChoiceScore, <<http://www.experian.com/marketing-services/data-digest-choicescore.html>>.

⁶⁵ CHOICESCORE BY EXPERIAN UNDER BANKED AND EMERGING CONSUMERS, <<http://datacardhub.adrearubin.com/market?page=research/datacard&id=268601>>.

⁶⁶ <<http://datacardhub.adrearubin.com/market?page=research/datacard&id=268601>>.

⁶⁷ According to the data broker's data card, two entities purchased this data: Achievecard, and Figi's Incorporated. Figi's Incorporated appears to be a food gift retailer. <<http://www.fbsgifts.com/about.html#figis>>.

the use of specific credit score information with good reason. From a consumer perspective, the same underlying principles still need to be at work: fairness, accuracy, transparency, and some reasonable limits on use.

Median Equivalency Score (Summarized credit statistics)

Experian's **Median Equivalency Score** "assesses the potential risk for seriously derogatory behavior."⁶⁸ Experian's material states: "The scores range from 360 to 840 (high score equals low risk) to accommodate the industry standard use of credit scores."⁶⁹

Summarized credit scores or statistics use geography to designate a credit statistic tied to a neighborhood.

"Summarized Credit Statistics are calculated by aggregating the available consumer credit data within a ZIP+4 geographic area. They do not communicate any individual consumer credit histories, but rather depict the consumer credit activity in a finite neighborhood."⁷⁰

This type of aggregation is typically done by analysis of census data overlaid with copious amounts of non-credit data, like consumer spending data. The summarized credit scores can then be used for marketing purposes broader than what the FCRA allows, because the FCRA applies to an individual's credit scores – not to this type of a neighborhood score.

An exemplar of the use of this summarized credit statistic is Experian's Summarized Credit Statistics mailing list.⁷¹ Experian described the list as follows:

"Summarized Credit Statistics data is derived from Experian's national consumer credit file and provides consumer credit activity in a neighborhood. The information is calculated by aggregating the available consumer credit data in each ZIP+4TM. Choose from more than 300 variables providing valuable information pertaining to tradeline status and specific types of tradelines.

Experian's Median Equivalency ScoreTM is a Zip+4 level score that helps you identify areas that may be more or less likely to have future derogatory credit activity. The score is statistically derived using payment information, utilization, mortgage, retail and other tradeline information aggregated at the ZIP+4 level."⁷²

⁶⁸ <<http://www.experian.com/assets/marketing-services/product-sheets/summarized-credit-stat.pdf>>.

⁶⁹ Id.

⁷⁰ Id.

⁷¹ Experian ConsumerView - Summarized Credit Statistics Mailing List. NextMark List ID #93574.

<<http://lists.nextmark.com/market?page=order/online/datacard&id=93574>>.

⁷² Id.

Often, scores of this type are used for lead generation. (Our analysis finds that lead generation is not the same thing as a pre-approved offer of credit. Risk scores might be used to generate leads, but the advertisements themselves are not formal pre-approved offers of credit.)⁷³

Additional Experian summarized credit statistic scores for ZIP + 4 neighborhoods include the:

- **National Risk Score**
- **National Equivalency Score**
- **National Bankruptcy Score**

The National Equivalency score is available to a consumer by either requesting an Experian file, or through a service like CreditSesame.⁷⁴

Risk IQ Score

The **Risk IQ Score** from AnalyticsIQ⁷⁵ uses summarized credit statistics to predict the “likely credit risk of individuals in a small geographic area.”⁷⁶ In this sense, it is similar to other risk scores that use summarized credit data that are applied to neighborhoods. The score is built from 100-plus sources, it uses 1,500 factors, and it is updated quarterly.

Most importantly, the Risk IQ score does not apparently use ECOA factors. It states directly in its materials “no protected class demographics are used in the model.”⁷⁷ This is a welcome statement in a risk model which does not likely fall under the FCRA. (As mentioned frequently in this report, the FCRA applies to individuals, not to neighborhood groupings.)⁷⁸

Consumer Profitability Score

⁷³ See e.g.,

<<http://lists.nextmark.com/market.jsessionid=480F77AC04EECF4E3BA8F18CB50CDFBF?page=order/online/datacard&id=93574>> (“Applications: Target candidates for invitations to apply for credit; Use as a predictive variable for acquisition and cross-sell models; Identify loyal prospects in ideal neighborhoods for publishing and continuity programs; Locate neighborhoods with recent and/or heavy credit purchase activity; activity may indicate families in new housing developments and neighborhoods undergoing revitalization where households have diverse product and service needs. Suggested users: Car dealerships/auto marketers; Catalogers and continuity clubs; Insurance providers; Investment planners; Tax services; Travel companies.” [Edited for space only. Copy of data card as of publication date available.]

⁷⁴ <<http://www.creditsesame.com>>.

⁷⁵ <<http://analytics-iq.com>>.

⁷⁶ <<http://analytics-iq.com/download/RiskIQ.pdf>>.

⁷⁷ Id.

⁷⁸ See supra 63 regarding difference between lead generation and formal offers of credit.

This score is designed to predict, identify, and target marketing prospects in households likely to be profitable and pay debt. Experian does not offer this score as a score subject to FCRA limits. Instead of being sourced from credit data, Experian sources this score from its proprietary ConsumerView database, which includes information about 235 million consumers and 117 million households from hundreds of data sources. The score is “rescored” or updated monthly.⁷⁹

While marketers may like these scores for identifying profitable consumers, these kinds of scores can potentially serve as unregulated proxies for credit risk. This score ranges from 1 – 13. Those households with a 1 are described as “High profitability, high likelihood to perform.” Those at 13 are designated as “Low profitability, unlikely to perform.”

It is important to understand that the **Consumer View Profitability Score** applies to a household. Household scores do not fall under the FCRA’s protections, because the FCRA applies to individuals, not households. Households at the lower end of the spectrum here are most affected by the absence of the rights that the FCRA gives to individual consumers. This score does not appear to use traditional credit file factors such as credit scores do. A credit score is limited in how it can be used for marketing. Companies who use a credit score must make a firm offer of credit or insurance. That is not the case with a risk score.

Experian describes the score:

“The ConsumerView Profitability Score combines a robust scoring model that offers high levels of refinement for selecting top profitability prospects with the best consumer database, ConsumerView, to deliver greater precision in predicting, identifying and targeting prospects at the household level.

The ConsumerView Profitability Score offers 13 levels with three high-profitability levels. This provides clients with additional precision in selecting the best prospects that will respond and comply with the terms of their Invitation-to-Apply lending, credit or continuity program offers.”⁸⁰

One of the ways Experian uses the consumer profitability score is in bundled targeted to different sectors. For example, Experian has a Healthcare bundle. This bundle is for healthcare organizations and marketers and includes the ConsumerView Profitability score, along with other Experian scores and a range of consumer demographic data. Here is Experian marketing pitch to healthcare companies:

“ConsumerView Healthcare: Healthcare organizations and marketers can leverage information about consumer’s lifestyles, interests and activities to help them distinguish relationships between various demographic and socioeconomic

⁷⁹ Experian Profitability Score (March 2014) <<http://www.experian.com/marketing-services/profitability-score.html>>.

⁸⁰ Id.

groups. This strengthens analyses, bolsters health risk assessments and makes consumer outreach initiatives more effective by identify those who are likely to be responsive to similar interventions, educational programs and communication initiatives (compliance, continuation of prescribed treatments, etc.) Individual characteristics such as age, marital status, education and occupation are provided along with other household insights including income, information about other members of the household, their dwelling type and length of residence among others. Furthermore, Mosaic USA lifestyle segmentation and a compact set of household expenditure propensities give healthcare marketers comprehensive visibility across a variety of dimensions.”⁸¹

Job Security Score

Scorelogix’s Job Security Score is an income risk-based score. The company describes the score as follows:

“Scorelogix is the inventor of the Job Security Score or JSS. The JSS is the industry’s first income-risk based credit score and the only score that predicts borrowers’ ability to pay by factoring their income stability. The JSS dramatically improves banks’ ability to reduce credit losses and marketers’ ability to reduce mailing costs.”⁸²

Scorelogix allows consumers to see their score, and the company includes an analytic report with the score. The score range is from 1- 1000, and the report includes a prediction of unemployment risk, a consumer’s relative ranking compared to others, key positive and negative risk factors influencing the score, and some additional information.⁸³

Consumer Prominence Indicator Score

Acxiom offers a **Consumer Prominence Indicator Score** that “quantifies the size of a specific consumer’s economic footprint, indicating the historical consumer purchasing and relative amount of marketing activity surrounding that individual.” The service appears to be aimed at marketers.⁸⁴

⁸¹ The Experian ConsumerView Healthcare (March 2014) <<http://www.experian.com/small-business/listdetails/consumerview-healthcare.html>>.

⁸² Scorelogix Home Page (March 2014), <<http://www.scorelogix.com>>.

⁸³ JSS landing page, (March 2014), <http://www.scorelogix.com/jss_landing_page02.asp?product_id=1>.

⁸⁴ See

<<http://lists.nextmark.com/market.jsessionid=0EB27D0D2B66AB1A48F0DA8EA9E66BBC?page=order/online/datacard&id=131838>> Title: Acxiom – InfoBase Consumer List Mailing List. Consumer Prominence Indicator noted in Selects. See also: *It’s All About the Data*, Slide presentation, Sandy Hurst, Acxiom Sales Team Leader, p. 11. <<http://webapps.franserv.com/spf11/downloads/8AcxiomData%20Services.pdf>>. See also Tom Hill, *The Future of Preferred Underwriting*, Society of Actuaries, Predictive Modeling Topics.

Discretionary Spending Index Score

Equifax offers a **Discretionary Spending Index (DSI)**. The DSI is a household scoring system based on the discretionary spending power of consumers. This service appears to be aimed at marketers.

“Discretionary Spending Index (DSI) is a continuous household-based score of 1 to 1000 that ranks households by likely spending capacity and spending behaviors. It enables marketers to rank customers and prospects by estimated spending power.

DSI can be used alone or incorporated into models where consumer spending is a factor. Marketers could use DSI to enhance account management, identify cross-sell opportunities, and provide appropriate offers.”⁸⁵

Invitation to Apply Score

These scores fall into the lead generation space, and as such, generally measure how likely a person is to respond to an offer. These generally do not fall under FCRA regulation. Several companies offer scores that fall generally in this category.

Fair Isaac at one time offered a marketing score (Qualify score) that allowed financial service marketers operating invitation-to-apply campaigns the ability to focus solicitations more narrowly on those more likely to respond. The score no longer appears to be in use. It is mentioned here as an historic exemplar.⁸⁶

Experian offers a product called **Veriscore** that helps to identify customer value potential.⁸⁷

“Assess the lifetime value of existing and prospective customers.

(August, 2013),

<<http://www.google.com/url?sa=t&rc=j&q=&esrc=s&source=web&cd=8&ved=0CFsQFjAH&url=http%3A%2F%2Fwww.soa.org%2Ffiles%2Fpd%2F2013-il-pref-under-sem-pred-mod-hill.pdf&ei=9IivU8KoNs3yqwGTjYCwBA&usg=AFQjCNFbGU5SNzReUTqy3zQMGYxIh50g8Q&bvm=bv.62922401,d.aWM>>.

⁸⁵ Equifax IXI Services (March 2014). <<http://www.ixicorp.com/products-and-services/customer-targeting-and-scoring/discretionary-spending-index-dsi/>>.

⁸⁶ Fair Isaac Qualify marketing score,

<http://www.businesswire.com/news/home/20040621005329/en/Fair-Isaac-Qualify-Score-Helps-Businesses-Reduce#Uy-K_FyaHxg>. This score appears to be defunct now, we mention it for historic purposes only. The score did not rely on credit data so the FCRA’s requirements for pre-screened offers of credit would not apply. For more information on prescreening, see the FTC publication at <<http://www.consumer.ftc.gov/articles/0148-prescreened-credit-and-insurance-offers>>.

⁸⁷ Experian-VeriScore (March 2014), <<https://www.experian.com/marketing-services/customer-value-marketing.html?cat1=>>>.

VeriScoreSM predicts response and lifetime value of new customers generated from alternate media sources such as call centers and registration forms. It evaluates their potential for fulfillment, cross-sell, up-sell and optimizes your database. Perfect for industries such as catalog, financial, fundraising, media, retail and telecommunications. Our models can be applied during merge/purge processing to produce a more targeted list without adding time to your production cycle. By allowing you to extract the best prospects from every list, VeriScore helps you find prospects most likely to respond and become loyal customers.”

InfoGroup Targeting Solutions, as part of its ITS Consumer Data services, creates a predictive response ranking from 1 to 9 based on consumer transactions and other information.

“Proprietary predictive response values ranking from 1-9 are based on *known transactions* taking into account recency of purchase, frequency of purchase activity, and dollars spent (RFM) within each market. The higher the RFM score indicates multiple purchases, frequent number of orders tied to dollars spent.”⁸⁸

Charitable Donor Score

Donor scores seek to classify and rank those who donate to charities. Donor scoring can be done internally within an organization, or donor scoring can be outsourced. Generally, donor research collects large amounts of information on potential donors from numerous sources including public securities filings, public ethics disclosure forms, probated wills, and other sources, including from non-profits’ existing donor lists.

SMR Research offers a **Donor Score**. The donor score is meant to identify the best prospects for large charitable donations.

“This Score predicts which U.S. households will make the largest contributions to charitable causes. The higher the score, the larger the donation usually will be.”⁸⁹

Blackbaud Sphere, offers analytic modeling to non-profits to assist with identifying a variety of donors.”⁹⁰ The company has robust analytics capacity and provides several models under its Target Analytics ProspectPoint modeling services. For example,

“The Major Gift model ranks and scores supporters and determines which of these individuals are most likely to make a major gift. It identifies not only which individuals have the capacity to make a major donation, based on overall wealth, income levels, and hidden assets, but also the propensity to give to the organization in significant amounts, as demonstrated by their profile and past behavior. The model is far more accurate than utilizing either capacity or

⁸⁸ See <<http://www.infogrouplistservices.com/b2c-data-solutions/its-consumer-data>>.

⁸⁹ <http://www.smrresearch.com/Charitable_Donors_Score_&_Lists.pdf>

⁹⁰ <http://internet.blackbaud.com/site/c.duIXLgOXJrIaE/b.8646093/k.99CB/Blackbaud_Sphere.htm>.

propensity alone and significantly reduces the risk that an organization will waste time and money investing in a nonproductive prospect.”⁹¹

DonorTrends offers a **DonorScore**.⁹² This model works with nonprofits’ internal databases to predict donor response:

“Our scientific DonorScores system assigns a value from 0 to 1,000 to each donor in your database. This value predicts the future actions each donor is likely to take. This enables you to target your donors more effectively to increase revenue and decrease cost.”⁹³

Household Segmentation Scoring Systems (Personicx, Mosaic, etc.)

Several companies offer targeting systems that match lifestyles, demographics, and spending habits with neighborhoods and households. These classifications are typically based on predictive analytics using varying models.

One of these companies is Claritas, which has a product called **PRIZM** that divides consumers into 66 segments with catchy names, such as Blue Blood Estates, Young Digerati, Gray Power, and Old Milltowns.⁹⁴

A similar system is **Personicx** by Acxiom. Personicx places each US household into one of 70 segments based on that household’s specific consumer and demographic characteristics.⁹⁵ These targeting systems qualify as consumer scoring systems even though they do not use a score because of the categorization of household by label.

Experian offers a similar product called **Mosaic**.⁹⁶

Collection and Recovery Scores

The use of collection and recovery scores likely falls under the FCRA, but it is unclear whether these categories of scores are actually exposed to consumers.⁹⁷

⁹¹ Blackbaud (March 2014), <https://www.blackbaud.com/files/resources/downloads/01.14.ANLY_%20ProspectPoint.datasheet.pdf>. See also <<https://www.blackbaud.com/ModelingExplorer>>.

⁹² <<http://donortrends.com>>.

⁹³ <<http://donortrends.com/performance-reporting/maap-report/>> See also <<http://donortrends.com/donorscores/>>.

⁹⁴ Claritas – MyBestSegments Segments Explorer, <<http://www.claritas.com/MyBestSegments/Default.jsp?ID=30>>.

⁹⁵ <<http://acxiom.com/personicx/>>. See Also *My Cluster*, <<https://isapps.acxiom.com/personicx/personicx.aspx>>.

⁹⁶ Experian. <<http://www.experian.com/marketing-services/consumer-segmentation.html>>.

⁹⁷ Consumer Finance Protection Bureau, Fair Debt Collection Practices Act and the Dodd-Frank Act (Bulletin 2013-08). There has been much discussion of the impact of debt collection on credit scores. This is beyond the immediate scope of this report. For more information, see <http://files.consumerfinance.gov/f/201307_cfpb_bulletin_collections-consumer-credit.pdf>.

FICO offers scores for delinquent accounts that predict whether the account is likely to pay, and if so, how much will likely be paid over a given time period.⁹⁸ The company calls these scores simply **FICO Collection Scores**.

“Adding more analytics for more precise decisions. FICO® Collection Scores are rapidly deployable analytics that typically boost collection performance by 15–20%. They include early-stage scores for cycle 1 and cycle 2 that rank-order accounts by their probability of rolling, as well as a late-stage score that ranks accounts by expected collection amount. FICO custom analytics include a wide range of predictive modeling (behavior, propensity, strategic default, attrition, etc.), decision modeling and optimization techniques.”⁹⁹

Experian has a product called **PriorityScore** for Collections.¹⁰⁰ It is designed to score and segment debt collection accounts.

Churn Scores

Churn scores seek to predict when a customer will move his or her business or account to another merchant (e.g., bank, cell phone, cable TV, etc.). These scores are abundant today. Any company that has historic customer sales data, can use the data to help build its own churn score.¹⁰¹ Many businesses create churn scores in-house or have an outside analytics company crunch the numbers for them.

Churn scores are interesting in many ways due to the unpredictability of customers.¹⁰² Churn scores are nevertheless well understood, to the point there is at least one patent application for a method of calculating churn.¹⁰³ Examples of companies that often have churn scores for customers are wireless telecommunications companies and cable providers, among many other business sectors.

Versium is one analytics company that creates churn scores for businesses.¹⁰⁴ Its churn scores are often custom scores made for specific businesses, and based on custom

⁹⁸ <<http://www.fico.com/en/products/fico-collection-score/>>.

⁹⁹ FICO, Insights White Papers, *Five Imperatives in a Shifting Collections Landscape*. Feb. 2013, No. 66. <<http://www.fico.com/en/products/fico-collection-score/>>.

¹⁰⁰ <<http://www.experian.com/consumer-information/debt-collections-strategies.html>>.

¹⁰¹ Emily Parkhurst, *What's Your Churn Score? Big Data Startup raises \$2.5M* (Sept. 10, 2013) Puget Sound Business Journal, <<http://www.bizjournals.com/seattle/blog/techflash/2013/09/whats-your-churn-score-big-data.html?page=all>>.

¹⁰² See, for example, Findable Consulting Blog, *3 Steps to Building Customer Churn Scores*, <<http://blog.findable.me/post/52410161427/3-steps-to-building-customer-churn-risk-scores>>. See also a discussion of creating the analytics for churn scores at <<http://www.analyticbridge.com/forum/topics/issues-with-predicting-churn?commentId=2004291%3AComment%3A72409>>.

¹⁰³ Eilam, Barak, Lubowich, Yuval, & Lam, Hila, *Method and Apparatus for Predicting Customer Churn*, US Patent Application 20090292583, <<http://www.freepatentsonline.com/y2009/0292583.html>>.

¹⁰⁴ Versium. <<http://versium.com/predictive->

enterprise data.¹⁰⁵ Versium is an exemplar of the trend of smaller analytics companies competing in a space that used to be reserved for much larger companies. Versium uses 300 billion attributes from over 8,000 compiled lists¹⁰⁶ covering up to, for example, 410 million unique emails, 240 million records of demographic data, 90 million mobile phone numbers, 120 million land line numbers, and 1.6 billion records of address history trail, among others. Some parts of this raw consumer data is going into score models for the churn score.

Another company with a churn score is Analytics IQ's **ChurnIQ** score.¹⁰⁷

Again, many churn scores exist. Due to the sheer number of businesses and analytics companies creating churn scores, the value ranges for the score may vary widely, as can the underlying factors used and the update schedule for the score. It is unlikely that many businesses make churn scores visible to their customers.

Category: Fraud Scores

Fraud scores are an important type of consumer risk score, and they comprise a major category of consumer scores. Fraud scores are prevalent. A large number of fraud scores are available today covering many risk types, with some of these scores in wide deployment. Those who received a phone call from their credit card company after making an unusually high credit card purchase have experienced a predictive fraud score in action. Fraud scores fall outside of the Fair Credit Reporting Act in almost every case. While there may be a fraud score that falls under the FCRA, the research undertaken for this report did not uncover such a score.

As a result, a consumer's fraud score — and most consumers will have multiple fraud scores — will not normally be available for viewing, or for correction or dispute. It is easy to understand the benefit of fraud scores. Companies that build predictive fraud scores have statistics showing how much reduction in fraud the score creates. In some cases, the amount of fraud reduction can be substantial, above 50 percent. There are downsides, though. False positives are highly problematic for the consumers saddled with them. This is particularly challenging for victims of identity theft — either financial, medical — who may have damaged or erroneous fraud scores imputed to them.

Fraud scores are part of the fabric of many businesses today, in particular retailers and financial sector businesses. A consumer with a fraud score that indicates high risk can have difficulty transaction business routinely. Declined credit purchases, declined loans, and declined financial or in some cases health services are among the most common impacts. False positives for fraud scores can vary significantly depending on the scoring

scoring/?PHPSESSID=844c277cc5f173b37a1bd0255c33caec>.

¹⁰⁵ WPF Interview with Versium (March 2014).

¹⁰⁶ Versium Data Sheet. <<http://versium.com/wp-content/uploads/2014/01/versium-com-data-sheet.pdf>> and <<http://versium.com/lifedata/>>.

¹⁰⁷ Analytics IQ. <<http://analytics-iq.com/products/loyaltychurn/>>.

model used and the factors fed into the model. Consumers subject to a false positive may find it difficult to clear their records because of the lack of transparency and formal rights.

Data brokers and analytics companies that sell fraud scores are secretive about the score factors and models. Little is known about them among the general public in comparison, to, for example, consumer credit scoring. It is not possible for consumers to approach FICO and request their Falcon Fraud score in the same way it is for FICO customers to request their FICO credit score. Consumers have no rights because the FCRA does not apply here.

This report discussed the problems with opacity of scores, particularly those that have noticeable impact on consumers. At one time, the credit score was secret due to concerns that consumers would game the credit system if they knew their scores. However, as stories grew of the abuse of credit scores, lawmakers took progressively stronger action to ensure consumers could see their credit scores, including structural protections such as constraining use of the scores and providing consumers a right of access. Fraud scores play a significant role in consumers' lives. We need a discussion about the fairness, accuracy, underlying factors used in the scores, and about the non-transparency of the scores.

FICO Falcon Fraud Manager

FICO sells an anti-fraud product that creates a near-real-time fraud score for consumers called **Falcon Fraud Manager**¹⁰⁸, which it describes as follows:

“Accurately detect fraud on payment card authorization and electronic payment transactions in a fraction of a second. Identify suspicious account holder, cardholder and, optionally, device behavior patterns, generating a score indicating likelihood of fraud.”¹⁰⁹

Falcon stands for Fuzzy Adaptive Logic Control/Decision Network. The Falcon score relies on a neural network, and FICO claims a high score accuracy.¹¹⁰ Accuracy rates vary, depending on the product and its usage. Fraud scores can hover around an 85 percent or higher accuracy rate. FICO states its false positive rate is around 4 percent.

Of course, claims of accuracy are common for consumer scoring products, but independent verification of the claims are rare. Generally, little is known about the values or ranges for fraud scores. Credit card purchase behavior, device behavior, and other known consumer demographics are likely candidates making up the fraud score at any given moment for an individual.

¹⁰⁸ FICO provided WPF an extensive in-person background on this product.

¹⁰⁹ <<http://www.fico.com/en/products/fico-falcon-fraud-manager/>>.

¹¹⁰ See for example, FICO page on Falcon Fraud Manager, multiple white papers available.

<<http://www.fico.com/en/products/fico-falcon-fraud-manager/>>. See also Fraud Score Accuracy, <<http://versium.com/fraud-score-accuracy/>> and FICO.

Other Fraud Scores

Corelogic has a range of **LoanSafe** products that check and score loan applications for fraud, among other mortgage fraud prevention services.¹¹¹

CoreLogic has another product called **ThirdParty Scorecard**. It assesses an agent's loan quality, assigning a risk score to each agent that can be compared against internal, local market and industry performance standards.¹¹² These scores can indirectly affect consumers, who have no way of know if their brokers or agents are viewed as trustworthy within the industry. A consumer using a broker who has a poor risk score may not be able to obtain a loan or may pay a higher price without knowing the real reason.

Interthinx has a range of mortgage fraud and verification products such as **FraudGuard** and **SafeCheck**.¹¹³

VISA Risk Manager / Visa Advanced Authorization— cardholder real-time risk scoring.¹¹⁴

ReD PRISM (Proactive Fraud Risk Management) (neural network). ReD PRISM® is a transaction monitoring and risk management tool for card issuers, merchant acquirers. This tool generates scores and reason codes for transactions.

“Patented neural network, pattern-recognition software, a fraud detection model and an Active Cardholder History Database that typically holds 30 days of cardholder activity.

The engine generates a score and reason codes for each transaction processed. Scores can be generated in real-time, to be part of an authorization system, and/or in near-real-time for post authorization analysis.”¹¹⁵

MasterCard fraud scoring solution, a collection of products under the umbrella of **Expert Monitoring Solutions**.¹¹⁶

Versium has a fraud score that has reported high accuracy levels of 85% with a false positive rate of 4%.¹¹⁷

¹¹¹ CoreLogic Mortgage Fraud Solutions, <<http://www.corelogic.com/solutions/mortgage-fraud-solutions.aspx>>.

¹¹² <<http://www.corelogic.com/products/data-repository-solutions.aspx#container-ProductDetails>>. See also, *Fifth Third Mortgage Correspondent Seller Guide* (Sept. 28, 2012), <<https://www.53.com/files/doc/cl/seller-guide/Seller%20Guides%2010.05.12.pdf>>.

¹¹³ <http://www.interthinx.com/solutions/mortgage-fraud-verification-services?pi_ad_id=36143850429&gclid=CMyljPiEvL0CFcqUfgod8xUAPQ>.

¹¹⁴ <<http://www.visa.com/visariskproducts/>> and <<http://visa.ca/merchant/security/layers-of-security/advanced-authorization.jsp>>.

¹¹⁵ ReDPrism <<http://www.redworldwide.com/fraud-prevention/red-prism/>>.

¹¹⁶ <<https://www.mastercard.com/us/company/en/whatwedo/products.html>>.

There are too many fraud scores to be comprehensive here. A sampling of other types of fraud scores include:

- **Medicare Advantage** is risk scoring for fraud
- **FICO Insurance Fraud Manager**¹¹⁸
- **LexisNexis FraudPoint** (applicant fraud prevention)¹¹⁹
- **Volusion** credit card fraud score¹²⁰
- **Kount Score** (Prevent fraudulent web purchases)¹²¹

Category: Custom Scores

Custom scores are those scores uniquely calibrated for a particular business, or that use a particular set of proprietary customer data, or both. The scores can use a combination of internal customer information from the business and information from data brokers and other external sources. Custom scores can be about almost anything to do with customer patterns.

Thanks to the sophistication of today's data brokers, availability of large streams of data, and accurate data analytics, custom scores are in wide use now. Their use may well increase over time as retailers and other businesses seek to increase understanding, segmentation, and targeting of new consumers and existing customer bases. Custom scores are typically closely held, and only limited information is little available on the public record.

The Emergence of Custom Scores and the Pregnancy Predictor Score Example

This report reviews only a portion of available consumer scores. A significant part of consumer scoring is entirely hidden from public view because of the emergence of custom scores. Custom scores can assess almost anything to do with customer patterns. If a business can leverage its own customer data with a custom score, it may have a unique asset.

Perhaps the most famous custom score thus far that became public is Target's **Pregnancy Predictor Score**. This score came to light due to the reporting of Charles Duhigg, who wrote a 2012 article for the New York Times, *How Companies Learn Your Secrets*. In

¹¹⁷ <<http://versium.com/fraudscore/>>. See also <<http://versium.com/resources-roc-curve/>>.

¹¹⁸ <<http://www.prnewswire.com/news-releases/fico-unleashes-new-analytics-for-fighting-americas-700-billion-healthcare-fraud-waste-and-abuse-problem-172324041.html>>. For general information about health fraud, see <<http://jama.jamanetwork.com/article.aspx?articleid=191726>>.

¹¹⁹ <<https://www.lexisnexis.com/risk/solutions/fraudpoint-fraud-prevention.aspx>>.

¹²⁰ <<http://onlinebusiness.volusion.com/articles/volusion-launches-fraud-score/>>.

¹²¹ <<http://www.kount.com/products/complete/kount-fraud-score/>>.

this article, Duhigg described in detail the Target pregnancy score and how Target developed it.¹²² Target's deep databanks of past customer behaviors across its broad customer base was the fodder for the score. Target used predictive analytic models to draw conclusions from the data.

Duhigg described how Target acquired such a robust customer database. Target – at least at the time Duhigg published his article describing the practice – assigned its customers a Guest ID whenever possible, which effectively linked purchases and activities over time. Duhigg also described what is now known to be a common practice among large retailers especially of adding outside data from data brokers to existing customer data. Duhigg describes it this way:

“Also linked to your Guest ID is demographic information like your age, whether you are married and have kids, which part of town you live in, how long it takes you to drive to the store, your estimated salary, whether you've moved recently, what credit cards you carry in your wallet and what Web sites you visit. Target can buy data about your ethnicity, job history, the magazines you read, if you've ever declared bankruptcy or got divorced, the year you bought (or lost) your house, where you went to college, what kinds of topics you talk about online, whether you prefer certain brands of coffee, paper towels, cereal or applesauce, your political leanings, reading habits, charitable giving and the number of cars you own.”¹²³

This process, called data appending, reverse data append, or data enhancement, is well-documented as a common practice. For example, a landmark California case, *Pineda v. Williams Sonoma*,¹²⁴ revealed reverse data append activity that is usually non-available to the public eye. The complaint states:

“Plaintiff visited one of defendant's California stores and selected an item for purchase. She then went to the cashier to pay for the item with her credit card. The cashier asked plaintiff for her ZIP code and, believing she was required to provide the requested information to complete the transaction, plaintiff provided it. The cashier entered plaintiff's ZIP code into the electronic cash register and then completed the transaction. At the end of the transaction, defendant had plaintiff's credit card number, name, and ZIP code recorded in its database.

Defendant subsequently used customized computer software to perform reverse searches from databases that contain millions of names, e-mail addresses, telephone numbers, and street addresses, and that are indexed in a manner resembling a reverse telephone book. The software matched plaintiff's name and

¹²² Charles Duhigg, *How companies learn your secrets*, New York Times (Feb. 16, 2012), <http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?_r=1&pagewanted=all>.

¹²³ *Id.*

¹²⁴ See, e.g., *Pineda v. Williams Sonoma Stores, Inc.*, No. S178241 (CA Sup. Ct, 2012), <<http://caselaw.findlaw.com/ca-supreme-court/1555490.html>>.

ZIP code with plaintiff's previously undisclosed address, giving defendant the information, which it now maintains in its own database. Defendant uses its database to market products to customers and may also sell the information it has compiled to other businesses."¹²⁵

This description was one of the first full explanations of the inner workings of data append to the public.

Using its customer data, Target's predictive formulas combed through customer patterns and identified approximately 25 products that, if purchased together in a certain period of time, suggest a likelihood of pregnancy. Predictive analytics allowed Target to assign a pregnancy score to shoppers who matched the criterion. It then used the scores to send ads to women whom Target predicted might be pregnant. The result reported by Duhigg was of a father storming into Target angrily with a handful of baby-related advertisements from the retailer stating that his daughter was not pregnant. Later, he discovered that his high-school aged daughter – unbeknownst to him – was indeed pregnant.

We know because of Duhigg's article about just one Target custom score. We do not know how many other customer scores Target uses. We do not know what other retailers use custom scores. If trends are any indication, the rapid upswing in analytics companies offering custom scores and custom analytics as well as access to massive, personally identifiable, consumer data sets hints that use of custom scores may at some point overtake off-the-shelf prefabricated scores. Custom scoring makes it even harder for consumers to learn how merchants use their data.

As stated throughout this report, some scores are less troublesome to consumers. But many may be problematic to a greater or lesser degree, and consumers do not know what they do not know about consumer scoring in general and about custom scoring in particular. We all deserve to live in a world where we have the opportunity to know how we are being sorted and sifted and to have transparency and fairness in any ranking process.

Category: Regulated Credit and Financial Scores

Traditional credit scores are regulated. In our taxonomy, a credit score is a type of consumer score. There are hundreds of credit scores available on the market today. Any credit grantor can develop and use its own general or specifically focused scoring system. Although consumers have rights to see credit scores, they may not know about the existence or use of those systems, especially for scores developed internally by a credit grantor and not purchased from external vendors.

Any businesses can develop a credit scoring system and can sell consumers the ability to see their scores, whether or not anyone actually uses the credit scores to make real world decisions about consumers. A household that wants to see multiple credit scores for

¹²⁵ Id.

household members could spend a considerable sum buying scores. Consumers may not know which scores are actually used and which are not. This has the potential to be lucrative for the business but potentially expensive for consumers, and this is the reason consumers need to be cautious about which credit scores they are purchasing.

The use of credit scores outside the immediate credit granting process is hard to assess. There is one report that an electric utility in Texas wanted to use credit scores to set electricity prices to some residential users. The Texas Office of Public Utility Counsel objected and the plan apparently never went into effect.¹²⁶ See the discussion about insurance scores.

FICO Score

A well-known credit score is the FICO Score.¹²⁷ There are numerous credit scores, but FICO appears to lead in terms of sales.¹²⁸ Fair Isaac states:

“The FICO® Score is the most widely used credit score in North America. Lenders purchased more than 10 billion FICO Scores in 2013, and 90 percent of all U.S. consumer lending decisions use the FICO Score. The 25 largest credit card issuers, the 25 largest auto lenders and tens of thousands of other businesses rely on the FICO Score for consumer credit risk analysis and federal regulatory compliance.”

The FICO score falls squarely under the Fair Credit Reporting Act.

FICO also has an **Expansion Score** that draws on alternative credit data such as bank account records, payday loan payment records, and installment purchase plans, to produce a credit score that may not be based solely on the contents of a credit report.¹²⁹ Other companies also offer similar credit scores for people with “thin” credit reports.¹³⁰

Vantage Score

The three national credit bureaus have jointly developed the VANTAGE score, which is also a popular credit score.¹³¹ The Vantage score is regulated under the Fair Credit Reporting Act.

¹²⁶ See <<http://www.liheap.ncat.org/news/Sept04/Texas.htm>>. Title: Texas OPC Says TXU Credit Scoring Puts Poor at Risk.

¹²⁷ <<http://www.myfico.com>>.

¹²⁸ Press Release, FICO, *New Version of the Industry standard FICO Score Will Be Available Beginning This Summer*, <<http://www.fico.com/en/about-us/newsroom/news-releases/new-version-industry-standard-fico-score-will-available-beginning-summer/>>.

¹²⁹ <<http://www.fairisaac.com/Fairisaac/Solutions/FICO+Expansion+Score/Expansion+Score+Overview/FICO+Expansion+Score.htm>>.

¹³⁰ <<http://biz.yahoo.com/brn/060908/19082.html>>.

¹³¹ <<http://www.vantagescore.com>>.

Beacon Score

A variant on the general credit score is Equifax's **BEACON** service, which seeks to predict the likelihood that a new or existing account will become delinquent within 24 months.¹³² The BEACON score is regulated under the FCRA.

When low Beacon scores are used to market to consumers

It is possible that an individual's low Beacon score can be used indirectly for marketing purposes through circumstances the authors of the FCRA did not predict. For example, consider an individual whose Beacon scores was too low to qualify for a cell phone contract. That individual could end up on a data broker list of "Cell Phone Turndowns." WPF first recorded this list in 2007,¹³³ and has confirmed the existence of a similar list in 2014.

The 2014 Cell Phone Turndowns list read in part:

"This file is comprised of individuals who attempted to set up a contract with a cell phone provider but did not meet the required Beacon score requirements. These consumers are ready and eager to receive offers and opportunities in the following categories: secured and sub-prime credit, Internet, legal and financial service, health insurance offers, home equity loans, money making opportunities, and pre-approved credit with a catalog purchase."¹³⁴

Even if a score begins life as subject to FCRA limits, it is easy for industry to track a consumer's activities and extract those that reflect a score at a particular level. This is what happened with the cellphone turn down list. Those consumers clearly have low credit scores. The resulting list of those denied phone would not fall under any regulation, and it could be used as a proxy for a regulated credit score.

Small Business Intelliscore

The focus of this report is on scores for individuals. There are also numerous scores for businesses available for purchase. It is worth mentioning at least one exemplar of this type of score. Experian offers a product called Small Business Intelliscore.¹³⁵ It uses commercial and consumer credit data to generate a risk score. This is an example of a scoring system that, for small businesses, fuzzes the line between consumer and commercial activities.

¹³² <https://www.eport.equifax.com/eport/eport_beacon.htm>.

¹³³ *Cell Phone Turndowns*, DirectListFinder2.0, NextMark ID 188161, <<http://listfinder.directmag.com>>, (June 16, 2007). PDF copy of the list available from WPF.

¹³⁴ *Cell Phone Turndowns*, NextMark ID188161,

<<http://lists.nextmark.com/market?page=order/online/datacard&id=188161>> (March 13, 2014).

¹³⁵ <http://www.experianbizinsight.com/data_enhancement/intelliscore.shtml>.

Tenant Scores

Numerous tenant scores are available. These scores give a history of evictions, and use credit bureau and other data.¹³⁶ CoreLogic is one of the companies offering a tenant score, theirs is called **CoreLogic SafeRent**.¹³⁷ Tenant scores fall under FCRA regulation.¹³⁸

Category: Identity and Authentication Scores

Identity and authentication scores are in widespread use. These are a form of modern eligibility scores, although current regulations do not view these scores as regulated eligibility scores.

Consumers encounter these scores, for example, when logging on to an online patient portal hospital system or to a financial institution for online banking. Some forms of ID or authentication scoring is invisible to the consumer. Entire businesses specialize in authentication. These kinds of scores are a subcategory of consumer risk scores. It is unusual for a consumer to find or see their ID or authentication score, but there are exceptions. The authentication of online users is a reasonable activity and a protection for both the consumer and the business. The problem arises when an innocent consumer – perhaps the victim of identity theft – loses access to accounts or is denied service and cannot readily learn the reason or correct the problem.

ID Analytics ID Score

ID Analytics is one of several companies that offer ID scores. Its **ID Score** “calculates the risk associated with an identity, allowing businesses to focus Identity Risk Management efforts on identities with the highest likelihood of fraud without alienating legitimate customers.”¹³⁹ ID scores are in widespread use, and can be troublesome to consumers when something is amiss. For example, if a person fails an ID score, it is an indication that they may either be an identity thief, or be an identity theft victim. At times, a low ID score will prevent a person from purchasing items like a cell phone, and can be troublesome in other verification situations.

ID Analytics is one of the few companies that allow consumers to view their ID score at no cost to the consumer.¹⁴⁰ This has been in place for about six years. Consumers cannot

¹³⁶ See for example <<http://myrental.com/reports/tenant-score/>>.

¹³⁷ <<http://myrental.com/aboutus/>>.

¹³⁸ See *FTC Warns Data Brokers That Provide Tenant Rental Histories They May Be Subject to Fair Credit Reporting Act* (April 3, 2013), <<http://www.ftc.gov/news-events/press-releases/2013/04/ftc-warns-data-brokers-provide-tenant-rental-histories-they-may>>. See also Privacy Rights Clearinghouse on Renter’s Rights for more on Tenant Screening and the FCRA. <<https://www.privacyrights.org/renters-guide-to-privacy-what-to-know>>.

¹³⁹ <<http://www.idanalytics.com/solutions/score.html>>.

¹⁴⁰ <<http://myidscore.2advanced.com>>.

learn the factors that go into the score nor dispute the score, but it is nevertheless a step in the right direction that consumers can see their ID score.

Fair Isaac has an ID product called **FICO Identity Resolution Manager**.¹⁴¹

Insurance Scores

Insurance scores, sometimes called credit-based insurance scores, fall under the FCRA. The scores use credit scores and credit information to analyze prices for automobile insurance and homeowners insurance. Insurance scores differ from credit scores, and it appears that insurance companies may have their own algorithms. More information about the use of credit scoring in insurance is available from the National Association of Insurance Commissioners¹⁴² and from Consumer Reports.¹⁴³

ChoiceTrust, a **ChoicePoint** company, offers to sell home and auto insurance scores to individuals.¹⁴⁴ **FICO** also has an insurance score product. A growing number of insurance carriers use custom scores that have been developed to meet that company's specific underwriting criteria.¹⁴⁵

It is unknown whether or how insurance companies use the insurance industry property claim databases (generally referred to as CLUE or Comprehensive Loss Underwriting Exchange) for scoring.¹⁴⁶ Consumers can request a copy of their CLUE reports under the FCRA.¹⁴⁷

Category -- Health Scores

Initial research for scoring done for this report in 2007 found few health scores. In 2014, research uncovered significant and high-impact consumer health scores in use. Health scores are now in full circulation with little consumer awareness. The same questions raised above about transparency, secrecy, factors, and use are relevant here. Other

¹⁴¹ <<http://www.fico.com/en/products/fico-identity-resolution-manager/>>.

¹⁴² <http://www.naic.org/cipr_topics/topic_credit_based_insurance_score.htm>.

¹⁴³ <http://www.consumerreports.org/cro/personal-finance/car-insurance-8-06/overview/0608_car-insurance_ov.htm>. See also <<http://www.insurancescore.com/>>.

¹⁴⁴ <<http://www.choicetrust.com/servlet/com.kx.cs.servlets.CsServlet?channel=welcome&subchannel=insscore#>>.

¹⁴⁵ <https://choicetrust-solutions.custhelp.com/cgi-bin/choicetrust_solutions.cfg/php/enduser/std_adp.php?p_faaid=617&p_created=1050502344&p_sid=fYAN1Nri&p_accessibility=0&p_lva=&p_sp=cF9zemNoPTEmcF9zb3J0X2J5PSZwX2dyaWRzb3J0PSZwX3Jvd19jbnQ9MTEmcF9wcm9keZ0mcF9jYXRzPTE3NiZwX3B2PSZwX2N2PTEuMTc2JnBfcGFnZT0x&p_li=&p_topview=1>.

¹⁴⁶ More information about CLUE is available from the Privacy Rights Clearinghouse <<http://www.privacyrights.org/fs/fs26-CLUE.htm>>.

¹⁴⁷ <http://personalreports.lexisnexis.com/fact_act_claims_bundle/landing.jsp>.

questions come into play as well. For example: can employers purchase health scores? Are health scores ever shared with debt collectors?

New health scoring systems that fall in the category of consumer scores will be developed and used in the near future. It is also possible to foresee the development of family and neighborhood health scores based either a combination of traditional medical histories, genetic data, census data, data broker lists, environmental data, or histories of actual health treatments that may fall outside of HIPAA.

Health records held by health care providers or insurers are subject to the federal health privacy rules known as HIPAA.¹⁴⁸ While these records are available for many non-consensual uses, the information in the records should not normally be available to data brokers and score creators. However, the HIPAA rules do not cover health information held by gyms, websites, banks, credit card companies, many health researchers, cosmetic medicine services, transit companies, fitness clubs, home testing laboratories, massage therapists, nutritional counselors, alternative medicine practitioners, disease advocacy groups, or marketers of non-prescription health products and foods. This vast class of largely unregulated health information is available as input to a health scoring algorithm. Further, consumers routinely disclose health information to companies that promise to provide coupons. Consumers rarely understand that companies can collect personal information that they can later sell.

Personal health records (PHR) maintained by companies outside HIPAA protections may also become a source of unregulated health information for scoring.¹⁴⁹ Information disclosed through web searches or Internet browsing also typically remains unregulated by HIPAA, and all of the information can be fodder for scores.

Affordable Care Act Individual Health Risk Score

Each individual in a health plan subject to risk adjustment under the Affordable Care Act (ACA) will be assigned a health risk score. This is a new score, and it is an important score especially because it is part of a federal program. In establishing the rules for health risk scores for individuals, the Health and Human Services Department effectively created a score that ultimately measures how sick a person is. The stated goal of the risk score is to create a relative measure of predicted health care costs for a particular enrollee. The scores are supposed to be phased out over the next four years.

The rules for the individual health risk score became official in March 2012, when the Department of Health and Human Services issued a final rule on reinsurance and risk adjustment under the Affordable Care Act.¹⁵⁰ The overall purpose of risk adjustment is to mitigate the impact of potential adverse selection and stabilize premiums in the individual

¹⁴⁸ Health Insurance Portability and Accountability Act, 45 C.F.R. Parts 160, 162, & 164.

¹⁴⁹ See World Privacy Forum, Personal Health Records: Why Many PHRs Threaten Privacy (2008), <<http://www.worldprivacyforum.org/2008/02/blog-legal-and-policy-analysis-personal-health-records-why-many-phrs-threaten-privacy/>>.

¹⁵⁰ <<https://www.federalregister.gov/articles/2012/03/23/2012-6594/patient-protection-and-affordable-care-act-standards-related-to-reinsurance-risk-corridors-and-risk>>.

and small group markets under the Affordable Insurance Exchanges that are part of the Affordable Care Act.

A key element of the risk adjustment is the calculation of a plan's average actuarial risk so that the plan's average risk can be compared to other plans. The scores will be important because they will determine whether a plan pays or receives funds through the premium adjustment system. A plan might have an incentive to assign its insured a higher score. The use or disclosure of that score for another purpose could harm an individual. Even disclosure of an honest score could be harmful. This is a new area and a new score, and there is much uncertainty about the use or misuse of the score.

A plan's average risk is based on the risk score of each enrollee in that plan. An individual's health risk score will be a measure of how much that individual is likely to cost the health plan. The risk score measures likely health costs and is, in a very general way, a proxy for how sick an individual is. How expensive an individual will be to insure is important to insurers and employers, and the score can easily be misused.

The HHS rule took some care to protect the privacy and security of an individual's risk score, including limits on the disclosure of identifiable elements when individual risk scores are passed on by a plan for use in State risk adjustment programs. Nevertheless, each individual in plans subject to risk adjustment will have his or her own health risk score.

The regulation is silent about individuals seeing their health risk score. If an insurer has a risk score for an individual, then it appears that it would be *Protected Health Information* as defined in the privacy rules issued under the rules of the Health Insurance Portability and Accountability Act (HIPAA). If that conclusion is correct, the score should be available to individuals under standard HIPAA rules. It is possible to foresee that an employer or lender or someone else with power over an individual might coerce the individual into obtaining his or her score and disclosing it.

FICO Medication Adherence Score (MAS)

Launched on June 23, 2011 by analytics firm Fair Isaac Corp., this score identifies a patient's propensity to adhere to a medication prescription plan during the next 12 months. It is a predictive score designed to let pharmacies and insurers know when or if a patient is at risk and needs a medication reminder.¹⁵¹ The score pulls from public data and from patients' prescription histories when available. The score ranges from 1-500, with a score above 400 indicating that a patient is likely to take medications as prescribed. Patients who score 200 or below may get a reminder, as a low score predicts non-adherence. The company created the scoring algorithm from a randomized sample set of

¹⁵¹ Jeremy M. Simon, *New medical FICO score sparks controversy, questions*, [Creditcards.com](http://www.creditcards.com/credit-card-news/fico-score-medication-adherence-1270.php) (July 28, 2011), <<http://www.creditcards.com/credit-card-news/fico-score-medication-adherence-1270.php>>. See also Tara Parker Pope, *Keeping score on how you take you Medicine* (June 20, 2011). NYT Blog Well, <http://well.blogs.nytimes.com/2011/06/20/keeping-score-on-how-you-take-your-medicine/?_php=true&_type=blogs&_r=0>.

Pharmacy Benefit Manager patients numbering in the millions. (More on what a PBM is here).¹⁵²

By the end of 2011, FICO scored 2 to 3 million patients, with an additional 10 million expected during 2012. New numbers were not available for later years.

Factors in the score include:

- Employment
- Homeownership
- Living situations
- Age
- Gender
- Family size
- Asset information (ex., likelihood of car ownership).

FICO states its analytics identify medication adherence risk by “using data from a range of publicly available third-party data sources. Because the FICO **Medication Adherence Score** requires minimal information from the patient, and no prescription claims or sensitive health information, the score can be generated for members of any patient population.”¹⁵³ FICO states that with just a name and home address it can “pull the remainder of the necessary information from publicly available sources.”¹⁵⁴ As of 2014 FICO also states “The Medication Adherence Score will use a patient’s prescription claims history when available and pull on other publicly available third-party data sources when no other information is present.”¹⁵⁵

These differences in what factors are used to create the score make a difference. If FICO calculates the score without any health information obtained from a covered entity regulated under the HIPAA federal health privacy rules,¹⁵⁶ then the information is not regulated as health information under HIPAA. This illustrates a limitation of the HIPAA privacy rules that allow information about patients to be used and disclosed, bought and sold, by data brokers and others without application of any health privacy rules.

If, however, FICO calculates a score about an individual based on any information – even just the individual’s name – obtained from a business associate under HIPAA rules, then it would appear that FICO uses protected health information under HIPAA and it would have to be a business associate of the provider or insurer that disclosed the information to FICO. This would bring this score under the HIPAA regulations.

¹⁵² For more on PBMs and HIPAA, see Gellman and Dixon, *A Patient’s Guide to HIPAA*, Basic Rights. World Privacy Forum, Sept. 2013. <<http://www.worldprivacyforum.org/2013/09/hipaaguide40/>>.

¹⁵³ FICO, Press release, *New FICO Analytics Predict Likelihood of Patient Adherence to Prescription Medication*, (June 23, 2011). <<http://www.fico.com/en/about-us/newsroom/news-releases/new-fico-analytics-predict-likelihood-of-patient-adherence-to-prescription-medication/>>.

¹⁵⁴ <<http://www.fico.com/en/products/fico-medication-adherence-score/>>.

¹⁵⁵ Id.

¹⁵⁶ 45 C.F.R. Parts 160, 162, 164.

Because of the two different scenarios that are possible here, it is impossible to tell from the outside just what FICO does with the score. It seems possible that FICO and HIPAA-covered entities could potentially organize the sharing of information to evade HIPAA's requirements. For example, FICO could take its entire list of individual scores, give that list to a HIPAA covered entity and allow the covered entity to select information about patients of the covered entity. That would result in no sharing of HIPAA-protected health information with FICO.

Possible customers for FICO's Medical Adherence Score are pharmaceutical manufacturers. Unlike health plans, drug manufacturers do not have any direct way of learning who is taking the drugs that they manufacturer. They need the assistance of intermediaries, like pharmacies and pharmacy benefit managers, to send prescription reminders. HIPAA allows these reminders.

Drug manufacturers fund many if not most prescription reminder programs. To send a reminder, the manufacturer pays a HIPAA-covered entity – most likely a pharmacy or pharmacy benefit manager – to contact the patient lawfully. The manufacturer must pay for the cost of the notice to the patient and provide an incentive to the intermediary. The full cost might be a few dollars per notice. If the manufacturer can identify those patients who are likely to refill prescriptions anyway, it can tell the intermediaries to send reminders only to those who have a low adherence score. The effect is to pay less to FICO and avoid paying a larger amount for a notice. We do not know if this reflects how the scores are actually used.

FICO states that patients can ask their health care providers if they have a score. For patients with a MAS score, FICO directs them to ask their health care providers about their opt out policies. Under HIPAA, patients should be able to request this score, as it should be Protected Health Information and subject to HIPAA transparency rules if a HIPAA covered entity maintains the score. However, an opt-out from a third party health score is uncharted territory for HIPAA. WPF's Patient's Guide to HIPAA has a section detailing how to request health records under HIPAA.¹⁵⁷ It is not always an easy or a simple process, and it can require a great deal of persistence just to find the right provider who has the information. It is not clear to us how providers might treat a MAS score request, and it is unknown if any would honor a request for an opt out. In short, it is somewhat disingenuous for FICO to direct patients to the HIPAA process when it is FICO that maintains a patient's MAS score.

Frailty Scores: General

Frailty Scores usually apply to the elderly. A good bit of research has been conducted using this score as a measure. As a result, a frailty score has become much more important in recent years. Research found that some frailty scores could predict mortality

¹⁵⁷ Gellman and Dixon , *A Patients Guide to HIPAA*, Right to Inspect and Copy Your Record, FAQs 18-24. <<http://www.worldprivacyforum.org/2013/09/hipaaguidepart2/>>.

within one year.¹⁵⁸ Separate research indicated some frailty scores can usefully predict the likelihood of patient post-operative surgical complications or readmission to a hospital. While the scores can predict care needs, the scores can also be used to simply project costs, and this raises questions about possible misuse in non-health scores or marketing activities. Unless a HIPAA-covered entity calculates a frailty score using health records, the score is not likely covered by the HIPAA health privacy rules.

CMS Frailty Adjustment Score

The Centers for Medicaid developed a frailty score in the late 1990s. In 2004, after refinement, the CMS frailty measure was extended to more Medicare managed care organizations.¹⁵⁹ CMS is a HIPAA-covered entity so the score should be subject to the HIPAA health privacy rules. After CMS developed its score, several other models of frailty scores developed.

Hopkins Frailty Score

Johns Hopkins University developed the **Hopkins Frailty Score**. Designed for use before surgery, the score would be calculated by a health care provider and would be subject to HIPAA. This predictive score in its original form has low factors compared to other scores and a small range. The factors are highly predictive, however, and this score is in widespread use.¹⁶⁰

It is unknown how many patients are assigned frailty scores, and it is unknown how many patients ever request their scores. Conceivably, a score held by a health care provider should be covered under HIPAA and patients should be able to request their score if one is there.

The concern with any predictive score, particularly a frailty score, is that it can escape into the hands of third parties where it can be used outside of the original intent of the score. The frailty score can be highly predictive, and therefore its use needs to be carefully guarded.

Other Health Scores

¹⁵⁸ Dave Levitan, *Frailty Score Predicts 1-Year Mortality But Not Procedural Complications in TAVR* (Sep. 17, 2012), <<http://www.tctmd.com/show.aspx?id=113395>>.

¹⁵⁹ John Kautter and Gregory C. Pope, *CMS Frailty Adjustment Model*, 26 M.S. Health Care Financing Review (2004-2005), <<https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/HealthCareFinancingReview/downloads/04-05winterpg1.pdf>>.

¹⁶⁰ See Johns Hopkins Medicine, *Level of Frailty Predicts Surgical Outcomes in Older Patients, Johns Hopkins Researchers Find* (May 12, 2010), <http://www.hopkinsmedicine.org/news/media/releases/level_of_frailty_predicts_surgical_outcomes_in_older_patients_johns_hopkins_researchers_find>.

A medical data breach revealed in 2011 that a company called Accretive collected detailed and sensitive health information about hospital patients in Minnesota via contract with those hospitals and then used that data to develop scores. The information included:

- Patient’s full name
- Gender
- Number of dependents
- Date of birth
- Social Security number
- Clinic and doctor
- A numeric score to predict the “complexity” of the patient
- A numeric score to predict the probability of an inpatient hospital stay
- The dollar amount “allowed” to the provider
- Whether the patient is in “frail condition”
- Number of “chronic conditions” the patient has

Patients had no knowledge of the use of the information for scoring:

“Upon information and belief, the hospitals’ patient admission and medical authorization forms do not identify Accretive by name or disclose the scope and breadth of information that is shared with it. Upon information and belief, patients are not aware that Accretive is developing analytical scores to rate the complexity of their medical condition, the likelihood they will be admitted to a hospital, their “frailty,” or the likelihood that they will be able to pay for services, among other things.¹⁶¹”

The Minnesota Attorney General found that company promoted its health data activities to investors as:

- Risk scoring of patients
- Has an “intense focus” on “reducing avoidable hospital admissions”
- Identifies the “sickest and most impactable patients” for “proactive management”
- Identifies “real-time interventions with significant revenue or cost impact”¹⁶²

The full details of Accretive’s activities are beyond the scope of this report because they involved matters beyond scoring. However, when the lawsuit ended, Accretive agreed to a ban from doing business in Minnesota for a period of time.¹⁶³

Personal Health Scores: WebMD, others

¹⁶¹ Complaint, *State of Minnesota vs. Accretive Health, Inc* (USDC Minn 2012).

¹⁶² Minnesota Attorney General, Press Release, Attorney General Swanson Sues Accretive Health For Patient Privacy Violations (Jan. 19, 2012), <<http://www.ag.state.mn.us/Consumer/PressRelease/120119AccretiveHealth.asp>>.

¹⁶³ Star Tribune.com, *Accretive is banned from Minnesota* (July 31, 2012), <<http://www.startribune.com/lifestyle/health/164313776.html>>.

A personal health score is a growing category of scores that, at the moment, are relatively casual and aimed primarily at enhancing a consumer's self-understanding. These scores do not carry the same underwriting weight as for example, a large sample-set based, formal statistical score would. The scores generally appear to be largely educational in nature, and voluntary.

Under the Affordable Care Act, wellness programs and health improvement are priorities. It is no surprise, then, that new health self-scoring activities for patient self-monitoring are coming online. These scores, by their nature, are typically generated by an online survey taken by the patient, with the resulting scores available to the patient. Although many variations exist of these sorts of more casual health scores, at this point most of the scores do not appear to be tied to benefit costs.

WebMD is a good exemplar of this kind of "personal" health score.¹⁶⁴

One Health Score is another exemplar.¹⁶⁵ This score allows a consumer to rate their physical activity and its benefits. This score has a range of 1-100, with a score of 60 and above indicating that the person being scored has a basic level of physical activity. Scores of 90 and over are generally attained by professional athletes.

These scores are likely not subject to HIPAA protections. If the scores derive from information supplied by a consumer, then they are not protected health information under HIPAA unless a HIPAA covered entity calculated the score. A commercial website offering services to consumers is normally not a HIPAA covered entity. Even if the website maintains health records for the consumer and with the consent of the consumer, use of the records is subject to the privacy policy and terms of service of the website. Most consumers would likely not understand that health information held on their behalf by commercial websites has no legal privacy protections. How the use of these health scores will evolve and whether they will "escape" into the hands of marketers and data brokers is not known.

Resource Utilization Group Scores

Medicaid uses resource utilization groups (RUG) to classify residents in nursing homes based on the relative resources that an individual is likely to use. Medicare pays for Part A skilled nursing facility stays based on a prospective payment system that categorizes each resident into a payment group (RUG) depending upon his or her care and resource needs. Skilled nursing facilities determine a RUG based on 108 items on an assessment of the resident known as the Minimum Data Set (MDS).¹⁶⁶ Calculated by a HIPAA covered entity (the nursing home), the score remains subject to HIPAA privacy rules.

¹⁶⁴ WebMD, *Health Manager, Personal Health Score* <<http://www.webmd.com/health-manager>>.

¹⁶⁵ One Health Score, <<http://www.onehealthscore.com/faq>>.

¹⁶⁶ See HHS Inspector General, *Review of Nursing Facility Staffing Requirements at Beverly Healthcare of Reading* (2005), <<http://www.oig.hhs.gov/oas/reports/region3/30400214.htm>>.

SF-36 Form

A standard health industry/research form is the SF-36 (the SF-12 is a shorter version). The SF-36 is a multi-purpose health survey that produces an 8-scale profile of functional health and well-being scores as well as physical and mental health summary measures and a health utility index. The SF-36 is used for surveys of general and specific populations, to compare the relative burden of diseases, and to differentiate the health benefits produced by different treatments. The data subject completes the survey.¹⁶⁷ Whether HIPAA protections apply to the information on the survey depends on who collects the information. If a HIPAA covered entity (health care provider or insurer) collects the survey, the information falls under HIPAA. If a patient completes the survey for a health researcher, it may or may not fall under HIPAA. Many researchers are not covered entities under HIPAA, and many fall under no privacy regulations at all.

This type of instrument could produce a consumer score depending on who uses it and the purpose. If used for general actuarial purposes by a health plan, the results would probably not qualify as a consumer score. If used by a medical device merchant to target likely wheelchair purchasers, it would qualify as a consumer score. If used in a research project, the SF-36 would not result in a consumer score.

Complexity Scores

Complexity scores are beginning to spring up for various patient types and situations (See Frailty score). Grants have been set aside for the development of new complexity scores, for example, work to create a Complexity Score to Identify Hospitalized Patients at High Risk for Preventable Adverse Drug Events was funded in 2013.¹⁶⁸ It is likely that complexity scores will be developed for many patients' situations. A complexity score used for treatment fall under HIPAA and does not qualify as a consumer score. A complexity score used for marketing or to set rates may be a consumer score.

An exemplar complexity score is the Aristotle Complexity Score. This score was developed over the course of five years by the Aristotle Institute. Used in its original context, this score is **not** a consumer score because it is for diagnostic use.

A group of 50 internationally accepted experts has been working for more than five years on a new method to evaluate the quality of care in Congenital Heart Surgery (CHS) that is called Aristotle. Senior, experienced congenital heart surgeons considered the possible risk factors for each procedure and assigned scores based on potential for mortality, potential for morbidity, and anticipated surgical difficulty.

¹⁶⁷ See <<http://www.sf-36.org/tools/sf36.shtml>>.

¹⁶⁸ See <<http://www.ashpfoundation.org/MainMenuCategories/PracticeTools/Drug-Therapy-Management-Complexity-Score-Index>> and <<http://www.ashpfoundation.org/PR2013ComplexityScore>>.

The Aristotle system, electronically available, has been introduced by both the European Association for Cardio-Thoracic Surgery (EACTS) and Society of Thoracic Surgeons (STS) as an original method to compare the performance of Congenital Heart Surgery (CHS) centers. Pediatric cardiologists have joined the project and are currently developing a complexity score for interventional cardiology procedures.¹⁶⁹

The Aristotle score, allows precise scoring of the complexity for 145 CHS procedures.¹⁷⁰ Again, as with the complexity score, if used in a clinical setting, these scores should fall under HIPAA and should be viewable by patients. Also as with frailty scores, complexity scores could be subject to abuse if layered into scores outside the health care context.

Category – Smart Grid and Energy Scores

The Internet of Things and the Smart Grid stands to generate a significant number of consumer scores in the very near future. At least one exemplar score is already in use, with many others set to come into use soon.

Peer-to-Peer Energy People Meter Score (EPM)

This score measures a residential customer's energy consumption and seeks to engage the customers to evaluate their own energy consumption patterns. Consumer scores arising from Smart Grid¹⁷¹ or Internet of Things¹⁷² usage is an emerging field. These scores are of great interest due to the approaching tsunami of information that connected devices in and out of the home, including cars, will provide. The **EPM score** is a proprietary score from Trove Data. The company has a range of analytics in the area of energy, not all of which qualify as consumer scores. The Energy People Meter score is of specific interest here. Trove Data describes it as follows:

“Trying to engage your customers? How can you score their actions as a customer? TROVE's Peer-to-Peer Scoring application and proprietary Energy People Meter (EPM) captures the attention of the customer and inspires them to evaluate their energy usage. By providing individual recommendations for each customer, this tool allows utilities to increase customer service and engage customers in new ways.”¹⁷³

¹⁶⁹ <<http://aristotleinstitute.org/aboutScore.asp>>.

¹⁷⁰ The Aristotle score: a complexity-adjusted method to evaluate surgical results. <<https://www.ncbi.nlm.nih.gov/pubmed/15144988>>.

¹⁷¹ See Comments of EFF and CDT to the Department of Energy on Implementing the FIPS in the Smart Grid, (Nov. 1, 2010), <<https://www.eff.org/files/DOE%20Comments-Nov1.pdf>>.

¹⁷² See Comments of EPIC to the FTC on the Privacy and Security Implications of the Internet of Things, (June 1, 2013), <<https://epic.org/privacy/ftc/EPIC-FTC-IoT-Cmts.pdf>>.

¹⁷³ Trove Data Utility Applications, <<http://www.trovedata.com/solutions/utility-applications/>>.

Trove Data lists its data sources used to compile its scores in a FAQ section. Its list includes meter data, satellite imagery, and appliance data. Its FAQ states:

“We aggregate and fuse thousands of data of attributes from hundreds of sources including the following:

- Meter data
- Demographic data
- GIS
- Distributed energy data
- Satellite imagery
- Financial data
- Wireless/Mobile data
- Appliance data
- Social behavior data
- Other disparate data sources that can provide deep insight into energy behavior patterns.”¹⁷⁴

While it is likely that other companies provide competing analytics scores in this area, Trove Data is an intriguing case study as a data aggregator because it was among the first to work with Smart Grid information.¹⁷⁵ A benefit of the Trove EPM score is in its environmental potential, allowing energy companies to fine-tune power supply and usage.

The initial EPM score is not exposed to the consumer, however, the company de-identifies consumer data and aggregates the information. The company has taken numerous steps to remove personally identifiable information from the data the utilities companies receive.¹⁷⁶ **Versium** has an energy score it is planning to introduce as well.

The issue of Smart Grid and other Internet of Things data flows and how to handle the privacy implications of these devices and the data they shed over the long term is complex. The Federal Trade Commission held a conference on this topic, which generated robust conversations on how privacy protections might be configured in Smart Grid scenarios.¹⁷⁷ FTC Chairman Edith Ramirez recommended privacy by design, simplified consumer choice, and transparency as important privacy starting points.¹⁷⁸ This is an area of consumer scoring where much can be done right now to be proactive in

¹⁷⁴ Trove Data FAQ, *What data sources does Trove use?* (March 15, 2014), <<http://www.trovedata.com/solutions/faq/>>.

¹⁷⁵ See an early article about the company, Katie Fehrenbacher, *GridGlo mines big data for real time energy apps*, GigaOm (May 10, 2011), <<http://gigaom.com/2011/05/10/gridglo-mines-big-data-for-real-time-energy-apps/>>.

¹⁷⁶ Interview, Trove Data, (March 2014).

¹⁷⁷ Federal Trade Commission, *The Internet of Things: Privacy and Security in a Connected World* (Nov. 19, 2013) (Conference, webcast, transcripts, and public comments). <<http://www.ftc.gov/news-events/events-calendar/2013/11/internet-things-privacy-security-connected-world>>.

¹⁷⁸ At 9-10, <http://www.ftc.gov/sites/default/files/documents/public_events/internet-things-privacy-security-connected-world/final_transcript.pdf>.

terms of mitigating any potential consumer downsides so as to reap the benefits of the data analysis.

Category - Social Scoring

Much has been made of social scoring over the past few years. Entire companies, web sites, books, college courses focus intensively social scoring as social scores themselves continue to proliferate. Social scores seek to measure influence, or Social Networking Potential (SNP) The scores derive from a specific type of algorithm called a “social algorithm.”¹⁷⁹ Michael Schaefer writes in his book *Return on Influence*:

“This trend of social scoring is creating new classes of haves and have-nots, social media elites and losers, frenzied attempts to crash the upper class, and deepening resentments.”¹⁸⁰

Some social scores enjoy a transparency that most of the scores mentioned in this report do not. **Klout**,¹⁸¹ **Booshaka**,¹⁸² **Kred**,¹⁸³ **PeerIndex**,¹⁸⁴ **PROskore**,¹⁸⁵ **SocialIQ**,¹⁸⁶ **Tweet Grader**,¹⁸⁷ and **Twitalyzer**¹⁸⁸ are among those providing social media analytics and metrics directly to the public. Among these, Klout is currently the most important and is in mainstream use. People can readily see their own – and others’ – Klout scores, which has created a complex dynamic. People can opt out of having a Klout score if they dare, but in some professions, not having a Klout score would be a professional liability.

Social scoring systems used by traditional data brokers are more difficult to find, evaluate, and quantify. Nevertheless, some information is available. Acxiom collects household social media predictors such as “Social media sites likely to be used by an individual or household, heavy or light user, whether they engage in public social media activities such as signing on to fan pages or posting or viewing YouTube videos.”¹⁸⁹ However, the metrics for this collection, noted in a 2013 GAO report, do not appear on Acxiom’s About the Data portal, so it is not possible to readily understand how Acxiom assesses social data. Analytics firm Versium offers a social influencer scoring, it uses its own observational scores.¹⁹⁰

¹⁷⁹ See for example <http://en.wikipedia.org/wiki/Viral_advertising#References>.

¹⁸⁰ Mark W. Schaefer. *Return on Influence: The revolutionary power of Klout, social scoring, and influence marketing* at Introduction(2012).

¹⁸¹ <http://klout.com/home>

¹⁸² <http://www.booshaka.com/product/overview>

¹⁸³ <http://www.kred.com>

¹⁸⁴ <http://www.peerindex.com>

¹⁸⁵ <http://proskore.com>

¹⁸⁶ <http://socialiq.com>

¹⁸⁷ <http://www.tweetgrader.webs.com>

¹⁸⁸ <http://twitalyzer.com/5/index.asp>

¹⁸⁹ Government Accountability Office, *Information Resellers, Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace*, at 52 (Table 2) (2013), <<http://www.gao.gov/assets/660/658151.pdf>>.

¹⁹⁰ <<http://versium.com>>.

It is unlikely that data brokers would ignore the important and relevant category of social scoring, but whether and how they use specific scores like Klout as factors in their own score calculations is an unknown. Data brokers that create their own custom, non-public social scores may be a larger category than is generally known, but it may take time for this information to come to public light, if it ever does.

Klout Score

The Klout score is the best-known social-media based score. Anyone who has a Twitter account usually also has a Klout score. People can see their Klout score and can opt out of the score if they wish.¹⁹¹ Large brands use the score to give free or discounted products and services to high scorers, called Klout Influencers. Common examples are reduced rates on a hotel room or free upgrades on flights. The score is fairly well-understood, and although the Klout algorithm is secret, the score itself is transparent. Mark Schaefer described Klout concisely:

“Klout compiles more than 100 different factors across dozens of social media platforms, pumps billions of pieces of data through its algorithms, and creates a personalized assessment of influence that ranges from 1 to 100. The world average is about 19. Someone with a score of 30 shows expertise, whereas a score of 50 or more means leadership and expert status. And a perfect 100? That is reserved for one person alone: Justin Bieber.”¹⁹²

The Klout score has ignited some privacy controversies. One early privacy snafu¹⁹³ occurred in 2011 when Klout was found to be inadvertently scoring minors. The company stopped the practice immediately and apologized. Other questions remain. What are the privacy consequences when people are socially scored and others can readily see those scores? It is the fundamental premise of this report that scoring is a modern decision-making methodology and shorthand that, relying on complex algorithms, is going to become important in modern life due to its prevalence and use. Given that the Klout score can be seen and has an opt out, some of the transparency and choice concerns that plague many of the other scores may diminish, but the secrecy of the Klout score's composition and opacity of use remain prime concerns.

An unusual factor in the Klout score is that those who have a score and have not opted out have a public score that anyone can see. Just to make a point, credit scores can be seen by the individual, but are not available to the public without constraint. The public nature of social scoring is something new, and it creates intriguing social effects that are absent in other types of scoring. This is at the heart of the complexity of what comprises

¹⁹¹ <<http://klout.com/corp/optout>>.

¹⁹² Mark W. Schaefer. *Return on Influence: The revolutionary power of Klout, social scoring, and influence marketing* at chapter 7 (2012).

¹⁹³ At one point, Klout was scoring minors. In a blog post, the CEO apologized and Klout discontinued the practice. <<http://blog.klout.com/2011/11/we-value-your-privacy/>>.

modern privacy. Anyone who really wants to can usually achieve a high Klout score, or at least a passable score, as tips for accomplishing this abound. In fact, on eBay, one can outright purchase Twitter followers to ostensibly enhance the score.¹⁹⁴

High scorers receive benefits – and these benefits can be financial or professional. Individuals with high Klout scores can be and have been treated preferentially by companies seeking to capitalize on their perceived clout. A high Klout score could help those seeking jobs in public relations, marketing, or other fields where Klout familiarity and understanding is helpful.¹⁹⁵

Mark Schaefer tells the story of Sam Fiorella, a top marketing executive who lost a job interview because his Klout score of 45/100 was deemed too low. When Fiorella followed up with the company after a post-interview period of silence, he was told by the company that his online influence was “not sufficient for the job requirements.”¹⁹⁶ Fiorella’s Twitter profile now does not disclose a Klout score, and he has written a book about social media. His analysis is that the social scores are fading in importance to some degree as more understanding about how brand influence in social media marketing works.¹⁹⁷ The use of any score in an eligibility decisions is tricky, and if the Klout score were to be used in this way with any consistency it would become the subject of much debate.

It may not be clear, however, how those with low or no Klout scores will fare in job or other marketplaces. An individual may never realize that he or she did not receive an interview, job, discount, premium, coupon, or opportunity due to a low score. It is hard to hear the dog that doesn’t bark. An individual denied credit based on a credit report is entitled to know the reason. An individual denied *something* based on a Klout score or the absence of a Klout score has no similar entitlement.

Beyond eligibility, the use of social scores for performance review or for public labeling is also a challenge. One sales conference posted a list of its speakers’ Klout scores in a top-40 roster.¹⁹⁸ The good news about the Klout score is that it is public. But perhaps because it is so public, available to all, new kinds of challenges are created. For example,

¹⁹⁴ Search term “Twitter followers” typed on eBay led to 35 items. One item advertised “4,000 REAL Twitter followers!.” Other items included books and other materials on how to increase Twitter followers. <http://www.ebay.com/sch/i.html?_trksid=p2050601.m570.11313.TR0.TRC0.H0.Xtwitter+followers&_nk w=twitter+followers&_sacat=0&_from=R40>. Search conducted March 15, 2014.

¹⁹⁵ For example, a social media marketing position. One such position was advertised on CareerBuilder, and listed among the job requirements: “Familiarity with the key tools involved in social marketing, including measurement devices such as Radian 6, and influence tools, such as Klout.”, <http://www.careerbuilder.com/JobSeeker/Jobs/JobDetails.aspx?source=jrp&APath=2.21.0.0.0&job_id=J3G0ZF6F6FQ17RCN03N&sc_emp1=js_jrp_viewjobdesc&IPath=ILKV0A>. Retrieved March 15, 2014.

¹⁹⁶ Schaefer, *Return on Influence*, at Chapter 1.

¹⁹⁷ Email exchange, March 2014. See also Fiorella’s book in which he discusses this issue at length, Pearson Publishing: *Influence Marketing: How to Create, Manage, and Measure Brand Influencers in Social Media Marketing* (2013).

¹⁹⁸ InsideSales.com, *Top 40 Klout Scores at the 2014 Sales Acceleration Summit* (March 12, 2014), <<http://www.insidesales.com/insider/social-selling-2/sales-acceleration-summit-top-40-klout-scores/>>.

opting out is not a realistic option for people in certain occupations. Having a low score could be professionally detrimental, just as having a high score could be beneficial.¹⁹⁹ A change in Klout's algorithm could have a major impact on an individual's status or marketplace opportunity without any notice or way for the individual to find out. Even if Klout scores lose their influence, individuals will be affected in various ways during the transition.

Another significant question is whether data brokers use or combine Klout scores and other publicly available scores in individual profiles of consumers. That answer is shrouded in obscurity, but it stands to reason that publicly available scores can be found and used by data brokers if they choose. If so, the Klout score may have more influence than even the company itself realizes, because the score could be used in algorithms that determine consumer placement and rank on lists for a wide range of consumer offers and non-offers. Klout scores could influence prices individuals pay through differential pricing algorithms.²⁰⁰ Data broker scores have too little transparency for any reasonable fact-finding in this area. It will be important to discover how data brokers are using social scores in their product ranges, and if these are publicly available scores, or custom social scores. The absence of transparency in this area is troublesome in the short and long term.

Employment Success Score

Researchers from a trio of universities figured out a way in 2012 to analyze a person's Facebook page to create a score predicting their job success.²⁰¹ Employers have not used this score in any formal way or to scale. If a score like this was developed for more widespread use, it would likely fall under the Fair Credit Reporting Act. (It would also be non-compliant with Facebook's Terms of Use.)

Tax Return Scores

The Internal Revenue Service scores tax returns using several different scoring systems. Computer programs calculate a numeric score for each tax return. The Discriminant Function System (DIF) score rates the potential for change as a result of an audit, based on past IRS experience with similar returns. IRS also has an Unreported Income DIF (UIDIF) score that rates a return for the potential of unreported income. IRS personnel uses the scores to select for audit and to identifying the items on these returns that are the

¹⁹⁹ Intriguingly, this applies not just to individuals, but also groups, businesses, schools, and other entities. For example, see Molly Greenberg, *You'll Never Guess Which DC Area College Has the Best Klout Score* (March 12, 2014). <<http://inthecapital.streetwise.co/2014/03/12/best-college-klout-scores/>>.

²⁰⁰ See generally, Robert Gellman, *Differential Pricing and Privacy: Good, Bad, or Otherwise?* (2014), <<http://www.concurringopinions.com/archives/2014/03/differential-pricing-and-privacy-good-bad-or-otherwise.html>>.

²⁰¹ Eve Tahmincioglu, *Facebook profiles predict job success* (Feb 22, 2012), Today.com, <<http://www.today.com/money/facebook-profiles-predict-job-success-1C8368043?franchiseSlug=todaysmoneymain>>.

best candidates for review.²⁰² Given the internal use of these scores and the mildly restrictive law²⁰³ on the privacy of IRS records, the IRS score may be of lesser immediate concern.

Category – Law Enforcement Scores, including Police, Transportation, Safety, and other

The government creates and uses risk scores for its work.²⁰⁴ A number of government scores exist around transportation, safety, anti-terrorism, and other law-enforcement related activities. This information, however, is very difficult to find in the public domain.

However, a substantive 2013 Rand report on predictive policing shed important light on a wide range of police use of predictive scoring techniques to determine which individuals would be most likely at high risk to offend in the future. It is an important baseline study on this emerging issue.²⁰⁵

Automated Targeting System Score

The screening of airline passengers by the Department of Homeland Security has been a subject of ongoing controversy for several years. The details of the system are still not fully publicly known, but what is known is that the program collects data about passengers and links the data with other sources of information to establish a risk score for each passenger. The Transportation Security Administration uses the scores to screen passengers.

The Privacy Impact Assessment for the program states:

“ATS provides equitable treatment for all individuals in developing any individual’s risk assessment score because ATS uses the same risk assessment process for any individual using a defined targeting methodology for a given time period at any specific port of entry.”²⁰⁶

²⁰² <<http://www.irs.gov/newsroom/article/0,,id=151888,00.html>>.

²⁰³ 26 U.S.C. § 6103. While this law has many loopholes, it should keep IRS records out of the hands of data brokers.

²⁰⁴ See Martin Bosworth, *FBI Uses Data Brokers, Risk Scores to Hunt Terrorists*, (July 11, 2007), <https://www.consumeraffairs.com/news04/2007/07/fbi_risk_scores.html> and Ellen Nakashima, *FBI Plans Initiative to Profile Terrorists*, Washington Post (July 11, 2007), <<http://www.washingtonpost.com/wp-dyn/content/article/2007/07/10/AR2007071001871.html>>.

²⁰⁵ Walter L. Perry, Brian McInnis, Carter C. Price, Susan C. Smith, & John S. Hollywood, *Predictive Policing: The role of crime forecasting in law enforcement operations* (2013) (Rand Corporation), <<https://www.ncjrs.gov/pdffiles1/nij/grants/243830.pdf>>.

²⁰⁶ Department of Homeland Security, *DHS Privacy Impact Assessment for ATS* (Aug. 3, 2007), <http://www.dhs.gov/xlibrary/assets/privacy/privacy_pia_cbp_ats_updated_fr.pdf>. See also Department of Homeland Security, *PIA Update*, (January, 2014), <<https://www.dhs.gov/sites/default/files/publications/privacy-pia-cbp-atsupdate-01312014.pdf>>.

More information on passenger screening is available from the Electronic Privacy Information Center²⁰⁷ and the Identity Project.²⁰⁸ A DHS report reviews legal and other problems with the Automated Targeting System.²⁰⁹ The DHS scores must be shared with airlines, but whether the scores leak into the commercial marketplace is uncertain.

Richard Berk Algorithm

Criminologist Richard Berk developed a predictive model to identify murderers. Kim Zetter of *Wired* wrote, “To create the software, researchers assembled a dataset of more than 60,000 crimes, including homicides, then wrote an algorithm to find the people behind the crimes who were more likely to commit murder when paroled or put on probation.”²¹⁰ Pennsylvania, Maryland, and Washington D.C. use the software.

Youth Delinquency Scores

The Foundation for Information Policy Research in the United Kingdom completed a report identifying the growth in children’s databases and assessing the data protection and privacy implications.²¹¹ The report describes structured assessment tools for the youth justice system in England and Wales that create profiles of young offenders by examining the factors that may have brought each youth into contact with the criminal justice system. The assessments are scored for adverse factors, and the score is used to predict the likelihood of re-offending.²¹² Whether any comparable U.S. scoring systems exist is unknown. What is known is that the UK scores are subject to the UK data protection law. Similar scores created by U.S. states would not necessarily fall under any privacy regulation.

Predictive Anti-Fraud Scores: US Postal Service Office of Inspector General

The US Postal Service Office of Inspector General has a predictive analytics team that uses predictive fraud scores to address point-of-sale fraud issues. As described, the Postal Service has a customized fraud detection tool. “Using more than 30 indicators to search a wide variety of data, the fraud detection tool flags and ranks instances of suspicious

²⁰⁷ <http://www.epic.org/privacy/airtravel>.

²⁰⁸ <http://papersplease.org/wp/>.

²⁰⁹ <http://www.dhs.gov/xlibrary/assets/privacy/privacy-secure-flight-122006.pdf> and DHS Privacy Impact Assessment ATS System, http://www.dhs.gov/xlibrary/assets/privacy/privacy_pia_cbp_ats_updated_fr.pdf.

²¹⁰ Kim Zetter, *Precog Software Predicts Crime*, *Wired* (Jan. 2013).

²¹¹ <http://www.wired.com/2013/01/precog-software-predicts-crime/>.

²¹² Foundation for Information Policy Research, *Children’s Databases – Safety and Privacy* (2006), http://www.fipr.org/childrens_databases.pdf.

²¹² *Id.* at 49-50.

activity, allowing investigators in the Postal Service's Office of the Inspector General to decide which leads to pursue."²¹³

Category -- Environmental Scores

EPA Health Risk Score

The EPA uses substantial predictive analytics tools, and has a **Human Toxicity Risk Score** that can be computed in aggregate, by neighborhood/per square mile.

In a groundbreaking series of articles in 2005, the Associated Press used the EPA data²¹⁴ to map the air quality risk scores for every neighborhood in the U.S. The AP mapped the EPA toxicity risk scores to socio-economic and racial factors for each neighborhood from the 2000 Census to determine the makeup of who was breathing the dirtiest air in America. The headlines across the country read, in some variation, that minorities suffer most from industrial pollution.²¹⁵

The results established important understandings about neighborhoods and toxicity, and the resulting snapshot of where and how factory pollution was impacting neighborhoods and people were deservedly much-discussed. These results are examples of beneficial uses of scores and what today would be called large datasets or "big data."

It is helpful that the EPA has a set of meaningful best practice guidelines for analyzing its data. The EPA has a Risk Characterization Handbook. It discusses EPA's use of risk characterizations in some detail. The EPA analysis of risk analysis is valuable here:

"Risk characterizations should clearly highlight both the confidence and the uncertainty associated with the risk assessment. For example, numerical risk estimates should always be accompanied by descriptive information carefully

²¹³ See US Postal Service, RFP, *USPS OIG Countermeasures and Performance Evaluations (CAPE) Data mining, predictive analytics, and Data Management Services* (Oct. 24, 2013), <<http://postalmag.com/datamining.pdf>>. See also: Shawn Hessinger – *Data Mining for Fraud at the US Postal Service*. Oct. 21, 2011.

<http://www.allanalytics.com/author.asp?section_id=1412&doc_id=234817>.

²¹⁴ See <<http://www.epa.gov/risk/health-risk.htm>>. From the AP article: "The scores aren't meant to measure the actual risks of getting sick or the actual exposure to toxic chemicals. Instead, they are designed to help screen for polluted areas that may need additional study of potential health problems, EPA said."

²¹⁵ David Pace, *More Blacks Live With Pollution*, Associated Press (Dec. 13, 2005), <http://onlinenews.ap.org/work/pollution/wrap.py?story=/linked_story/part1.html>. See also http://www.nbcnews.com/id/10452037/ns/us_news-environment/t/minorities-suffer-most-industrial-pollution/ The EPA uses toxic chemical air releases reported by factories to calculate risk for each square kilometer of the United States. The scores can be used to compare risks from long-term exposure to factory pollution from one area to another.

The scores are based on:

- _The amount of toxic pollution released by each factory.
- _The path the pollution takes as it spreads through the air.
- _The level of danger to humans posed by each different chemical released.
- _The number of males and females of different ages who live in the exposure paths.

selected to ensure an objective and balanced characterization of risk in risk assessment reports and regulatory documents.”²¹⁶

Further, the EPA created excellent documentation on how the analysis of its own data is to be used.²¹⁷ The documentation is for its own researchers, and is of note here because of its quality.

It stated, in part:

“The methods used for the analysis (including all models used, all data upon which the assessment is based, and all assumptions that have a significant impact upon the results) are to be documented and easily located in the report. This documentation is to include a discussion of the degree to which the data used are representative of the population under study. Also, this documentation is to include the names of the models and software used to generate the analysis. Sufficient information is to be provided to allow the results of the analysis to be independently reproduced.”²¹⁸

These recommendation could be readily applied to consumer scores and would increase fairness and transparency for many of the scores.

AIQ Green

IQ Analytics’s **AIQ Green** scoring tool “identifies prospects with a high propensity to show interest in environmentally friendly products.”²¹⁹ This is a marketing score.

Category - Other Consumer Scores

The borders of consumer scoring are not fully clear. We view our definition as a work in progress. It may be too broad or too limited. Individuals may be affected in some way by patterns of usage yet to develop or in ways that are hidden from public view. If so, then the definition for consumer scores may need to change. These are some other scores that we found but did not otherwise include here.

- **Brand Name Medicine Propensity Score**
- **Rx Online Search Propensity**
- **Casino Gaming Propensity Score**

²¹⁶ U.S. Environmental Protection Agency, *Risk Characterization Handbook. Prepared for the U.S. Environmental Protection Agency by EPA’s Science Policy Council At A5 (2000)*, <<http://www.epa.gov/spc/pdfs/rchandbk.pdf>>.

²¹⁷ U.S. Environmental Protection Agency, *Policy for Use of Probabilistic Analysis in Risk Assessment*, (May 15, 1997), <<http://www.epa.gov/spc/pdfs/probpol.pdf>>.

²¹⁸ Id at 2.

²¹⁹ <http://analytics-iq.com/download/AIQ_Green_PDF.pdf>.

- **Economic Stability Indicator Financial**
- **Prescriptions by Mail Propensity**
- **Underbanked Indicator**

Potential Scores

eBay feedback score: This score is public, standard-less, and subjective, but eBay publishes the scores, and vendors and purchasers use the scores to make decisions about each other. At the least, it falls at the lower end of the spectrum of concern. eBay has its own rules for monitoring the scores, with opportunities for appeal. Crowd-source scores like eBay's are likely to grow in popularity, and they may be worthy of more detailed analysis in the future as a separate class of scores.

SenderScore: This score comes from a database of email sender reputation maintained by a commercial company.²²⁰ According to the sponsoring website, the score derives from a proprietary Return Path algorithm, and represents a domain's overall performance against metrics important to both ISPs and recipients of email. This score represents the overall health of an email system as it appears to receiving systems.

Non-included Scores:

A host of proprietary Business-to-Business and business-focused scores are available. We acknowledge the considerable number of these scores, but exclude them on the basis of our definition of consumer scores.

Part IV: The History of Scoring: How the Credit Score and Consumer Scores Began, and Why it Is Relevant Today

The credit score is the progenitor of all consumer scores. The scoring story begins in 1941 with credit scoring, and continues today with the broadening of scoring to encompass consumer scoring in finance, insurance, health, and more.²²¹

The beginnings of consumer scoring

Today's broad array of activities in consumer scoring grew from credit scoring activities, which date back to 1941. It was then that a mathematician named David Durand used

²²⁰ See <<https://www.senderscore.org>>. *SenderScore – Search for Email Sender Reputation* (July 21, 2007).

²²¹ This discussion of the development of scoring is not intended to be comprehensive, but rather a look at the highlights of how scores have developed.

discriminant analysis to produce a scoring system to help predict risk in the granting of credit. His 1941 publication in the National Bureau of Economic Research is widely viewed as the first known account of using an analytical model to score credit risk.²²² One company in particular, Fair Isaac and Company, saw the business potential of scoring and developed scoring products for sale to business beginning in 1950. Fair Isaac was early to monetize the field, and as such the company is deeply intertwined with the popularization of scoring in business settings.

Credit scoring becomes entrenched

Over the course of the next few decades, mathematicians tweaked and refined models, tried new ones, compared models and combined scores, and in general pushed the entire body of research forward dramatically. These mathematical advances mirrored rapid advances in computing. The combination of computing and scoring allowed for increasingly rapid deployment of scoring in the credit environment. Credit scoring received a formal nod when the Equal Credit Opportunity Act (ECOA) cited credit scoring systems in one of its amendments.²²³ By 1979, William Fair of Fair Isaac estimated that between 20 and 30 percent of all consumer credit decisions were made by credit scoring.²²⁴ By the 80s and 90s, scores had been adopted widely in the U.S. and were spreading across the world, particularly in mature economies such as the U.S. and Europe. In 2006, a press release noted that FICO scores were used by U.S. lenders to make decisions about more than 75 percent of mortgage loan originations.²²⁵ During the creation, spread and use of the credit score from about 1941 to 2000, the score was largely secret to consumers. A decision by the FTC in the 1990s sealed the secrecy of the

²²² David Durand, *Risk Elements in Consumer Installment Financing, Study #8* (1941), National Bureau of Economic Research. Interestingly, a 1942 letter reporting on 1941 research activities written by National Bureau of Economic Research board member C. Reinold Noyes reveals that the bureau did not understand the profound implications of what they had published: “Undoubtedly the most important event to record in this report on the National Bureau of Economic Research for the year 1941 is the appearance of Simon Kuznets’ National Income and its Composition, 1919-38...” Durand’s work received a small, passing mention in the letter. *Report by our Representative on the Board of Directors of the National Bureau of Economic Research*, 32 *American Economic Review* 519-521 (1942) (Supplement, Papers and Proceedings of the 54th Annual Meeting of the American Economic Association).

²²³ 12 CFR § 202.6(b)(2)(ii). “To qualify as an empirically derived, demonstrably and statistically sound, credit scoring system, the system must be:

(i) Based on data that are derived from an empirical comparison of sample groups or the population of creditworthy and noncreditworthy applicants who applied for credit within a reasonable preceding period of time;

(ii) Developed for the purpose of evaluating the creditworthiness of applicants with respect to the legitimate business interests of the creditor utilizing the system (including, but not limited to, minimizing bad debt losses and operating expenses in accordance with the creditor’s business judgment);

(iii) Developed and validated using accepted statistical principles and methodology; and

(iv) Periodically revalidated by the use of appropriate statistical principles and methodology and adjusted as necessary to maintain predictive ability.”

²²⁴ Credit Card Redlining 1979, p. 183-184. As quoted in Noel Capon, *Credit scoring systems: a critical analysis*, 46 *Journal of Marketing* 82-91 (1982).

²²⁵ Fair Isaac, Press Release, *Fair, Isaac ‘Demystifies’ FICO Scores with list of Score Factors, Web-Based Explanation Service* (June 8, 2000) (“FICO scores are used by U.S. lenders to make billions of credit decisions each year, including more than 75 percent of mortgage loan originations.”).

credit score, but was later reversed. The history of how the credit score became public is an important precedent for current eligibility-related scores that are currently not available to consumers.

How the formerly secret credit score became available to the public

Scores were unknown to most consumers through the 50s, 60s, 70s, and 80s. Trickle of a score that was that could be used to deny a person credit but which was not revealed to consumers began to leak out slowly to some policymakers, particularly around the time ECOA passed. But scores had not entered the minds of most people.

In May 1990, the Federal Trade Commission wrote commentary indicating that risk scores (credit scores) did *not* have to be made available to consumers. But when scoring began to be used for mortgage lending in the mid 90s,²²⁶ many consumers finally began hearing about a credit score, most for the first time, and most when they were being turned down for a loan.²²⁷ A slow roar over the secrecy and opacity of the credit score began to build.

By the late 90s, the secrecy of credit scores and the fact that people could not see the underlying methodology or factors that went into the score or the range of the score to determine how the number should be interpreted was a full-blown policy issue. Beginning in 2000, events pushing toward increased credit score disclosure began to escalate, culminating in a rapid-fire series of events that eventually dismantled credit score secrecy and non-disclosure.

It is fair to say that a good deal of the escalation of events began when E-Loan, an Internet lender, took the extraordinary step of making credit scores public in February 2000 via a web site.²²⁸ The scores were free, and the word got out quickly to consumers. In one month, the site attracted more than 25,000 customers, and a lot of attention.²²⁹ The web site was shut down after six weeks; Fair Isaac, at that time, had a rule prohibiting the disclosure of FICO scores to consumers unless they were turned down for a loan.²³⁰ But although the site was been shut down, consumer appetite for their scores had been whetted. This incident was a tipping point due to how it popularized the score issue among consumers.

²²⁶ In 1995 Freddie Mac and Fannie Mae endorsed the use of credit scores as part of the mortgage underwriting process. This had a substantial impact on the use of credit scores in the mortgage loan industry. See for example Kenneth Harney, *The Nation's Housing Lenders might rely more on credit scores*, *The Patriot Ledger* (July 21 1995).

²²⁷ See for example, comments of Peter L. McCorkell, Senior Counsel to Wells Fargo, to the Federal Trade Commission, August 16, 2004 in response to FACT Act Scores Study.

²²⁸ A good first-hand account of the E-Trade web site incident may be found in an E-Loan press release: E-LOAN, Inc., *A full credit score disclosure pioneer, calls for national legislation; New credit score disclosure law is a giant step forward for California consumers, but consumers everywhere else in America remain in the dark* (June 27, 2001).

²²⁹ Brian Angell, *A Score to Settle: Consumer demand is high for credit scores. What's the holdup?* *American Banker-Bond Buyer*(August 2001).

²³⁰ *Id.*

That same month, on February 22, 2000, California senator Liz Figueroa introduced SB 1607 which would give Californians access to their credit scores. Specifically, the bill required lenders to give customers a copy of the credit score obtained to solicit a loan or accept a loan application.²³¹ Bowing to the growing pressure, Fair Isaac began to release some information about the factors that were used in its credit scoring model, FICO, in June 2000,²³² but they did not release the actual score at that time. One of the arguments they made was that too much disclosure would allow manipulation of the score.²³³

Governor Grey Davis signed the credit score disclosure legislation in September 2000, and the law took effect July 1, 2001.²³⁴ An uncomfortable situation then arose for federal lawmakers: Californians were the only ones who had access to their credit score. It was a classic recipe for national legislation on credit score disclosure.

In 2002, the FTC reversed its 1990 decision and concluded that consumers should be able to see their credit scores. As of December 2004, the Fair Credit Reporting Act as modified by the Fair and Accurate Credit Transactions Act, or FACTA, ended score secrecy formally, and required consumer reporting agencies to provide consumers with more extensive credit score information, upon request.²³⁵ Also made available to the public was the context of the score (its numeric range), the date the score was created, some of the key factors that adversely affected the score.²³⁶ The Federal Trade Commission is required by FACTA to study various aspects of credit scoring, insurance scoring, disparities, modeling, and more.²³⁷ Much still remains unknown about scoring models, even those that fall under FACTA such as credit scoring models. The formulas, which are important in verifying many aspects of the scoring model, are still secret.

Ongoing disclosure challenges and other issues with consumer credit scores

During the FACTA process, a growing trend was captured via the public comment process, that is, that the use of credit scores was greatly expanding to other areas of

²³¹ See <http://info.sen.ca.gov/pub/99-00/bill/sen/sb_1601-1650/sb_1607_bill_20000222_introduced.html> for the bill as introduced, and <http://info.sen.ca.gov/pub/99-00/bill/sen/sb_1601-1650/sb_1607_bill_20000930_chaptered.html> for the final version of the bill as chaptered.

²³² Fair, Isaac 'Demystifies' FICO Scores with list of Score Factors, *Web-Based Explanation Service*, PR Newswire, June 8, 2000

²³³ Fair, Isaac, Freddie to offer credit info to consumers, *Credit Risk Management Report*, June 12, 2000. Vol. 10, No. 11. Regarding arguments for not releasing the score to consumers, see Bonnie Simnock, *Fair Isaac site offers credit score details*, *American Banker* –Bond Buyer Association, November 6, 2000.

²³⁴ C.A.R.'s Credit Scoring Bill Signed into law by Governor Gray Davis; Landmark Legislation Gives Consumers Access to Credit Scores. PR Newswire, October 2, 2000.

²³⁵ 15 U.S.C. § 1681 g(f).

²³⁶ For a guide to consumer rights granted under FACTA, see *Appendix F: Summary of Consumer Rights Under the Fair Credit Reporting Act*, FTC, November 19, 2004. <<http://www.ftc.gov/opa/2004/11/facta.shtm>>.

²³⁷ The Federal Trade Commission has already released some reports, which may be found at <<http://www.ftc.gov/>>.

business. One area of concern was the use of credit scores to determine homeowner and auto insurance rates. Some individuals who had good driving records, for example, all of a sudden, upon renewal, were receiving much higher insurance rates due to a weak credit score. This practice has been the subject of much discussion, study, consternation, and some lawmaking, with varying results. During this general period of time that FACTA was being debated, the crime of identity theft began to become known and understood. New laws regarding the setting of insurance rates by a credit score impacted by identity fraud have been percolating through the states now as a result.

Disclosure of credit scores is now a non-issue. But while credit scores have been made public, it is not so with all other consumer scores. Consumers who want to see their identity score, their Z score, or many other scores cannot. Consumers who inquire about scores, or even the existence of a possible score, are not always told whether or not a score is being used. Similar if not identical arguments are used today to keep some consumer scores secret as were used to keep the credit score secret. While the credit score and its use, has been regulated by FACTA and now also by Basel II, this is not so for the broadening range of consumer scores that are increasingly attaching themselves to consumers.

The heightened availability and almost complete lack of oversight and regulation of the newer consumer scores combined with almost complete opacity regarding consumer scores' (minus credit and some forms of insurance scores) models, factors, ranges, validation, bias, sample size, and so forth has created a swath of non-disclosure and secrecy that consumers are at this point largely unaware of.

Conclusion

Because consumers cannot see most of the new consumer scores, cannot know the factors underlying many of the scores, there is no real application of Fair Information Principles to many of the new and unregulated consumer scores. Consumers who do not know about the existence or use of consumers scores cannot have any say in who used the scores, or how. Scores affect the lives of consumers, but only with reform will consumers receive rights to protect their interests.

The data business is changing and is becoming much more sophisticated. Consumer scores are a significant way that this is happening. Consumer scoring has substantial potential to become a major policy issue as scores with unknown factors and unknown uses and unknown validity and unknown legal constraints move into broader use.

Secrecy, fairness of the factors, accuracy of the models, and the use of sensitive information are some of the key issues that must be addressed. It is exquisitely unlikely that self-regulation will solve all of the dilemmas consumer scoring introduces. However, we already have at least a partial model for what would constitute fair regulation from the history of the credit score. The protections consumers receive with respect to credit

scores need to be expanded to all consumer scoring, and the rules for credit scores may warrant some reexamination as well.

About this Report and Credits

Authors: Pam Dixon, Executive Director, World Privacy Forum and Robert Gellman, Privacy Consultant.

Editing, proofing, and reading: Tim Sparapani, Linda Ackerman, Nathan Good, Blake Hamilton.

Cover illustration: John Emerson of Backspace.

Publication date: April 2, 2014.

Published at: www.worldprivacyforum.org

Version: 1.2

URL for report: <http://www.worldprivacyforum.org/category/report-the-scoring-of-america/>

Future version and updates will be located at the above URL.

Background interviews for this report were conducted from a period of 2007 to 2014. The World Privacy Forum thanks the US Federal Trade Commission and the many companies, experts, analysts, and others we interviewed for this report.

Appendix A

Timeline: Highlights in Scoring

1941 David Durand publishes first account of the use of discriminant analysis to produce a scoring system for the use in granting credit.²³⁸

1950 Glueck, S. and Glueck, E.T. *Unraveling Juvenile Delinquency*, Cambridge: Harvard University Press. Develop a point scoring method.

1950 Stanford University researchers Bill Fair and Earl Isaac set up a new business in a San Rafael garage.²³⁹

1963 Myers and Gorgy (compared discriminant analysis and regression analysis)

1971 Orgler used regression analysis to create a behavioral score to evaluate outstanding loans based on past performance. (Orgler, Yair E. *A Credit Scoring Model for Commercial Loans*, *Journal of Money, Credit and Banking*, 2 November 1970, 435-45.)

1977, 1978 Eisenbeis presented assessment of the use of discriminant analysis in business, finance, and economics in general.²⁴⁰

1980 Wiginton one of first published accounts of logistic regression applied to credit scoring.

1981 Grablowsky and Talley (1981) compared linear discriminant analysis and probit analysis by using data from a large U.S. Midwestern retail chain.²⁴¹

1984 Breiman et al publish on recursive partitioning, or decision trees.

1991 Safavian and Landgrebe – survey of recursive partitioning in scoring, including artificial intelligence.

1992 Blackwell and Sykes described the use of behavioral scoring to determine credit limits.

1993 Leonard described an expert system for detecting fraudulent use of credit cards.

²³⁸ D.J. Hand, *Statistical Classification Methods in Consumer Credit Scoring: A Review*, 160 *Journal of the Royal Statistical Society* 532 (1997).

²³⁹ Edmund Sanders, *California Firm that developed FICO credit scores is still sailing*, *Orange County Register* (September 26 1997).

²⁴⁰ D.J. Hand, *Statistical Classification Methods in Consumer Credit Scoring: A Review*, 160 *Journal of the Royal Statistical Society* 532 (1997).

²⁴¹ *Id.*

- 1994** Rosenberg and Gleit-- applications of neural networks to corporate credit decisions and fraud detection.
- 1995** Henley described a fraud score card built by a linear regression analysis model.
- 1995** FAIR ISAAC introduces its first model of a Small Business Credit Score.²⁴²
- 1996** Henley and Hand develop an adaptive metric “nearest neighbor” method for credit scoring.
- 2000** In February California senator Liz Figueroa introduces legislation to allow consumers to see their credit scores. The bill is signed into law in September.
- 2001** July 1: Californians have the legal right to view their credit scores.
- 2003** The FACT Act is enacted December 4, consumers nationwide given the legal right to view credit scores, score range, and some additional rights.
- 2004** FTC requests public comments on the use of credit scores in setting insurance rates.
- 2007** Antibody scoring models in development, Theodore Crooks.
- 2008** Klout is born, and social scoring becomes a reality.
- 2011** FICO launches Medical Adherence Score, one of the first major medical consumer scores. The score does not have to rely on medical files for its predictions.
- 2012** Charles Duhigg breaks the existence of the Target *pregnancy predictor score* in a New York Times feature article; raises awareness of predictive analytics and use of masses of factors in scoring algorithms.
- 2013** HHS creates The Health Risk Score for individuals using the Affordable Care Act (Obamacare) program.
- 2014** FTC holds alternative scoring models conference, first high-level attention to non-FCRA scores.
- 2018** The target date for phasing out the Health Risk Score. (Planned).

²⁴² Allen N. Berger, W. Scott Frame. *Small Business Credit Scoring and Credit Availability*, 45 *Journal of Small Business Management* (2007).

Appendix B

Score Taxonomy

In minds of consumers, there is just one score, the credit score. But the credit score is just one final outcropping of a layered and complex taxonomy of scoring. This taxonomy can assist consumers in seeing the full range and depth of scoring activities that exist, and may impact them.

This taxonomy is also important in understanding the scores this report focuses on. This report is focused on *consumer scores* that are used for *consumer purposes*. Or to use the taxonomic language, consumer scores derived from formal predictive models and used for consumer-related purposes, that is, used in a way that impacts a non-clinical (non-medical) decision about a consumer or a group of consumers.

I. Predictive Statistical Models

II. Formal Scoring Models

III. Consumer Scoring Models

IV. Consumer Scoring Model Type (application, behavioral, or combined)

V. Consumer Scoring Function: the broad function of the score card, as follows:

Propensity score cards: will the consumer, for example, default, what is the propensity of a certain result. Credit scoring is a propensity scoring function. Health Scoring is a propensity function if it falls under the full taxonomy preceding this point.

Response score cards: will the consumer respond to a direct marketing offer

Usage score cards: will the consumer use the credit (or other) product if given the product

Attrition score cards: will the consumer continue with the lender, especially if there is some special offer available for an introductory period only.

Customer profit scoring score cards: estimates the total profitability of the customer to the lender

Product profit score cards: seeks to estimate the profit the lender makes on this product from the customer²⁴³

VI. Source of the Score Model and score (Generic, custom, or vendor supplied score)

VII. The Specific Type of Score (fraud, credit, etc.) Here, the term credit refers to the broad type of score.

VIII. Application of Score (what purpose is the score used for)

Consumer-related: test: does the score impact a decision about an individual consumer or a group of consumers?

Research-related: (esp. Health research)²⁴⁴ test: is the score used to primarily to understand or explain a process or a disease and never used to make a decision about an individual consumer beyond a clinical medical decision? (If a financial or risk decision is taken, then the score becomes a consumer score, not just a clinical score.)

IX. Actual Scores (This includes all specific scores resulting from the taxonomy, Z score, Falcon score, FICO score, etc.) *Note: this report is focused on Consumer-related scores, or scores that are used for consumer purposes. If at any point a pure research-related score is used in a consumer score model as a predictive factor and the resulting final score is used for consumer purposes, the final score would be considered a blended consumer score and would be included in the consumer category.* See Taxonomy step VII.

²⁴³ The discussion of the function of score cards in this segment is derived closely from LC Thomas, RW Oliver, DJ Hand, *A Survey of Issues in Consumer Credit Modeling Research*, 56 Journal of the Operational Research Society(2005).

²⁴⁴ There are a number of health-related scores in particular that are originally created solely for research purposes, particularly public health research. Scores of this type are not considered in this report. But if the research score is later combined with a consumer score and is used for consumer purposes, then that score would be a blended consumer score and would be considered in the report. Because of how scoring models operate, it is possible that some pure research-related health scores later became part of some consumer-related scores. Any characteristic, such as a health score, can be input into a consumer scoring model. In this way, a pure research score can *contribute to* a consumer-related score. It would be nearly impossible to determine how many health scores originally created for research purposes have or are being used in this way. Scoring models generally do not reveal formulae to this level of specificity.

**RESPONSE TO WRITTEN QUESTION OF SENATOR MENENDEZ
FROM ALICIA PUENTE CACKLEY**

Q.1. Can data brokers legally compile, aggregate, or sell data that has been acquired through an illegal hack?

A.1. GAO has not conducted work to determine the extent to which data brokers are collecting, compiling, aggregating, or selling data that was acquired through illegal hacks, or the legality of such actions. However, we reported in March 2019 (GAO-19-230) that, except in certain circumstances, companies are generally not required to be transparent about the consumer data they hold or how they collect, maintain, use, and secure these data. Further, we recommended more than a decade ago that Congress consider whether to expand more broadly the class of entities explicitly required to safeguard sensitive personal information, including considering whether information resellers should be required to safeguard all sensitive personal information they hold (GAO-06-674). Even still, statutes like the Computer Fraud and Abuse Act provide some protection by making the knowing unauthorized access of computers a crime, and FTC has used its enforcement authority to address some instances of unfair or deceptive behavior in the sale of information or its use in advertising. Notably, in 2014, FTC alleged that a data broker sold hundreds of thousands of loan applications that contained sensitive data, including consumers' names, addresses, phone numbers, employers, Social Security numbers, and bank account numbers (including routing numbers) to entities that it knew had no legitimate need for such data. FTC alleged that, as a result, at least one of those purchasers used the information to withdraw millions of dollars from consumers' accounts without their authorization. FTC and the involved companies settled this case in 2016, which included monetary judgments and a permanent ban for all defendants on selling, transferring, or otherwise disclosing consumers' sensitive personal information.

**RESPONSES TO WRITTEN QUESTIONS OF SENATOR WARREN
FROM ALICIA PUENTE CACKLEY**

Q.1. In response to the Equifax data breach, I opened an investigation into the causes, impacts, and response to the exposure of personal data of nearly 150 million Americans. Equifax and other credit reporting agencies collect consumer data without permission, and consumers have no way to prevent their data from being collected and held by private companies. My investigation found that Equifax failed to adopt standard cybersecurity measures, in large part because Federal law incentivizes pursuit of profits over the protection of sensitive data.

Your written testimony notes, “[The Fair Credit Reporting Act (FCRA)] protects the security and confidentiality of personal

information collected or used to help make decisions about individuals' eligibility for credit, insurance or employment. FCRA limits resellers' use and distribution of personal data."¹ This law, however, is not specifically designed to address cybersecurity threats.² In your view, how should Federal regulators address this gap in the oversight and enforcement of privacy safeguards?

A.1. There is currently no comprehensive Federal statute to address consumer privacy, which is one reason that Federal regulators are limited in their ability to address potential gaps in current law. In a 2013 report (GAO-13-663), we recommended that Congress consider updating the consumer privacy framework to reflect the effects of changes in technology and the marketplace—changes that have included new and greater cybersecurity threats. Criteria for developing such a framework could include the Fair Information Practice Principles—and a key principle is that personal information should be protected with reasonable security safeguards against risk such as loss or unauthorized access, destruction, modification, or disclosure.

Q.1.a. How would legislation to establish and provide Federal authority and resources to monitor data security practices of credit reporting agencies and data brokers benefit consumers?

A.1.a. Stronger Federal oversight of data security practices could help to ensure that consumer reporting agencies and data brokers better safeguard all sensitive personal information, which could protect consumers from identity theft and other effects of data breaches. To strengthen such oversight, our February 2019 report on consumer reporting agencies (GAO-19-196) recommended that Congress consider giving FTC civil penalty authority to enforce Gramm-Leach-Bliley Act's (GLBA) safeguarding provisions. In addition, we have long held that data protections should apply broadly. For example, in 2006 (GAO-06-674), we noted that much of the personal information maintained by information resellers that did not fall under FCRA or GLBA was not necessarily required by Federal law to be safeguarded, even when the information is sensitive and subject to misuse by identity thieves. We therefore recommended that Congress consider requiring information resellers to safeguard all sensitive personal information they hold.

Q.1.b. In your view, would legislation to impose strict liability penalties for breaches involving consumer data at credit reporting agencies and data brokerages lead to improvements in consumer data security? Would consumers benefit if such penalties were imposed on data brokers?

A.1.b. GAO has not reviewed the issue of how strict liability penalties for breaches involving consumer data at consumer reporting agencies and other information resellers would affect consumer data security or consumers. However, we have highlighted the importance of providing agencies with civil penalty authority, which can also be a strong enforcement tool. In our February 2019 report

¹ Written testimony of Alicia Cackley to the U.S. Senate Committee on Banking, Housing, and Urban Affairs, June 11, 2019, <https://www.banking.senate.gov/imo/media/doc/Cackley%20Testimony%206-11-19.pdf>.

² Letter from Acting Federal Trade Commission Chair Maureen Ohlhausen to Senator Elizabeth Warren, October 3, 2017.

on oversight of consumer reporting agencies (GAO-19-196), we recommended that Congress consider giving FTC civil penalty authority to enforce GLBA's safeguarding provisions. Currently, to obtain monetary redress for these violations, FTC must identify affected consumers and any monetary harm they may have experienced. However, harm resulting from privacy and security violations (such as a data breach) can be difficult to measure and can occur years in the future, making it difficult to trace a particular harm to a specific breach. FTC currently lacks a practical enforcement tool for imposing civil money penalties that could help to deter companies from violating data security provisions of GLBA and its implementing regulations. Such deterrence could benefit consumers because companies may be motivated to develop stronger procedures for data security that would protect consumer data from theft and security breaches.

Q.2. Despite there being laws in place to regulate consumer credit reporting, your written testimony notes that there are “no Federal laws designed specifically to address all the products sold and information maintained by [data brokers].”³ Given the limited ability of individuals to access, control, and correct their personal data, as well as the limited legal framework to regulate data brokers, would the inadequacy of current laws be addressed by regulating data brokers under the Fair Credit Reporting Act?

A.2. GAO has not conducted work specifically assessing the advantages and disadvantages of regulating all information resellers (data brokers) under the Fair Credit Reporting Act. In 2013 (GAO-13-663), we noted gaps in Federal privacy law—including that it did not always cover consumer information used by information resellers for marketing purposes or other uses not covered by provisions of the Fair Credit Reporting Act. We recommended that Congress consider strengthening the consumer privacy framework to address these gaps, but we did not recommend a specific regulatory scheme for doing so.

Q.2.a. Credit reporting agencies make billions of dollars collecting and selling information about consumers, but consumers have little ability to control how their personal information is collected and used by these agencies. How would legislation to give consumers more control over personal financial data and to create a uniform, Federal process for obtaining and lifting credit freezes benefit consumers? Would consumers benefit if such legislation also applied to currently unregulated parts of the industry, such as data brokerages?

A.2.a. While consumers currently do not have a uniform, Federal process for credit freezes, the Economic Growth, Regulatory Relief, and Consumer Protection Act required the three nationwide consumer reporting agencies to place and lift freezes at no cost to the consumer. Freezes must be placed within 1 business day, and lifted within 1 hour, of receiving a telephone or electronic request. However, consumers must contact each of the three agencies individually and request the freeze. Consumers obtain a PIN from each

³Written testimony of Alicia Cackley to the U.S. Senate Committee on Banking, Housing, and Urban Affairs, June 11, 2019, <https://www.banking.senate.gov/imo/media/doc/Cackley%20Testimony%206-11-19.pdf>.

company, which enables them to lift or remove a freeze at a later date. Before the 2018 Act, consumers typically had to pay \$5–\$10 per agency to place a credit freeze. In our March 2019 report (GAO–19–230) on data breaches and limitations of identity theft services, some experts had noted cost and inconvenience as some of the limitations to a credit freeze.⁴ The new law addresses these concerns to some degree by making credit freezes free and requiring these consumer reporting agencies to lift freezes expeditiously on request.

In terms of less-regulated segments of the information reseller industry—most notably, companies or data not covered by FCRA—our 2013 recommendation to Congress (GAO–13–663) suggested updating the consumer privacy framework in ways that could address this gap. In particular, two key elements we said such legislation should consider are (1) the adequacy of consumers’ ability to access, correct, and control their personal information in circumstances beyond those currently accorded under FCRA; and (2) whether there should be additional controls on the types of personal or sensitive information that may or may not be collected and shared.

**RESPONSES TO WRITTEN QUESTIONS OF SENATOR SCHATZ
FROM ALICIA PUENTE CACKLEY**

Q.1. Are data sets collected by data brokers getting into the blood stream of credit, employment, and housing decision making, in a way that evades FCRA?

A.1. GAO has not conducted work to determine the extent to which information collected by data brokers is being used to make credit, employment, and housing decisions in ways that do not comply with the Fair Credit Reporting Act (FCRA). However, in a 2018 report on financial technology (GAO–19–111), we evaluated consumer protection issues related to FinTech lenders’ use of alternative data—that is, data not traditionally used by the national consumer reporting agencies in calculating a credit score—to make loan decisions.¹ Five of the 11 FinTech lenders we interviewed said they used alternative data to supplement traditional data when making a credit decision, with one using it exclusively. These lenders told us that they obtain the data from borrowers, data aggregators, national databases, or other sources. Consumers may face risk of harm due to inaccurate credit assessments when FinTech lenders use alternative data to underwrite loans. Inaccurate data or models could classify borrowers as higher credit risks than they actually are. This could result in those borrowers paying unnecessarily high interest rates (and increase risk of default), or it could result in creditworthy borrowers being denied credit. While FCRA requires that borrowers have an opportunity to check and correct inaccuracies in their credit reports, borrowers could face challenges checking and correcting alternative data, which typically are not shown in credit reports. Further, it may not be transparent to consumers

⁴ GAO, *Data Breaches: Range of Consumer Risks Highlights Limitations of Identity Theft Services*, GAO–19–230 (Washington, DC: March 27, 2019).

¹ GAO, *Financial Technology: Agencies Should Provide Clarification on Lenders’ Use of Alternative Data*, GAO–19–111 (Washington, DC: Dec. 19, 2018).

and regulators what specific information alternative credit-scoring systems use, how such use affects consumers, and what consumers might do to improve credit access and pricing.

Q.2. Under current law, do companies that collect and sell information about consumers have any duty to consumers about how that information will be used?

A.2. The legal obligation to consumers related to the use of consumer information varies based on the content and context of that use. No comprehensive Federal privacy law governs the collection, use, and sale of personal information by private-sector companies. While there are Federal laws addressing commercial privacy issues, they are generally narrowly tailored to specific purposes, situations, types of information, or sectors or entities—such as data related to financial transactions, personal health, and eligibility for credit. These laws include provisions that can restrict how certain companies use consumer information they collect or sell—by, for example, limiting the disclosure of certain types of information to a third party without an individual’s consent.

For example, FCRA—which applies to personal information used for certain eligibility determinations—gives consumers the right, among other things, to opt out of allowing consumer reporting agencies to share their personal information with third parties for prescreened marketing offers. Another example is the Gramm-Leach-Bliley Act, which imposes certain sharing and disclosure restrictions on financial institutions or entities that receive nonpublic personal information from such institutions. For instance, a third party that receives nonpublic personal information from a financial institution to process consumers’ account transactions generally may not use or resell the information for marketing purposes. Similarly, other laws, such as the Health Insurance Portability and Accountability Act of 1996 and the Children’s Online Privacy Protection Act of 1998, also restrict how consumer information can be used, but they too apply narrowly to specific entities or types of information.

Q.3. If consumers are discriminated against or harmed because of how that data is used, who is responsible?

A.3. While the responsible party, if any, is going to vary based on the facts and circumstances of each case, our January 2019 report on internet privacy (GAO–19–52) examined some examples of Federal Trade Commission (FTC) enforcement actions taken against companies related to internet privacy.² In these enforcement actions FTC alleged each company’s practices were unfair, deceptive, a violation of the Children’s Online Privacy Protection Act (COPPA), a violation of a settlement agreement, or a combination of these reasons. In that report we found that between July 1, 2008, and June 30, 2018, FTC filed 101 internet privacy enforcement actions, 15 of which included COPPA enforcement actions against a variety of companies. Of the 101 internet privacy actions, we reported that 51 involved internet content providers, 21 involved software developers, 12 involved the sale of information or

²GAO, *Internet Privacy: Additional Federal Authority Could Enhance Consumer Protection and Provide Flexibility*, GAO–19–52 (Washington, DC: Jan. 15, 2019).

its use in advertising, 5 involved manufacturers, 1 involved an internet service provider, and 11 involved a variety of different products, such as those provided by rent-to-own companies or certification services. In nearly all 101 cases companies settled with FTC, which required the companies to make changes in their policies or practices as part of the settlement. We reported that during that 10-year period, FTC leveled civil penalties against 15 companies for alleged violations of COPPA regulations totaling \$12.7 million. These civil penalties ranged from \$50,000 to \$4 million with an average amount of \$847,333. We also reported that FTC can seek to compel companies to provide monetary relief to those they have harmed and during that period FTC levied civil penalties against companies for violations of consent decrees or obtained monetary relief to consumers from companies for a total of \$136.1 million. These payment orders ranged from \$200,000 to \$104.5 million and the average amount was \$17 million.³

Q.4. If a data broker is breached and a consumer suffers harm from identity theft, who is liable?

A.4. As with the broader case of consumer harm, liability in identity theft cases is a matter of the facts and circumstances of each individual case. GAO hasn't examined liability specifically with regard to data breaches. However, as noted above, in our January 2019 report (GAO-19-52) we found that 12 of FTC's internet privacy enforcement actions between July 1, 2008, and June 30, 2018, involved the sale of information or its use in advertising. Notably, in 2014, FTC alleged that a data broker sold hundreds of thousands of loan applications that contained sensitive data, including consumers' names, addresses, phone numbers, employers, Social Security numbers, and bank account numbers, including the bank routing numbers, to entities that it knew had no legitimate need for such data.⁴ FTC alleged that, as a result, at least one of those purchasers used the information to withdraw millions of dollars from consumers' accounts without their authorization. FTC and the involved companies settled this case in 2016, which included monetary judgments and a permanent ban for all defendants on selling, transferring, or otherwise disclosing consumers' sensitive personal information without consent.⁵

Q.5. Do you think Federal law should require companies that collect and use consumer data to take reasonable steps to prevent unwanted disclosures of data and not use data to the detriment of those consumers?

³However, this sum does not represent the amount of money that consumers actually received or that was forfeited to the U.S. Treasury. In some cases, including the payment order for \$104.5 million, FTC suspended the judgment because of the defendants' inability to pay.

⁴See Complaint, Federal Trade Commission v. Sitesearch Corporation, dba LeapLab *et al.*, No. 2:14-cv-02750-NVW (D. Ariz. Dec. 22, 2014), <https://www.ftc.gov/system/files/documents/cases/141223leaplabcmt.pdf>; see also Complaint, Federal Trade Commission v. Ideal Financial Solutions, Inc., *et al.*, No. 2:13-cv-00143-MMD-GWF (D. Nev. Jan. 28, 2013), <https://www.ftc.gov/sites/default/files/documents/cases/2013/02/130220ifscmt.pdf>.

⁵See Stipulated Final Order for Permanent Injunction and Settlement of Claims, Federal Trade Commission v. Sitesearch Corporation, dba LeapLab, a Nevada corporation; *et al.*, No. CV-14-02750-PHX-NVW (D. Ariz., Feb. 5, 2016), https://www.ftc.gov/system/files/documents/cases/160218leaplalorder_0.pdf; see also Order Granting in Part Motion for Summary Judgment and Motion for Default Judgment, Entering Final Judgment, and Closing Case, Federal Trade Commission v. Ideal Financial Solutions, Inc., *et al.*, No. 2:13-cv-00143-JAD-GWF (D. Nev. Feb. 23, 2016), <https://www.ftc.gov/system/files/documents/cases/160309idealfinancialorder.pdf>.

A.5. While GAO has not taken a position on whether Federal law should require all companies to take measures to protect all consumer data and to not use that data to the detriment of consumers, we have previously recommended in GAO-13-663 that Congress consider strengthening the current consumer privacy framework. In making our recommendation, we noted that current privacy law is not always aligned with the Fair Information Practice Principles. One of these principles directly addresses unwanted disclosures: “security safeguards” is the principle that personal information should be protected with reasonable security safeguards against risks such as loss or unauthorized access, destruction, use, modification, or disclosure. Other principles address not using a consumer’s data to the detriment of that consumer: for example, “use limitation” is the principle that data should not be used for other than a specified purpose without consent of the individual or legal authority.

In addition, GAO has made a number of specific recommendations for modifying Federal law that relate to protecting consumer data held by private companies.

- In May 2019 (GAO-19-340), we recommended that Congress consider providing the Internal Revenue Service (IRS) with explicit authority to establish security requirements for paid tax return preparers’ and Authorized e-file Providers’ systems.⁶
- In February 2019 (GAO-19-196), we recommended that Congress consider providing the Federal Trade Commission with civil penalty authority for the safeguarding provisions of the Gramm-Leach-Bliley Act, which would help the agency act against data security violations by financial institutions.⁷
- In June 2006 (GAO-06-674), we recommended that Congress consider requiring information resellers to safeguard all sensitive personal information they hold—not just information covered under the safeguarding provisions of the Fair Credit Reporting Act and Gramm-Leach-Bliley Act.⁸

RESPONSES TO WRITTEN QUESTIONS OF SENATOR CORTEZ MASTO FROM ALICIA PUENTE CACKLEY

Q.1. What does it mean for financial markets now that FINRA can essentially predict and decide in real time, or near real-time investor behavior? What does it mean for other financial and technical sectors?

A.1. In a March 2018 GAO forum (GAO-18-142SP), we highlighted the use of artificial intelligence (AI) in financial services, including market surveillance oversight activities.¹ At the time of the forum, the Financial Industry Regulatory Authority (FINRA) was developing a prototype AI-based system, called the Dynamic Surveil-

⁶GAO, *Taxpayer Information: IRS Needs to Improve Oversight of Third-Party Cybersecurity Practices*, GAO-19-340 (Washington, DC: May 9, 2019).

⁷GAO, *Consumer Data Protection: Actions Needed to Strengthen Oversight of Consumer Reporting Agencies*, GAO-19-196 (Washington, DC: Feb. 21, 2019).

⁸GAO, *Personal Information: Key Federal Privacy Laws Do Not Require Information Resellers to Safeguard All Sensitive Data*, GAO-06-674 (Washington, DC: June 26, 2006).

¹GAO, *Technology Assessment: Artificial Intelligence: Emerging Opportunities, Challenges, and Implications*, GAO-18-142SP (Washington, DC: March 28, 2018).

lance Platform, which used supervised machine learning capabilities to learn and detect different patterns of market anomalies to enhance the ability to detect instances of potential illegal manipulation of the securities and options markets. With new AI-based tools, as well as future data enhancements to increase the visibility of each trading transaction offered by a new consolidated audit trail being developed, regulators were hopeful that employing machine learning capabilities will help identify future intentional manipulation of the markets.

During the forum, industry participants and regulators highlighted both benefits and challenges offered by the use of AI tools in the marketplace. Benefits included enhanced surveillance monitoring (by an entity internally as well as externally by financial regulators) and tools to better detect and prevent improper market conduct and enforce existing laws and regulations in the marketplace. At the same time, challenges and growing pains associated with technological advances of AI-based tools also exist. For instance, banking regulators and other industry observers said that banks are reluctant to move quickly in implementing AI tools for lending operations due to concerns about meeting requirements under existing laws and regulations (*e.g.*, requirements stemming from fair lending laws that prohibit discriminatory practices on lending, whether intentional or not, based on race, gender, color, religion, national origin, marital status, or age).

Q.2. What are some of the gaps in currently existing law with respect to how enforcement agencies deal with this multitude of laws and what should we be thinking about in the Banking Committee as we prepare to potentially consider broader privacy legislation drafted by the Commerce Committee?

A.2. Many existing privacy statutes in the United States were developed before the advent of many current technologies and before companies were collecting and sharing such vast quantities of consumer personal information. We reported in a 2013 review of information resellers (GAO-13-663) that we believed that gaps exist in the current statutory privacy framework, and we believe this remains true today.² In particular, the current framework does not fully address changes in technology and marketplace practices that fundamentally have altered the nature and extent to which personal information is being shared with third parties. Moreover, while current laws protect privacy interests in specific sectors and for specific uses, consumers generally have little control over how their information is collected, used, and shared with third parties for marketing purposes.

If Congress considers broader privacy legislation to strengthen the consumer privacy framework, we believe that among the issues that should be considered are:

- the adequacy of consumers' ability to access, correct, and control their personal information in circumstances beyond those currently accorded under the Fair Credit Reporting Act;

² GAO, *Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace*, GAO-13-663 (Washington, DC: Sept. 25, 2013).

- whether there should be additional controls on the types of personal or sensitive information that may or may not be collected and shared;
- changes needed, if any, in the permitted sources and methods for data collection; and
- privacy controls related to new technologies, such as web tracking and mobile devices.

At the same time, we recognize that different legislative approaches to improving privacy involve tradeoffs and believe that any strengthened privacy framework should also seek not to unduly inhibit the benefits to consumers, commerce, and innovation that data sharing can accord.

**RESPONSES TO WRITTEN QUESTIONS OF SENATOR
MENENDEZ FROM PAM DIXON**

Q.1. In the hearing, you stated it is of “grave concern” that data not covered by HIPAA is ending up in the hands of data brokers.

Q.1.a. Are medical billing companies selling non-HIPAA data to brokers?

A.1.a. We are most familiar with third-party medical billing companies that inappropriately use HIPAA data for fraudulent purposes. We are less familiar with medical billing companies selling non-HIPAA data. The risk of HIPAA data misuses, however, is significant by itself.

One major modality medical billing companies have used is to fraudulently use HIPAA data to bill Medicare/Medicaid directly, apart from original billing tasks. In another model, medical billers may simply overcharge for services. These activities are a form of medical identity theft, and typically results in fraudulent changes to the health file. The Office of the Inspector General wrote a brief but seminal report about billing companies in March, 2000.¹ In the report, the OIG noted the complex problems with medical billing, including problems with transparency and auditing. There continue to be many cases relating directly to problems with medical billers.²

OIG has established voluntary compliance guidance for medical billing, but the guidance dates from 1998.³ HBMA has established medical billing credentialing and training for companies, which currently functions as a set of best practices.⁴ We believe much more can be done here, for example, we would like to see many more credentialed members of HBMA, and more encouragement from Congress for either certification or some additional form of oversight for medical billing companies.

Medical billing deserves an update from OIG and from Congress. It would be a particularly productive area to update.

¹ See <https://oig.hhs.gov/oei/reports/oei-05-99-00100.pdf>.

² See, for example, the 2015 Medicaid case: <https://www.justice.gov/usao-wdnc/pr/owner-medical-billing-company-indicted-health-care-fraud-and-aggravated-identity-theft>; and the more recent case from July 2019: <https://www.justice.gov/usao-sdoh/pr/medical-billing-company-owner-sentenced-prison-health-care-fraud>.

³ See <https://www.oig.hhs.gov/fraud/docs/complianceguidance/thirdparty.pdf>.

⁴ See <https://www.hbma.org/content/certification/hbma-compliance-accreditation-program/accredited-companies>.

Q.1.b. How pervasive of a problem is medical identity theft?

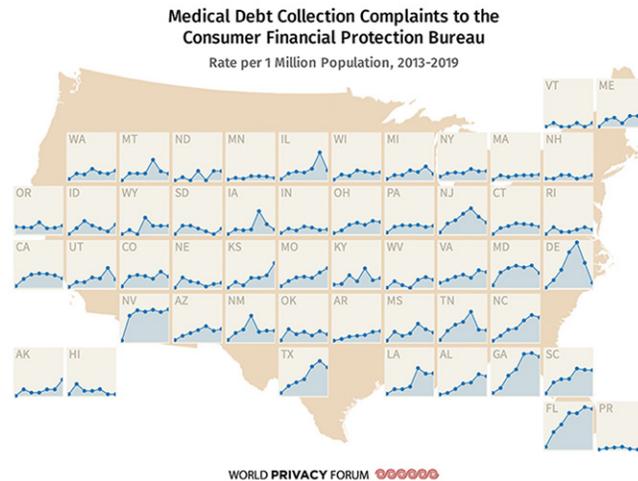
A.1.b. We first identified medical identity theft as a problem in testimony to NCVHS in 2005, then wrote the first known report on the topic in 2006.⁵ We continue to research the field, and can now give you precise quantifications of the problem, State by State.

In January 2020 we will publish our *State of Medical Identity Theft report*, which follows our 2017 *Geography of Medical Identity Theft report*.⁶ We published an interactive data visualization of medical identity theft in the United States, by State that accompanied the report.⁷

In our 2020 report, we again have found pervasive incidents of medical identity theft across the United States, with some States showing more serious problems. We have included two screen shots of our pre-publication data to give you a visual view of the numbers. The numbers from 2013–2018 are final, and the numbers for 2019 run to Dec. 1. Our January report with the final 2019 numbers will have nearly identical statistics as the screenshots attached here.

As you can see from the data, medical identity theft is now present in all States. This data has been adjusted per population rate. We note persistent patterns of medical identity theft through the southeastern corridor, with hot spots in Texas, Georgia, Florida, South Carolina, and Nevada. We note that New Jersey was a hot spot, but has seen improvement in recent years, as has Illinois.

Medical Identity Theft complaints, 2013–2019:

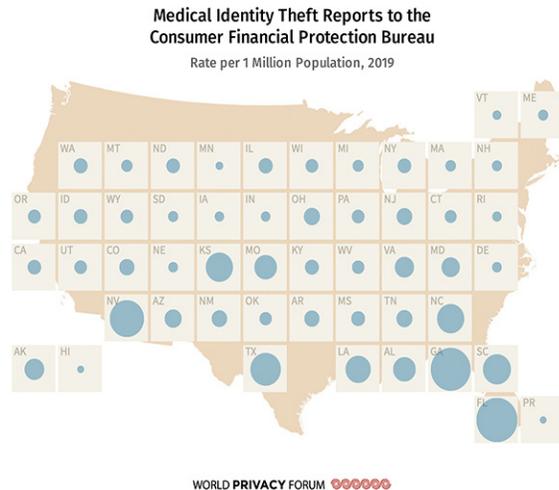


⁵ See <https://www.worldprivacyforum.org/2006/05/report-medical-identity-theft-the-information-crime-that-can-kill-you/>.

⁶ See <https://www.worldprivacyforum.org/2017/12/new-report-the-geography-of-medical-identity-theft/>.

⁷ World Privacy Forum, *Medical Identity Theft Mapped by State: Data Visualization*. <https://www.worldprivacyforum.org/2017/12/medical-identity-theft-reports-to-the-consume-financial-protection-bureau/>.

Medical Identity Theft Complaints, 2019
Rate per 1 Million Population



Q.1.c. When patients are victims of medical identity theft, what recourse do they have to correct errors on their files?

A.1.c. Patients can use their rights under the FCRA to correct the financial aspects of their healthcare provider records. However, patients do not have commensurate rights under HIPAA to delete or correct errors in their medical records. Under HIPAA, patients can request the addition of an amendment to their records. An amendment request does not have to be honored by the healthcare provider. Amendment requests do not mandate the removal or correction of information, they simply allow consumers to dispute the information. Healthcare providers typically do not delete information in a health file.

There are some workarounds. A responsible healthcare provider can remove inaccurate information from a patient's record and leave only a numeric cross reference to the information introduced by the fraudulent activities. For example, if a patient was fraudulently billed for having cancer, the patient's health record would reflect that error. The healthcare provider could remove that and other related information introduced by the fraudulent activity, and sequester it into a new "John or Jane Doe" file, leaving only a numeric cross reference. This is one of the several best practices for handling errors in records resulting from medical identity theft.

However—this issue needs to be addressed legislatively so that there is a national standard for how to assist victims in correcting their health records after medical identity theft has introduced errors. Ultimately, a national-level solution will improve data for the entire health system as well as help victims. This is a gap that needs to be addressed.

Q.1.d. Typically, how often do these cases go unresolved?

A.1.d. Anecdotally, many cases go unresolved. We are aware of many patients over the years who have chosen to ignore the problems, because they simply could not resolve them. Part of the way we know this is from ongoing phone calls over the years since the first publication of our report in 2006. We have found that there is a high degree of variability in healthcare providers' responses. We believe a uniform procedure for correction could improve outcomes for victims and providers alike.

Q.2. You also mentioned that we need to do more to ensure that consumers are notified when a data broker suffers a breach that exposes consumers' sensitive information.

Q.2.a. Given that data brokers often do not have a direct relationship with consumers, what do you think is the best way for Congress to ensure that consumers are notified when their data is exposed by a breach?

A.2.a. Data brokers should have specific requirements to make breach notification to consumers. It is not reasonable that data brokers cannot find a way to contact consumers who are not their direct customers, but nevertheless have lists and APIs filled with highly identifiable personal data of these same consumers, including email addresses, home addresses, phone numbers, and sometimes social media handles. Of all entities, data brokers have the information on hand to make appropriate breach notification—even those that do not have a direct relationship to the consumer.

Q.2.b. Is there a way for consumers to better control how their data is shared with brokers, perhaps by requiring some sort of affirmative consent?

A.2.b. Requiring consent in some circumstances and providing a uniform opt-out with enforcement procedures and penalties for non-compliance would be helpful for better controlling data management among data broker companies.

Currently, there is not a uniform, comprehensive, or simple way for consumers to control how their data is shared with brokers, nor to opt out. Not all data brokers provide an opt out. Those that do can be difficult for most consumers to find. To opt out of all data brokers operating in the United States is not possible today. Even if it were possible, most consumers would need to be an extraordinary amount of time to find and request data broker opt outs. A central data broker registration point would be helpful to solve this problem.

Vermont passed a modest but important data broker registration law that did not include opt-out requirements. However, the registration law is still helpful so that consumers know what data brokers are operating in their State. A handful of other States have passed some limited opt-out requirements, for example, some States allow members of the judiciary and law enforcement the right to opt out of data broker databases.

Both data broker registration and opt-out requirements have roles to play in improving consumer control.

Q.3. The World Privacy Forum's website says "Some commercial data brokers allow some categories of consumers to opt out of some limited uses and disclosures of personal information." That quote

does not inspire confidence in consumers that they have control over their data.

Q.3.a. Does the data broker industry have a comprehensive and uniform opt-out policy for consumers?

A.3.a. No. The data broker industry does not have a uniform or comprehensive opt-out policy for consumers. The data broker industry has a poor record of how they handle opt outs. Here are some of the key issues:

- Opt-outs often require additional identity information, including digital scans of Government IDs, which consumers are rightly concerned about giving to a data broker.
- Some sites charge opt-out fees. For example, the DMA charges a fee to consumers to opt out. Consumers should be able to opt out free of charge.
- Data brokers—many of them—make the opt-outs so difficult that the hurdle is too high for any but the most persistent and determined consumer. See the FTC complaint we wrote in regards to this issue.⁸ There are also a lot of nudges to redirect people from opting out.
- We have worked with many survivors of crime and domestic violence regarding data broker issues. When we work with individuals to try to opt out, we find that it takes people about 40 hours on average to get through all of the opt-outs. And that is a first pass of just the larger data brokers that do allow opt-outs.
- Not all opt-outs “take.” The rates for opt-out failure vary widely by site.
- FCRA compliance among data brokers is woefully low; data brokers that are offering background checks often disclaim responsibility by noting that consumers can only search for themselves. How are these sites ensuring no FCRA violations are occurring? Where is the oversight on this?
- And on top of all of this, can consumers even find all of the data brokers to opt-out from?

Q.3.b. What is the best approach for giving consumers power over their data given that current data broker opt-out options are “quite limited” and that it is nearly impossible to tell the effect an opt-out will actually have?

A.3.b. First, it is important to institute multifactoral solutions. Data brokers present complex problems and challenges for consumers. There isn’t a “single silver bullet” solution that will capture everything.

Second, there are many small solutions which, if put in place, would facilitate meaningful improvements for consumers regarding data brokers. When taken together, if a thoughtful grouping of solutions could be enacted, it would be helpful. (Opt out plus registration plus data breach requirement plus oversight, *et cetera*.)

⁸See <https://www.worldprivacyforum.org/2009/04/public-comments-request-for-declaration-regarding-fairness-of-opt-out-methods-and-investigation-into-axiom-ussearch-publicrecordsnow-and-usa-people-search-consumer-opt-outmethods-for-compliance-with/>.

Third, self regulation has utterly failed in the data broker industry. We do not need to spend any more time on this. It hasn't worked, and is not likely to work.

Fourth, data brokers have many business models. It is a complex sector, and the definitional boundaries are challenging to set. There is not one sole definition anymore of a data broker. It makes sense at this point to consider a variety of regulatory strategies to match the type of data broker. For example, People Search data brokers should be required to provide opt-outs to consumers. Data brokers creating aggregate credit scores should be subject to the FCRA in their uses of household-modeled scores. (The FCRA will need to be expanded for this to happen.)

Solutions that will help:

1. Legislation that requires data brokers to not use or disclose consumer data for any fraudulent or criminal purpose, and requires data brokers to not use consumer data in a discriminatory way or for any discriminatory purpose.
2. Legislation requiring data brokers to provide an opt-out to consumers. All People Search data brokers should be required to provide an opt-out.
3. Legislation mandating a comprehensive, unified opt-out in content and format.
4. Legislation providing for a unified registry of all categories of data brokers (Vermont State statute, exemplar.)
5. Expansion of the FCRA to expand definitions of eligibility to ensure that household or aggregate credit scoring and other meaningful consumer scores are regulated.
6. Legislation that requires all data brokers to provide data breach notification to consumers.
7. Legislation that requires data brokers to maintain security standards, and actively set requirements for meeting security targets, benchmarks, and show security improvements.

Q.3.c. What happens to a consumer's data once they have opted out?

A.3.c. Consumers' data, after they have placed an opt-out request, is most frequently suppressed in some way. The opt-out data is frequently still held by the data broker, but when data brokers "suppress" the data, they do not allow it to be visible to the public for a period of time.

A number of data brokers require opt-outs to be repeated after a period of time, and there are no rules of the road for what period of time will be involved. It can be 1 year, 2 years, 3 years, *et cetera*. Consumers are on their own to keep track of how often they will have to go through the opt-out process.

**RESPONSES TO WRITTEN QUESTIONS OF SENATOR WARREN
FROM PAM DIXON**

Q.1. In response to the Equifax data breach, I opened an investigation into the causes, impacts, and response to the exposure of personal data of nearly 150 million Americans.

Equifax and other credit reporting agencies collect consumer data without permission, and consumers have no way to prevent their data from being collected and held by private companies. My investigation found that Equifax failed to adopt standard cybersecurity measures, in large part because Federal law incentivizes pursuit of profits over the protection of sensitive data.

Q.1.a. Your written testimony notes, “Credit scores and predictions are being sold that are not regulated by [The Fair Credit Reporting Act (FCRA)]” and that “The technology environment is facilitating more scores being used in more places in consumers’ lives, and not all uses are positive.” Your proposed solutions include bringing unregulated forms of credit scoring under the FCRA and studying new areas of eligibility that need to fall under the FCRA. Given the limited ability of individuals to access, control, and correct their personal data, as well as the limited legal framework to regulate data brokers, would the inadequacy of current laws be addressed by regulating data brokers under the Fair Credit Reporting Act?

A.1.a. It would be of great help for Congress to clarify that aggregate credit scores should already be regulated under the FCRA, and to study new areas of eligibility. These actions would provide for significant improvements in solving some of the more egregious issues related to credit and other “grey area” eligibility decisions. These changes, should Congress take action, would remedy certain aspects of the current problems. I agree that these changes would not address every challenge posed by data broker activities. But these changes would capture a good portion of some of the more serious and systemic problems consumers are facing.

In 2013, WPF testified before Congress about non-FCRA or unregulated credit scores, warning that they were problematic and could create consumer harm. In 2014, we wrote a report called *The Scoring of America* that more fully documented the non-FCRA credit scores. We have found that in 2019, unregulated credit scores are now widespread and are being used on data broker lists and in electronic data append services. We are deeply concerned that the use of unregulated credit scores is poised to create substantial, widespread consumer harm as the use of these scores becomes an entrenched business practice.

I would like to respond in additional detail to your questions.

First, regarding issues relating generally to data availability, even though unregulated credit scores use third-party data, which now circulates in abundance, this use does not automatically mean the scores are unregulated. The alternative credit scores such as those offered by PRBC are regulated credit scores. Alternative data is considered regulated just as if it were credit bureau data. This creates a strong basis for determining that it is not just the use of traditional credit bureau data that causes the applicability of the FCRA to a score. Using third-party data therefore does not constitute a condition under which a score does not fall under FCRA regulation.

Second, household-level scores may still be applied to an individual consumer. Even though companies and credit bureaus creating and using unregulated versions of credit scores make great efforts to explain that the scores are “aggregated” to a household level data, or census block-level data, or ZIP+4 data, it does not

mean that the data will not be used as a proxy for a credit score of an individual living at that address.

If an aggregate credit score is applied to an individual at a decision-making point that would be regulated if it were a traditional credit score, then the credit score, even if it is an aggregate, ZIP+4 modeled score, still must be regulated under the FCRA because it is being applied to an individual. We stress that as long as a person's home address is known, then a ZIP+4 credit score can be applied to that person as an individual. Additionally, any person who gives a general ZIP Code at a point of purchase, for example, could be scored in near real-time and decisions can be made about that person as an individual based on the ZIP Code of the neighborhood they live in. In this way, too, unregulated credit scores may be applicable to individuals.

Note the following exemplars:

A. Equifax Aggregated FICO Scores.¹

B. TransUnion offers TransUnion Audiences. This is what the company calls a summary level view of credit profiles at a geographic (ZIP+4) level. This is TransUnion's version of an unregulated credit score, and the scoring is offered as a service.

"Our consumer finance audiences are aggregated and de-personalized using ZIP+4 microgeographies to achieve a high level of targeting effectiveness while maintaining regulatory compliance."²

and

"TransUnion audiences are sourced from anonymized, aggregated consumer credit data, delivering valuable credit behavior intelligence. Built from TransUnion's consumer database consisting of more than 230 million U.S. records, aggregated credit data provides a summary-level view of credit profiles at a geographic (ZIP+4) level. TransUnion audiences target the consumers most likely to have the financial ability to qualify and respond."³

C. Analytics IQ offers a GeoCreditIQ product,⁴ which is its version of an unregulated consumer score. Analytics IQ states that:

"Credit-related data, even summarized at a geographic level, should always come directly from the source—U.S.-based credit bureaus. That is the approach AnalyticsIQ takes to create the foundation of our GeoCreditIQ data. By working directly with the bureaus, our GeoCreditIQ data is extremely accurate and predictive. With GeoCreditIQ marketers get the best of both worlds. The data correlates highly to actual credit

¹ See <https://www.equifax.com/business/aggregated-fico-scores/>.

² TransUnion Audience Buying Guide, <https://www.transunion.com/resources/transunion/doc/insights/buying-guides/TU-digital-audience-buying-guide-july-2018.pdf>.

³ Nielsen Data as a Service Data Partners, TransUnion. <http://sites.nielsen.com/daas-partners/partner/transunion/>.

⁴ Analytics IQ. <https://analytics-iq.com/what-we-do/>. For a more detailed description, see: <https://analyticsiq.com/downloads/analyticsiq-productsheet-geocreditiq.pdf>.

scores, however, it is less restrictive and very powerful in everyday marketing activities.”⁵

D. Experian offers its Premier Aggregated Credit Statistics score. The “The Premier Aggregated Credit Statistics product is derived from the credit profiles of more than 220 million credit-active consumers and averaged at the ZIP-Code level.”⁶ Experian states that this score is “Beneficial to virtually any industry, including debt collections, education, government, financial services, capital markets and data analytics.”⁷ Experian states that customers can “Get unprecedented insight into the credit health of neighborhoods across the United States.” And it also states that it can be used for debt collections, which typically is applied at an individual level. It has used its data to score the top 25 neighborhoods with the most mortgage debt, for example.⁸ Experian’s ZIP Code credit score is offered as a service.

E. NextMark sells a data broker list of “Summarized Credit Scores FICO-Like Mailing List.”⁹ The data card states: “Summarized Credit Scores are used to help our clients target segments of the population at varying levels of credit worthiness. It is carefully built upon the historic financial transaction data of hundred of millions of consumers, aggregated at the ZIP+4 level.” The data card has further recommendations for use:

“Recommendations for Banking, Insurance and Automotive Industries:

Overlay summarized credit scores on your database to determine credit worthy, or subprime for special finance offers.

Recommendations for mortgage industry:

Subprime Program: Identify consumers with debt and credit challenges: Choose summarized credit FICO-like ranges of less than 600, specific loan dates and loan amounts or LTV. . . .”

F. The Dataman Group has “Modeled Credit Score Prospect Lists.”¹⁰ The lists include a profitability score, and uses layers of data to score at the household level.

“This new ConsumerView Profitability Score list select helps identify households likely to pay their debts and ranks households by profitability, allowing marketers to target the best prospects based on:

Profitability

Approval Rates

⁵ Analytics IQ GeoCreditIQ brochure, <https://analytics-iq.com/downloads/analyticsiq-product-sheet-geocreditiq.pdf>.

⁶ Experian Premier Aggregated Credit Statistics. Available at <https://www.experian.com/consumer-information/premier-aggregated-credit-statistics.html>.

⁷ Supra note 5.

⁸ Experian Blog Post, ZIP Codes with the Highest Mortgage Debt, July 22, 2019. <https://www.experian.com/blogs/ask-experian/research/zip-codes-with-the-highest-mortgage-debt/>.

⁹ Nextmark, <https://lists.nextmark.com/market.jsessionid624D63468C12F73E52082D474F1C49C9?page-order/online/datacard&id=281247>.

¹⁰ Dataman Group, Modeled Credit Score Lists, <https://www.datamangroup.com/modeled-credit-score-lists/>.

Response Rates

The scores align very closely to bonafide Credit Scoring—and with this file—no preapproval is needed!

The ConsumerView Profitability Score combines a robust scoring model that offers high levels of refinement for selecting the most profitable prospects combined with our top-notch Consumer Database. This gives you greater precision in predicting, identifying and targeting prospects at the Household Level.”

These are just a few exemplars of the ways in which unregulated credit scores are being used today.

Third, credit scores may only be pulled for purposes strictly defined in the FCRA; they cannot be used for general marketing purposes. It is already established policy, and law, that credit scores cannot be used for general marketing purposes except in situations expressly defined by the FCRA. Given that unregulated credit scores are accurate proxies for regulated credit scores, the use of aggregate ZIP+4 credit scores for expansive marketing purposes currently violates established law and public policy about uses of credit scores. If credit scores were meant to be used for expansive marketing purposes, then the FCRA would permit such uses.

And finally, despite the apparent applicability of the FCRA to aggregate credit scores, we do not see mechanisms that have been made available to consumers for making the uses of these scores transparent. We do not see prominent efforts by credit bureaus to allow consumers to see their ZIP+4 credit scores, nor household scores, nor reveal who has requested their unregulated credit score. We do not see mechanisms for consumers to correct errors in their unregulated scores, or to prevent other abuses the FCRA and ECOA were designed to address. We do not know how or if the credit bureaus are affirmatively tracking, monitoring, and policing the uses of unregulated credit scores, and we are greatly concerned that these scores may also be easily used both applied at an individual level and used for eligibility purposes. We do not see the credit bureaus and others reporting publicly their technological proof of compliance with the FCRA regarding the unregulated credit scores.

Unfortunately, consumers are not able to avoid the harms involved with unregulated credit scoring. The lists and databases of millions of consumers appended with their unregulated credit scores occur without consumers’ knowledge or ability to correct the data. Financial, educational, employment, and other opportunities based on a person’s unregulated ZIP+4 or household credit score may have profound impacts on individuals, but they will not be able to use existing FCRA tools to remedy the problems posed by this category of credit scores.

If Congress clarified the FCRA to bring aggregate credit scores clearly under the auspices of the FCRA, with no interpretational grey areas, it would provide meaningful, significant improvement. Aggregate credit scores would no longer be able to be used for marketing purposes, these types of credit scores would not be able to be quietly applied illegally to individual consumers, and an avenue of growing harm would be closed.

Q.1.b. Credit reporting agencies make billions of dollars collecting and selling information about consumers, but consumers have little ability to control how their personal information is collected and used by these agencies. How would legislation to give consumers more control over personal financial data and to create a uniform, Federal process for obtaining and lifting credit freezes benefit consumers? Would consumers benefit if such legislation also applied to currently unregulated parts of the industry, such as data brokerages?

A.1.b. When identity theft remedies were being put in place from the mid-1990s through the early 2010s, I observed in real-time how these remedies beneficially impacted consumers through the many phone calls that came in to World Privacy Forum. After State security freeze laws were enacted, consumers with multistate identity theft issues experienced significant relief, as did single-state victims of identity theft. Security freeze laws have worked well for consumers, particularly those with serious identity theft in their present or past. If a uniform Federal process took the strongest and best of the State laws and created rapid setting and lifting of security freezes, that could be beneficial.

It would be beneficial for security freezes to apply across data brokerages as well. This would assist in cases of identity theft, and it would assist with safety considerations. We have found that in particular, victims of crime, including domestic violence and stalking among other crimes, as well as elected officials and law enforcement officers, have safety considerations that apply to data broker data.

Q.2. Your written testimony calls for legislation to facilitate setting due process standards that would fill in meaningful gaps in privacy protections. Along with Professor Jane Winn, you suggest legislation that would give the Federal Trade Commission additional authorities to regulate practices in connection with personal data. Relatedly, I have introduced legislation to give the Federal Trade Commission more direct supervisory authority over data security at credit reporting agencies.

Q.2.a. How would legislation to establish and provide Federal authority and resources to monitor data security practices of credit reporting agencies and data brokers benefit consumers?

A.2.a. Legislation that would provide Federal authority and resources to monitor data security practices of CRAs and data brokers could benefit consumers in several ways; by setting guardrails for the data broker sector generally, by giving consumers more agency in the overall process, and by requiring data brokers and CRAs to manage data using processes documented to facilitate ongoing improvements in outcomes.

By way of background, the current debate over what Federal information privacy legislation should look like is often based on the assumption that there are only two models to choose from: a market-based approach or a hierarchical rights-based approach. Applying Nobel Laureate Elinor Ostrom's principles of governance design (Nives Dolsak, Elinor Ostrom & Bonnie J. McCay, *The Commons in the New Millennium* (2003) and a pragmatic understanding of scientific knowledge as socially constructed makes it possible to find

a middle path between a market approach or a hierarchical approach to information governance.

Successful examples of governance mechanisms that lie on this middle path include privacy standard setting processes, as you noted in your question. Such collaborative standards-setting efforts should not be confused with privacy self-regulation, which is one example of a market approach that lacks accountability because, as the economist Anthony Ogus pointed out in *Rethinking Self-Regulation*, (Oxford Journal of Legal Studies, 1995), private self-regulation is per se captured from its inception.

The term “voluntary consensus standards” has a specific meaning that is already defined in law. The U.S. Food and Drug Administration has been using voluntary consensus standards that comply with due process requirements as articulated in the Office of Management and Budget (OMB) Circular A-119 for more than 20 years, which has resulted in more than 1,000 recognized standards applicable to medical devices. The World Trade Organization (WTO), *Agreement on Technical Barriers to Trade* is a core document that outlines how standards may be set by independent parties in a fair and appropriate manner that does not create transactional or other barriers. These ideas have applicability to data ecosystems and privacy risks.

Within the framework of due process guarantees set out in OMB Circular A-119, Federal regulators today have the power to recognize compliance with voluntary, consensus standards as evidence of compliance with the law for specific, limited regulatory purposes. Federal regulators may only use voluntary consensus standards to create such safe harbors if the standards can be shown to have been developed through processes whose openness, balance, consensus, inclusion, transparency and accountability have been independently verified.

When the interface between Federal legislation and voluntary, consensus industry standards is working correctly, then the private sector (inclusive of all private sector stakeholders) takes the lead in developing appropriate, context-specific standards for solving policy problems. Next, regulators take the lead in assessing whether those private standards meet the needs of the American public as well as the industry players that developed them. These assessments will ideally be conducted in an ongoing manner, and can realistically include monitoring that is in real time or near real time. Finally, courts stand by ready to serve as independent arbiters of the behavior of both industry and Government.

Beyond the standards approach, another important set of measures relates to governance that ensures ongoing improvement targets are set and achieved. See my response to B, below.

Q.2.b. In your view, would legislation to impose strict liability penalties for breaches involving consumer data at credit reporting agencies and data brokerages lead to improvements in consumer data security? Would consumers benefit if such penalties were imposed on data brokers?

A.2.b. Credit Reporting Agencies and data brokers have a heightened responsibility to ensure data integrity on all fronts, including responsibilities related to data security, data integrity, and data

breaches. Strict liability requirements can have a place in highly sensitive data settings to ensure the highest standards of data integrity are being met.

Much has been learned in the last 25 years about data protection and digital ecosystems. Data protection laws that have already been enacted in 123-plus countries have grown to have significant similarities, even when aspects of the law have been adapted to unique county-level conditions. See for example, the work of Graham Greenleaf on this topic. Data breach requirements are spreading globally.

However, despite all of the work on privacy and data protection, baseline governance principles that have demonstrated worth in other settings such as environmental, manufacturing, and law enforcement contexts, have generally not yet been applied in the privacy realm. This is a rich area for exploration regarding legislation.

By themselves, strict liability requirements are not enough to create reliably good results in the long term if the goal is to substantively improve outcomes for consumers and for the businesses that must comply with data breach laws. A comprehensive governance system is needed that will facilitate the creation of specific and appropriate benchmarking and improvement processes to achieve improvement goals.

Here, we point to the expansive and demonstrably productive work of W. Edwards Deming, including his system (and principles) of management¹¹ and his process cycle of continual improvement.¹² If legislation were to go beyond strict liability and also enshrine such types of ongoing improvement processes as part of the principles of governance within a privacy or data breach context, it would go far to creating a more mature and effective approach to data systems and processes. Over time, while strict liability will have certain baseline compliance effects, it is primarily a tool for deterrence. It does not fully work to complete the job of bringing businesses up to significant levels of improvement. For this to happen, affirmative governance structures also need to be in place. Given that privacy is still catching up to other business systems thought in other sectors, enshrining ideas of continual improvement would be helpful in creating an environment where better systems of data governance can be created.

RESPONSES TO WRITTEN QUESTIONS OF SENATOR SCHATZ FROM PAM DIXON

Q.1. Are data sets collected by data brokers getting into the blood stream of credit, employment, and housing decision making, in a way that evades the FCRA?

A.1. Yes, data sets regarding consumers that are held by data brokers are being used for credit, employment, and housing decision making in ways that may evade the FCRA. Going one step further, data broker data is being used to create consumer scores being used in eligibility situations, and this also evades the FCRA, or closely skirts it. In our *Scoring of America* report we documented

¹¹ See <https://deming.org/explore/fourteenpoints>.

¹² Plan, Do, Study, Act; <https://deming.org/explore/p-d-s-a>.

many of the various data streams that data brokers utilize in gathering consumers' personal data, and we documented the scores themselves.

In particular, *aggregate* or *modeled* credit scores are particularly challenging in regards to FCRA compliance. These are scores that are typically modeled on ZIP+4, census block, or the household level. They are often marketed as comparable to regulated credit scores. When household credit scores are applied to the individual, I believe this violates the FCRA. When the household credit scores are used in eligibility circumstances at the individual level, this, too, I believe is a violation of the FCRA. In my testimony, I discussed the FICO Aggregate Credit Score. It is not the only such score in this category.

Tracking the proliferation of aggregate and modeled credit scores is one way to see the significant potential for skirting of the FCRA. Questions abound:

- How many of these scores are being used in eligibility circumstances?
- How are these scores being used in marketing or other circumstances?
- How are the companies policing the use of these scores?
- To whom or what entities have the scores been sold?
- How can the companies producing aggregate credit scores affirmatively demonstrate that their product is only being used in full compliance with the FCRA?

There are limited ways available to track data broker data. However, one of the ways to get a glimpse of it is to review the data broker data cards that are available via the list broker or data broker websites. Examples include:

- NextMark List Finder: <https://lists.nextmark.com/market>.
- Exact Data Consumer Lists: <https://www.exactdata.com/consumer-mailing-lists.html>.
- InfoUSA Consumer Lists: <https://www.infousa.com/lists/consumer-lists/>.
- Dataman Consumer Lists: <https://www.datamangroup.com/national-consumer-database/>.
- Experian Consumer Sales Leads: <https://www.experian.com/small-business/sales-leads.jsp>.

This is a very small selection of offerings of detailed consumer data available via lists. I note that this is just one aspect of data brokering. It happens to be the easiest to demonstrate at this time; however, many other data broker activities occur out of sight, for example, data APIs, which provide the "list" on demand and will likely replace older list methods fairly soon.

And to reiterate, it is crucial to understand that the production of consumer scores is a way to condense raw data broker data into numeric shorthand. Unregulated consumer scores can be as challenging to the FCRA as the original raw data, and can cause harms when misused in eligibility circumstances.

Q.2. Under current law, do companies that collect and sell information about consumers have any duty to consumers about how that information will be used? If consumers are discriminated against or harmed because of how that data is used, who is responsible?

A.2. There is not yet a broad, comprehensively applicable rule applicable to duties of care regarding the use of consumer data. There are some sectoral protections in place. Additional pressures from the States have created a very narrow pathway for some rules in some circumstances. We note that California's law, the CCPA, has numerous exemptions and loopholes, and thus, even in California there is not a broad law that will apply routinely to all data brokers. Because of this, there is no question that there are meaningful gaps in consumer protection at the State and Federal level.

At the Federal level, the answer to the questions of duty and responsibility depends on what entity is holding the data, what sectoral regulations are in place, and for unregulated companies, what the privacy policy of that company states. For example, HIPAA-covered entities do have a duty to patients about how protected health information will be used. Entities engaging in FCRA-covered activities also have some duties to consumers about information use. As good as the FCRA is, in some ways, as I mentioned in testimony, it has lost some of its effectiveness due to what has become the "household" vs. individual loophole. In the public sector, the Privacy Act does make some stipulations about data use.

For companies that are not regulated under a sectoral regime, the FTC can enforce privacy policies that are posted by companies under its FTC Act §5 authority; but this has its limits, and does not provide for a proactive requirement of certain duties to consumers regarding data use.

Vermont, in enacting its first-in-nation 2018 data broker legislation, made incremental steps at a State-level toward creating at least some duty regarding consumer data when it required data brokers to not use consumer data for committing fraud, or in a discriminatory way. This is not a comprehensive protection, but it remains an important exemplar.

Q.3. If consumers are discriminated against or harmed because of how that data is used, who is responsible? If a data broker is breached and a consumer suffers harm from identity theft, who is liable?

A.3. The answer to both of these questions will depend on the circumstances of the discrimination or harm, and the complexities of resolving this issue are no small matter. In an FCRA context, consumers who experience harm because of improperly conducted background checks, for example, have recourse. In this situation, an employer may be the responsible party, or the background check provider. But outside of the FCRA context, harms can accrue that are unregulated, which makes the assignation of responsibility more difficult in some circumstances.

For example, when a business uses an aggregate or household credit score to determine eligibility for a financial service or product, and chooses to decline the consumer for a service or product, unless the consumer had a way to know about this declension, they would not be likely to learn about the harm. In this situation, the

creator of the aggregate or household score, the seller of the score to the institution that used it, and the institution may possibly have some responsibility, but this is not yet litigated under the FCRA, and Congress has not yet clarified the issue of aggregate or modeled credit scores. Until and unless we have additional clarity, it will be very difficult to have bright-line responsibility assignments in this and other areas.

Regarding data brokers and unregulated scores generally, there is a need for more bright-line rules in regards to responsibilities and duties, including nondiscrimination.

Currently, outside of the State of Vermont, and as of 2019, also California, which have both passed basic data broker registration laws, the answer to this question is not straightforward whatsoever, and in large part, it is fair to say it is undetermined. In most cases, consumers are unlikely to be able to determine with specificity how their information was compromised, or what party created the risk. In the case of consumer data held by data brokers, it would be very difficult for consumers to know which data brokers held their data, much less which had breached their data. Specific data broker breach requirements and other protections would help ameliorate some of these problems.

Q.4. Do you think Federal law should require companies that collect and use consumer data to take reasonable steps to prevent unwanted disclosures of data and not use data to the detriment of those consumers?

A.4. Yes. There are no reasonable arguments against providing proper security for consumer data at all stages of its lifecycle in a business. And there are no arguments against prohibiting using data in a detrimental, discriminatory, or unfair way. It is essential to provide for fair data uses and prevention of harm regarding consumer data; without such provisions, consumer trust will eventually be lost. Abusive data practices where data is used in detrimental, discriminatory, or unfair ways in consumers' lives is not sustainable in a digital economy.

**RESPONSES TO WRITTEN QUESTIONS OF SENATOR CORTEZ
MASTO FROM PAM DIXON**

Q.1. Are there firms that you think are utilizing algorithms to expand access for affordable credit or useful financial products that are beneficial? If so, which ones?

A.1. Some beneficial examples in this context are found in the area of "thin file" consumer scoring products. These types of credit scores are well understood in the marketplace. Typically called "alternative credit scores," thin file credit scores are almost always brought in as regulated scores under the FCRA. Alternative credit scores typically use a small alternative data set to calculate thin file scores. Utility payments, rent payments, phone bill payments, and other types of steady payments are used as predictors for credit risk for people who may not have purchased a home, a car, and may not have an extensive credit history for a variety of reasons.

Exemplars include the FICO UltraFICO,¹ and ID Analytics use of alternative credit data,² particularly the Credit Optics Full Spectrum.³ These products utilize *alternative data* to provide credit score analysis, and at last check, the companies consider the products to be regulated under the FCRA.

Thin file or alternative credit scores should not be confused with aggregate credit scores. Companies building aggregate credit scores typically do not see these models as regulated under the FCRA, because these scores apply to households, not individuals. This is a loophole in the FCRA, as the FCRA only applies to individuals. Aggregate credit scores that are created at a household level are not regulated, but they nevertheless might be applied to individuals by companies seeking an unregulated predictive score.

Aggregate credit scores can use hundreds and up to more than a thousand factors, and can be quite accurate. In short, aggregate credit scores can act as an unregulated proxy for the traditional credit scores originally regulated under the FCRA. This is in contrast to thin file, alternative credit scores, which are regulated scores that can be beneficial to previously unscored consumers or consumers with minimal credit histories.

Q.2. Do you believe that people should get to see their unregulated credit reports and scores just as they do their regulated scores?

A.2. Yes, people should be able to see their unregulated credit reports and scores. For example, we should be able to see our FICO aggregate credit score. We should also be able to see our Experian neighborhood risk score, as this score is used to create a variety of metrics about households and those living in that household. Any score used in matters relating to eligibility, or used to determine the character, reputation or creditworthiness of an individual should be available and not secret.

Q.3. What does it mean for financial markets now that FINRA can essentially predict and decide in real time, or near real-time investor behavior? What does it mean for other financial and technical sectors?

A.3. FINRA is a key exemplar of modern real-time governance. It didn't begin that way, but the system has evolved in important ways. We think that FINRA is just the beginning of the "real-time governance" movement, where high volumes of data analysis and governance is what a lot of compliance reporting is going to start looking like in the United States and elsewhere.

As a self-regulatory organization under the Securities and Exchange Act ('34 Act), FINRA is authorized to issue rules under Section 15A(b)(6) of the 1934 Act in order to ". . . prevent fraudulent and manipulative acts and practices, to promote just and equitable principles of trade, and, in general, to protect investors and the public interest and Section 15A(b)(9) of the Act."

Q.4. In the past, FINRA produced periodic summarized reports to support its mission. This was fine, and entirely appropriate for a

¹See <https://www.fico.com/en/products/ultrafico-score>.

²See <https://www.idanalytics.com/solutions-services/credit-risk-solutions/alternative-credit-data/>.

³See <https://www.idanalytics.com/solutions-services/credit-risk-solutions/>.

paper-based economy and era. From the 1930s when the modern U.S. securities law framework was established through to the present, regulators such as the Securities and Exchange Commission and SROs such as the New York Stock Exchange and the National Association of Securities Dealers (whose SRO powers were eventually transferred to FINRA) had no choice but to rely on periodic reporting from regulated entities as their primary source of information. Staff members of regulated entities spent huge amounts of time boiling down vast quantities of raw data into highly simplified, abstract form for reporting. Then staff members of regulators tried to develop an accurate understanding of the complex reality summarized in the reporting forms through a combination of analysis of the reporting forms and selective audits. These paper-based reporting and regulatory processes were normal and appropriate and used throughout the American economy and world for most of the 20th century.

The computerization of American financial markets was driven in the late 1960s and 1970s by the “paperwork crunch” on Wall Street. As trading volumes increased, paper-based clearing and settlement systems became overloaded, making it impossible to settle all of 1 day’s transactions before the start of the next trading day. The first response to the paperwork crunch was to close markets earlier, which was obviously not a solution that appealed to either financial firms or their clients.

By the end of the 1970s, clearing and settlement systems were running on mainframe computers and American banks, brokerage firms and insurance companies were world leaders in the computerization of their back-office systems. The regulatory financial reporting obligations of these firms were met through a combination of reports generated by mainframe computer systems and information collected and summarized by staff members. These reporting and regulatory oversight processes were based on point-in-time, low-resolution snapshots of the business operations of regulated entities. Regulators could see the equivalent of the tip of an iceberg and were forced to guess the characteristics of the submerged portion of the iceberg. The executives running regulated entities were in much the same position.

In his book, “Seeing Like a State,” Harvard political science Professor James Scott wrote a book, articulated the challenges that modern regulators face when forced to make decisions on the basis of the kind of highly compressed summaries of complex realities found in periodic reporting by regulated entities. The regulator can literally “see” only what is presented in the summary, and on the basis of that kind such summaries, make educated guesses about where to look more closely for evidence of violations of law.

Following the Stock Market Crash of 1987, regulators began working with regulated entities to better understand the operation of their computer systems and to integrate the functioning of those computer systems more directly into their regulatory oversight activities. As regulators gained greater direct access to the information begin generated by the information systems operated by regulated entities, they gradually were able to “see” something closer to what the executives of regulated entities could see.

By the 2000s, financial market regulators such as the SEC and FINRA were developing the capacity to collect and analyze raw data feeds directly from regulated entities. This brings us to today, where FINRA is using the availability of increased technological capacity to acquire real-time transaction data regarding TRACE—eligible securities (Trade Reporting and Compliance Engine). Instead of receiving periodic reports, those subscribing to FINRA's TRACE reporting system now have firehoses of real-time data to manage and analyze.

In the FINRA real-time environment, regulators now have to develop their own capacity to analyze these data feeds and draw their own inferences from them, which requires huge investments in computing capacity and staff with relevant subject matter expertise. After these systems are fully operational, then in theory what regulators should be able to “see” whatever executives at regulated entities can “see.” The starting point of the dialogue between regulators and regulated entities can focus on comparing the results of the regulators' analyses and the regulated entities' analyses of the same raw data generated by the regulated entities' computer systems.

FINRA's TRACE reporting system was developed specifically to assist with this process. To meet its primary mission, FINRA will need to continue to ensure that the kinds of compliance problems they look for, such as concealed shell companies, achieve maximum benefits from the data volume and velocity “real time” affords. “Real time” does not automatically equal “better” unless foundational work has been done to ensure that the data has been properly tagged and organized to facilitate compliance reporting and response. For example, compliance alerts in real-time systems are typically based on some form of trigger. Various kinds of data tags and identifiers are particularly important to construct properly to fulfill this task. With proper triggers in place, real-time data firehoses can be purposefully and reliably analyzed at scale and at speed in order to create accurate real-time governance feedback.

The ability of regulators to request real-time data from regulated entities and to engage in real-time analysis of that data for evidence of compliance or violations of the law by the regulated entities represents the beginning of a new era of “real-time governance.” In a real-time governance system, regulators should be able to respond almost as quickly as regulated entities to evidence of a risk of noncompliance. The expansion of real-time governance in the United States and around the world promises a fundamental breakthrough in risk management: citizens should be able to enjoy the best quality goods and services and the benefits of rapid technological innovation while at the same time also being provided better protection from risks.

In order to lay a foundation for continuous improvement of real-time governance systems, regulators and regulated entities will need to collaborate to increase the standardization of data formats. Back in the 1970s, when each financial service firm was installing its own mainframe computer, it was not uncommon for each firm to acquire custom-developed, bespoke software application. Standards were developed for transaction data so that first it could send and receive order and execution information from exchanges and

other firms quickly and accurately, but there was no need to standardize other parts of the firms' computer systems.

By the 2000s, the result was significant diversity across firms in the way that some of the information relevant to their reporting obligations was generated and stored. Limited standardization of data formats and software architectures across regulated entities increases the challenges to regulators to move to real-time governance because of their need to compare compliance-related behaviors across different firms with different computer systems.

Lack of standardization of data formats hampered regulators' ability to respond to the 2008 collapse of Lehman Brothers and the 2010 Flash Crash. Regulators' efforts to track down the course of large volumes of computer-generated orders were hampered by the difficulty of comparing data generated by different firms. One problem in particular had to do with lack of standardization in how customers that were "legal persons" (*e.g.*, corporations), were identified. The same corporation's name might be entered into different firm computers differently due to the use of nonstandard abbreviations or even typographical errors. The lack of global standards for identifying common ownership of financial accounts by business entities quickly and accurately was hampering tax and anti-money laundering regulatory efforts as well.

In 2011, the Depository Trust & Clearing Corporation (DTCC) and the Society for Worldwide Financial Telecommunications (SWIFT) launched a collaborative, global standard-setting effort that led to the creation of the "Global Legal Entity Identifier" standard. This standard has been endorsed by the Financial Stability Board and the G20 and designated as International Organization for Standardization ISO standard 17442. Some jurisdictions outside the United States have begun mandating the use of LEI numbers in certain financial service markets in order to increase the effectiveness of regulatory oversight processes (*e.g.*, EU Markets in Financial Instruments Directive known as MiFID II).

Any legal entity anywhere in the world can obtain quickly, easily and cheaply a globally unique 20 digit LEI number from the LEI issuer of their choice, and be confident that it will be accepted by regulators and counterparties around the world for compliance purposes. The LEI Regulatory Oversight Council and the Global Legal Identifier Foundation (GLEIF) jointly administer the LEI system. This includes the oversight of a global network LEI issuers that compete with each other to issue LEI numbers to entities; providing the Global LEI Index, an open, searchable database of LEI numbers, and monitoring emerging technologies and updating the standard as needed to accommodate them.

The LEI ROC and GLEIF provide a clear example of the kind of transparent, accountable and inclusive governance processes that are needed to insure that real-time governance serves the public and is not captured by industry or leveraged by owners of proprietary technologies. The LEI ROC and GLEIF operate in all global markets simultaneously to reduce compliance burdens on regulated entities, amplify the effectiveness of national and global regulators' efforts to protect the public and are completely transparent to end users.

But the public, the regulators that represent the public interest, and private firms cannot enjoy any of those benefits of real-time governance without a very large, one-time investment by the private sector in business process reengineering. That is because all private enterprises today have some system for identifying themselves to their counterparties and keeping track of their counterparties that was developed before the global legal entity identifier standard was developed. The problem from a software programming perspective is similar to the Y2K problem at the end of the 1990s: software programs that only allocated two digits for storing information about years had to be modified to accommodate four digit years in order to insure that the year 2000 was not interpreted by the software as 1900 instead. In a similar manner, all business software systems will have to make a one-time change to adopt GLEI and phaseout whatever other system they were using. Depending on how a firm's computer system is organized, this may require undertaking a long, slow, difficult process to achieve what appears to be a simple and obvious outcome to anyone not familiar with the challenges of business processing reengineering.

With regard to the ability of FINRA or any other regulator working with real-time data feeds to fulfill their public service mission through real-time governance processes, increasing standardization of data formats is an essential part of the process of increasing the accuracy of regulators' ability to predict the behavior of investors, regulated entities and markets generally. The kind of predictions that the use of big data and artificial intelligence make possible are statistical inferences about the probability of different outcomes. The use of data analytics would permit a regulator to estimate the probability that certain data revealed a violation of the law.

Using real-time data flows and real-time governance processes in this way permits regulators to engage in provable, fact-based, and "risk based" regulation. This would permit regulators to adjust dynamically and in real-time their allocation of scarce enforcement resources to those situations where they would create the most value for the public. They could use real-time governance mechanisms to identify those situations where the regulator believes the probability of a violation of the law occurring is the highest and the risk of harm to the public as a result of that violation is the highest, and concentrate their resources there.

The migration by regulators to real-time governance in effect levels the playing field with regard to what the executives of regulated entities know and what regulators know. In addition, regulators gain deeper insight into the behavior of markets generally because unlike the executives of regulated entities who can see in detail only their own firms' internal operations, regulators will be able to learn from comparing detailed, accurate information about operations of all regulated entities.

As regulators give up the 20th century system of regulation based on information contained in point-in-time, low resolution snapshots of the behavior of regulated entities and move to real-time governance instead, regulators will be able to use whatever resources they have more effectively, the public will be better protected and regulated entities will benefit from greater predictability and consistency of regulatory enforcement actions.

It is difficult to overstate the potential significance of the move from 20th century command and control bureaucratic regulatory processes to real-time governance process not just in financial services but in every sector of the American economy and across global markets. In the 19th century, governments could only act as a “night watchman state” because of their limited capacity to regulate the economy. By the 20th century, the modern regulatory State had come into being and could act to protect the public from tainted food, poisonous medicines and lethal workplaces. The Administrative Procedure Act of 1946 was enacted to insure that the power of the modern regulatory State was exercised in a manner consistent with the rule of law.

The fundamental advances in accountability and effectiveness ushered in by the APA such as notice and comment rulemaking cannot meet the challenge of insuring that regulatory power exercised through real-time governance processes also conforms to the rule of law. In order to lay a statutory foundation for the transparent, accountable and inclusive exercise of regulatory power through real-time governance processes, a fundamentally new approach to regulation is required.

Such a new legislative interface would be congruent with the APA but would explicitly authorize regulators to leverage voluntary, consensus standards developed by private standard-setting organizations that have committed to observing due process. Public-private collaborations between Federal regulators and private sector standard developing organizations have been taking place for decades with the framework of Office of Management and Budget Circular 119–A governing Federal Participation in the Development and Use of Voluntary Consensus Standards and in Conformity Assessment Activities and most recently updated in 2016. This new approach to regulatory governance is discussed in more detail in the information privacy law context in Pam Dixon and Jane Winn, *From Data Protection to Information Governance* (forthcoming 2019) and Jane Winn, *The Governance Turn in Information Privacy Law* (July 11, 2019), <https://ssrn.com/abstract=3418286>.

Real-time financial sector analysis is no longer a single-jurisdiction endeavor. It requires multilevel cooperative efforts. The example of the Global LEI standard demonstrates that the use of a legislative interface through which regulators and private standard-setting organizations can collaborate to achieve real-time governance that serves the public can work any context, not just information privacy law. It also demonstrates that the transparency, accountability and inclusiveness of real-time governance can be supported by cooperative efforts with global standard-setting organizations as well as American standard setting organizations. How these cooperative efforts are accomplished requires careful and methodical decision making and planning—private organizations and the public sector both need to be fully committed to insuring the fundamental fairness of their own processes. FINRA’s system gives us a view into the implications of the world to come, and the depth of its new technical and policy requirements.

Q.4. Do you believe that there should be something similar to the “legitimate interest” basis for data processing in the United States

and, if so, how should we think about nonconsent-based processing for entities that have no consumer relationship such as data brokers?

A.4. Data processing that is not based on consent is an important issue to address, because it is going to become front and center in the predictive world we are moving into. It is not reasonable to think that individuals will be able to consent to every bit of processing of their data. That being said, we still need structures that ensure nondiscrimination and people-beneficial uses of data. Processing varies in levels of importance depending on the context and use of the processing and data, among other factors.

We now have some experience with legitimate interests processing via the GDPR in Europe. Legitimate interest-based processing has proven to be a challenging issue to implement, and the results have been uneven thus far. Because of the implementation issues with the GDPR, I prefer the idea of routine uses as outlined conceptually in the Privacy Act of 1974. The United States routine uses model allows for data processing within limits, based on the context, but prohibits other uses outside of the known context and requires affirmative consent as the uses and data become more sensitive.

One of the questions that immediately arises regarding both legitimate interest and routine uses is: who gets to decide what is a legitimate interest, or what is a routine use? This is an important question in a democratic society, and is one of the biggest decisions that needs to be determined in a democratic process. In the Privacy Act, the concept and structure of routine uses allows for individuals, businesses, and other entities to have a voice in what those routine uses look like, but it is the Government that has the ultimate authority to make bright-line decisions.

The details of deciding upon routine uses can be managed by utilizing a combination of sectoral legislation to decide the brightest lines (like the floor for HIPAA) and the addition of due process voluntary consensus standards that would allow all stakeholders to have a fair and robust dialogue to create the more granular rules for what constitutes fair routine uses in more particularized settings. Voluntary consensus standards are due process standards, where all stakeholders have a say in what those “routine uses” should look like. This kind of standards work is in contrast to industry self regulation, where only industry has a role in the process and key stakeholders (such as consumers) might not be included.

Again, in some areas, and applying the routine use idea broadly, beyond the confines of the Privacy Act, Congress will need to make the general bright line boundaries for some “routine uses.” At a more granular level, multistakeholder work can set the finer boundary lines, with input from all stakeholders. Anything that goes beyond a checkbox will involve a more time-intensive process, but one that is well worth the effort.

Q.5. How effective are the GDPR’s provisions surrounding profiling and automated decision making, and is that something we should emulate in the United States?

A.5. AI and machine learning systems require a lot of data, and they can present a variety of meaningful risks, including serious potentials for bias and inappropriate manipulations. The approach the GDPR took to automated decision making is understandable given the risks, yet the approach is also proving to be problematic. I spent over a year as a member of the OECD's AI Expert Group (AIGO). The AIGO group was tasked with providing extensive technical input into the OECD Principles on AI, which have now been ratified by the United States and other OECD countries, see: <https://www.oecd.org/going-digital/ai/principles/>.

Something that became very apparent throughout the discussions of AIGO was that the GDPR approach to AI processing brings many noncompetitive restrictions to data use and analysis. The OECD final guidelines took a broader approach than the GDPR, one that respected human values and privacy, and also innovation and economic growth. It is important that democratic societies such as the United States stay highly competitive with other jurisdictions in regards to AI and Machine Learning. The Belt and Road Initiative (BRI) countries (<https://www.worldbank.org/en/topic/regional-integration/brief/belt-and-road-initiative>) are focused on winning the AI and Machine Learning race, and this focus on achieving AI dominance should not be underestimated.

The United States faces an ethical dilemma. That is: do we handle data as aggressively as nondemocratic jurisdictions do in order to stay competitive? Or, do we protect privacy and take potential risks with our ability to compete? Or is there another way? We cannot take a stance of abusing the privacy, autonomy, and trust of the American people. And we must also innovate and lead in new technologies of prediction. After long consideration, I believe it is imperative that we find the third way, a way that allows us to retain privacy, autonomy, and democratic values while still innovating and staying competitive. This is both worthwhile and possible.

Legislating AI as a broad command and control statute is not possible due to the complexity and variety of AI systems. We believe that an approach where lawmakers determine a set of general principles, then implement those principles with fair standards setting processes using OMB Circular A-119 as a due process model, will work well for addressing the complex challenges AI analytics poses at a granular level.

This is an admittedly complex topic, and we do have forthcoming research on governance of privacy in complex ecosystems. In the meantime, a paper written by Jane Winn, who is a law professor in the United States and has taught short courses in China for many years, articulates some of these issues (and potential solutions): *The Governance Turn in Information Privacy Law* (July 11, 2019), <https://ssrn.com/abstract=3418286> or <http://dx.doi.org/10.2139/ssrn.3418286>.

Q.6. What are some of the gaps in currently existing law with respect to how enforcement agencies deal with this multitude of laws and what should we be thinking about in the Banking Committee as we prepare to potentially consider broader privacy legislation drafted by the Commerce Committee?

A.6. There are several meaningful gaps in existing law regarding enforcement agencies:

- A. Too-narrow of enforcement authority at the FTC
- B. Enforcement gaps between existing sectoral laws
- C. Enforcement gaps of new sectors

Regarding the FTC's enforcement authority, this issue has been well-discussed in Congress. The primary issues are the limitations of The FTC Act to address the full range of modern privacy problems, and the limitations created for the FTC under Magnuson-Moss, which limits the FTC's rulemaking power. The Magnuson-Moss vision of how the FTC should operate is not a viable position for the FTC to be held to today, particularly in light of the privacy and security concerns attending the fast-moving data ecosystem.

Nevertheless, there is a school of thought that the FTC should not be the Nation's main privacy enforcement authority due to its constraints. This leads us to the idea of a new structure. We favor the creation of a Federal oversight board with responsibility for privacy—for example, a 12-member board with broad enforcement oversight. An overarching administrative privacy enforcement council or board would be in a position to spot issues across sectors, agencies, more readily identify a broader variety of gaps, and direct resources.

Regarding enforcement gaps between existing sectoral laws, we see three pathways to enforcement. First, focused laws to fill in the gaps, accompanied with clear enforcement authority. Second, voluntary consensus guidelines at the State and Federal level with Government oversight, again, directed at the gaps where there is the most need. Third, we see a role for certification and other tools to assist with enforcement, again, with Government oversight.

Third, it would make sense to conduct an analysis to identify any new sectors or potential sectors that need separate rules. Data brokers may be such a sector, so may certain kinds of platforms. It is an understatement to note that discussions about regulating a group of businesses would be an incredibly contentious discussion on all sides. Nevertheless, it would still be a good idea to at least have the discussion, because it is both reasonable and possible that at some point in the future certain types of businesses and platforms might be considered a sector unto themselves.

Q.7. How can we ensure the consumer is informed about scoring, profiling, and other decisions that are made about them in their daily lives while balancing the need to not put the entire onus on the consumer?

A.7. Requirements for quality controls such as labeling, certification, audit and documentation, bias and accuracy testing, among other measures are some of the mitigations that could be put in place to reduce informational risks without placing the burden entirely on consumers. Rules that require affirmative disclosure of meaningful consumer scores is important, as are rules that allow consumers to request disclosure of smaller scores. We include below a partial list developed from our original Scoring of America report:

- There should be no secret consumer scores. Anyone who develops or uses a consumer score must make the score name, its

purpose, its scale, and the interpretation of the meaning of the scale public. All categories of factors used in a consumer score must also be public, along with the source category of information used in the score.

- Scores used for meaningful decision making about consumers should be subject to quality controls, ideally stipulated in Federal standards.
- The creator of a consumer score should state the purpose, composition, and uses of a consumer in a public way that makes the creator subject to Section 5 of the Federal Trade Commission Act. Section 5 prohibits unfair or deceptive trade practices, and the FTC can take legal action against those who engage in unfair or deceptive activities.
- Any consumer who is the subject of a consumer score should have the right to see his or her score and to ask for a correction of the score and of the information used in the score. It is the responsibility of business to know when they are using a score to make a decision about a consumer.
- Those who create or use consumer scores must be able to show that the scores are not and cannot be used in a way that supports invidious discrimination prohibited by law.
- Those who create or use scores may only use information collected by fair and lawful means. Information used in consumer scores must be appropriately accurate, complete, and timely for the purpose.
- Anyone using a consumer score in a way that adversely affects an individual's employment, credit, insurance, or any significant marketplace opportunity must affirmatively inform the individual about the score, how it is used, how to learn more about the score, and how to exercise any rights that the individual has.
- A consumer score creator has a legitimate interest in the confidentiality of some aspects of its methodology. However, that interest does not outweigh requirements to comply with legal standards or with the need to protect consumer privacy and due process interests. All relevant interests must be balanced in ways that are fair to users and subjects of consumer scoring.
- The Congress and the FTC should continue to examine consumer scores and most especially should collect and make public more facts about consumer scoring.
- The FTC should investigate the use of health information in consumer scoring and issue a report with appropriate legislative recommendations.
- The FTC should investigate the use of statistical scoring methods and expand public debate on the proprietary and legality of these methods as applied to consumers.
- The Consumer Financial Protection Bureau should examine use of consumer scoring for any eligibility (including identity verification and authentication) purpose or any financial purpose. CFPB should cast a particular eye on risk scoring that evades or appears to evade the restrictions of the FCRA and on the use and misuse of fraud scores. If existing lines allow

unfair or discriminatory scoring without effective consumer rights, the CFPB should change the FCRA regulations or propose new legislation.

- The CFPB should investigate the selling of consumer scores to consumers and determine if the scores sold are in actual use, if the representations to consumers are accurate, and if the sales should be regulated so that consumers do not spend money buying worthless scores or scores that they have no opportunity to change in a timely or meaningful way.
- Because good predictions require good data, the CFPB and FTC should examine the quality of data factors used in scores developed for financial decisioning and other decisioning, including fraud and identity scores. In particular, the use of observational social media data as factors in decisioning or predictive products should be specifically examined.
- The use of consumer scores by any level of government, and especially by any agency using scores for a law enforcement purpose, should only occur after complete public disclosure, appropriate hearings, and robust public debate. A government does not have a commercial interest in scoring methodology, and it cannot use any consumer score that is not fully transparent or that does not include a full range of Fair Information Practices. Government should not use any commercial consumer score that is not fully transparent and that does not provide consumers with a full range of Fair Information Practices.
- Victims of identity theft may be at particular risk for harm because of inaccurate consumer scores. This is a deeply under-researched area. The FTC should study this aspect of consumer scoring and try to identify others who may be victimized by inaccurate consumer scoring.

Q.8. Should some types of data, such as biometric information, even be allowed to be shared with third parties?

A.8. If data—or knowledge derived from that data—is sensitive enough, it should not be shared with third parties unless there are specific protective rules and risk mitigations in place. Some data is too sensitive to simply allow to be freely shared, either because as data it is sensitive, or as combined with other information, it could lead to knowledge impacting an individual’s ability to make a living or purchase a home, or other issues related to eligibility under the FCRA.

Working with data types we know well, consider the Social Security Number. In the 1980s, the SSN had grown to very broad uses in the United States. As a result, at a time when the United States was moving from a paper-based world to a digital world, certain types of crimes—particularly identity theft—were greatly facilitated by the relative availability of SSNs. An early trickle of identity theft legislation in the mid-1990s turned into a torrent of legislation in short order around the use, storage, and protection of the SSN.

SSNs are still used today, but many beneficial protections are now in place. Yes, SSNs are still used by third-parties, for example, by credit bureaus. But generally, SSN uses are much more re-

stricted now. For example, SSNs have been removed from being printed on Medicare cards and on drivers' licenses. Data types and potential for uses need to be evaluated for risks to make a determination about risks related to sharing.

In taking this a step further and discussing knowledge derived from data, think of the mosaic of information that outlines an individual's reputation and character such as that which would be revealed in a comprehensive background check. This is why the FCRA protections around background checks are so important. Background checks may be undertaken, but not without the subject's knowledge, and there is a procedure for disputing errors. Where safety rails do not exist, then more risk exists for that data or knowledge.

Regarding the biometric portion of your query, I would like to respond in some detail. It is an important question.

All biometric data, including genetic data, rises to the level of high sensitivity. As such, WPF proposes that biometrics be designated as a *technology of very high concern*, and be subjected to meaningful safety guardrails. The United States is one of the few countries where biometric technologies have not yet been as pervasively implemented as they have been in other jurisdictions. But it is very unlikely that the United States will fully escape the use of biometrics, as seen in airport biometric entry/exit programs, among other biometrics programs.

Because of the significant risks inherent in the uses of the technology, biometrics—including facial recognition—should be classified as a high-risk technology, and procedural safety protections that are well-tested and understood in other high-risk contexts should be adapted for biometrics and put in place as guardrails.

The guardrails we are proposing are similar to those found in existing safety regulations in the United States and Europe.

Regulatory Safety Structures that Act as Guardrails for Biometric Systems (Facial Recognition)

The protections fall into three key areas: pre-and post-market safety and quality regulations, use controls, and a consumer complaint mechanism.

Pre-and Post Market Safety and Quality Regulations:

The following pre and post-market safety regulations for biometrics are derived from the existing legislative models of RoHS, REACH, and the Chemical Safety for the 21st Century Act (updates U.S. Toxic Substances Control Act) as well as the Fair Credit Reporting Act. Finally, the consumer complaint mechanisms at the CFPB and CDC provide the model for the post-market consumer complaint reporting.

- **Classification: Biometrics would be classified as a “technology of very high concern.”**
- **Applicable to full supply chain:** The regulations would apply to the full supply chain and to any entity that produces, develops, sells, assembles, distributes, installs, and uses biometric systems.

- **ID risks and reporting requirements:** Biometric entities would be required to identify risks in the technology and document and report those risks to the applicable Government body.
- **Testing requirements:** Biometric technologies available for use would be required to be tested and evaluated by NIST for accuracy and bias on a regular basis, at a minimum, this review would be updated annually.
- **Proven safe prior to launch:** The technology must be proven safe and fit for purpose prior to launch, and must be cleared for market by the appropriate Government oversight body. For facial recognition, a nondiscrimination analysis would need to be performed.
- **Product labeling:** The biometric product would be labeled for accuracy and for bias. (Facial recognition.)
- **Certification and training requirements would apply.**
- **Ongoing monitoring:** The full supply chain of vendors and implementors must agree to ongoing monitoring and documentation for compliance. Monitoring can be in real time, or near real time.

Use controls:

Biometric technology is deployed in specific use cases. Some use cases are not objectionable, however, some uses cases are objectionable and pose threats of discriminatory impact or other harms.

- Some use cases of biometrics would not be allowed due to safety considerations, or lack of functionality. For example, body cameras equipped with real-time facial recognition are viewed by biometricians and a majority of law enforcement as a high-risk use case. This particular use case has both legal and technical problems.
- Allowed use cases would have significant definitional controls and procedural requirements. For example, biometrics used in law enforcement investigatory settings would be subject to the procedures set forth at the Federal level. At the State level, the Bureau of Justice Assistance procedures for biometrics use, for example, could be required (<https://www.bja.gov/Publications/Face-Recognition-Policy-Development-Template-508-compliant.pdf>.)
- Voluntary Consensus Standards could be used in conjunction with legislation to establish ongoing multistakeholder evaluation of emerging use cases.

Post-Market Consumer Complaint Reporting:

- Voluntary Consensus Standards could be used in conjunction with legislation to eUsing the adverse event reporting model and the consumer complaint model, biometrics technologies would have a dedicated post-market monitoring mechanism at the Federal level.
- Consumers and others would be able to submit complaints to a central structure.

- As with the structure of the existing Consumer Financial Protection Bureau (CFPB) consumer complaints database, complaints would be available for viewing within a matter of a week, and the complaints would be available for download and analysis. This data will provide ongoing insight into problem areas and detailed implementation feedback.

Key Underlying Safety Statutes

RoHS: EU Directive, also implemented in some U.S. States.

- As of July 2019 all RoHS deadlines active; Directive is now applicable to any business that sells electrical or electronic products, equipment, sub-assemblies, cables, components, or spare parts directly to RoHS-directed countries, or sells to resellers, distributors or integrators that in turn sell products to these countries, is impacted if they utilize any of the restricted 10 substances.
- Requires products to be cleared for market prior to launch and meaningful compliance documentation/recordkeeping from all parties in the supply chain, regularly updated information, mandatory compliance labeling.
- In the United States, California, Colorado, Illinois, Indiana, Minnesota, New Mexico, New York, Rhode Island, and Wisconsin have enacted RoHS-like and e-waste regulations.

REACH: EU Regulation

- Applies to essentially every product manufactured, imported, or sold within the EU.
- REACH regulates chemical substances, particularly those known as Substances of Very High Concern (SVHC). Substances considered carcinogenic, mutagenic, toxic for reproduction, or bioaccumulative fall under SVHC criteria.
- EU manufacturers and importers are required to register all substances produced above a set yearly volume to:
- ID risks associated with the substances they produce.
- Demonstrate compliance in mitigating the risks to ECHA.
- Establish safe use guidelines for their product so that the use of the substance does not pose a health threat.

Chemical Safety for the 21st Century Act: United States, Federal

- Requires pre-manufacture notification for new chemical substances prior to manufacture.
- Where risks are found, requires testing by manufacturers, importers, and processors
- Requirements for certification compliance
- Reporting and record keeping requirements
- Requirement that any person manufacturing (including imports), processes, or distributes in commerce a chemical substance or mixture and who obtains information which reasonably supports the conclusion that such substance or mixture presents a substantial risk of injury to health or the environment to immediately inform EPA, except where EPA has been

adequately informed of such information. (The EPA screens all TSCA b§ 8(e) submissions.)

ADDITIONAL MATERIAL SUPPLIED FOR THE RECORD



301 Dave Ward Drive, Conway, AR 72032
501-342-1000 www.axiom.com

July 1, 2019

Chairman Mike Crapo
U.S. Senate Committee on Banking, Housing, and Urban Affairs
534 Dirksen Senate Office Building
Washington, D.C. 20510

Ranking Member Sherrod Brown
U.S. Senate Committee on Banking, Housing, and Urban Affairs
534 Dirksen Senate Office Building
Washington, D.C. 20510

Dear Chairman Crapo and Ranking Member Brown:

I write to you as the Chief Data Ethics Officer of Axiom. We followed your Committee's June 11, hearing entitled "Data Brokers and the Impact on Financial Data Privacy, Credit, Insurance, Employment and Housing." While we were not surprised to hear our name mentioned in the context of the hearing, we were surprised to hear it mentioned in the context of having declined an invitation to testify since, to our knowledge, we did not receive an invitation.

At the outset, please know Axiom supports passage of a strong and balanced national privacy law and is committed to helping Congress develop one that will protect consumers and allow businesses to continue to use information in a responsible and ethical manner.

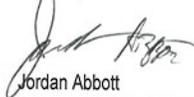
Axiom is proud of the work we do. We provide the data, technology, and services needed to power exceptional customer experiences everywhere. These services, in turn, power the economy in exceptional ways, such as improving the accuracy of data; helping businesses direct their goods and services to customers who are interested; and helping to detect and prevent fraud. Axiom provides these services to for-profit private sector companies, non-profits, and United States government agencies.

Axiom has testified many times on Capitol Hill regarding our business and we have met with numerous Congressional leaders and staff over the years to talk about what we do. We also regularly meet with agencies such as the FTC and other regulators to help them also better understand our business. We always welcome the opportunity to make sure that those who are interested in the work we do are also well-informed about how we do it.



We would be happy to bring an Acxiom executive team to Washington to meet with you and discuss our business with you in greater detail. As you might expect, we have ideas for the framework of a federal privacy law that are designed to benefit both consumers and businesses.

Sincerely,



Jordan Abbott
Chief Data Ethics Officer
jordan.abbott@acxiom.com
(501) 342-0356



June 27, 2019

Chairman Mike Crapo
U.S. Senate Committee on Banking, Housing, and Urban Affairs
534 Dirksen Senate Office Building
Washington, D.C. 20510

Ranking Member Sherrod Brown
U.S. Senate Committee on Banking, Housing, and Urban Affairs
534 Dirksen Senate Office Building
Washington, D.C. 20510

Dear Chairman Crapo and Ranking Member Brown:

This letter responds to your June 11, 2019 letter to the Association of National Advertisers (“ANA”) regarding data collection in the digital economy. We appreciate the opportunity to provide the Committee with information about how ANA members are involved in these issues.

ANA makes a difference for individuals, brands, and the advertising industry by driving growth, advancing the interests of marketers, and promoting and protecting the well-being of the marketing community. Founded in 1910, the ANA provides leadership that advances marketing excellence and shapes the future of the industry. The ANA’s membership includes more than 1,850 companies and organizations with 20,000 brands that engage almost 50,000 industry professionals and collectively spend or support more than \$400 billion in marketing and advertising annually. The work of the nonprofit ANA Educational Foundation (AEF), which has the mission of enhancing the understanding of advertising and marketing within the academic and marketing communities further enriches this ecosystem.

We understand that the Committee’s particular focus is on “data brokers.” ANA has over 1,100 client-side marketers and more than 750 marketing solutions provider members. These include leading marketing data science and technology suppliers, ad agencies, law firms, consultants, and vendors. While “data brokers” does not have a simple or clearly defined definition, to the extent the Committee deems any of these organizations “data brokers,” ANA has insight into the data broker “industry” by virtue of our representative position in regard to these marketing solutions providers. Our representation, however, is solely focused on these companies’ advertising functions, which do not include all of the functions that “data brokers” may undertake.

We do wish to offer some additional context in response to your letter. First, as you know, more than six years ago the Federal Trade Commission (“FTC”) and Congress each undertook a thorough examination of data brokers’ information practices through a series of inquiries about such companies’ business models, data collection activities, and use of consumer information. Those examinations resulted in detailed, official reports and findings that described

the data broker business and proposed suggestions for the industry.¹ We note that much of the information set forth in our responses below is therefore likely not new as the government has already systematically investigated the data broker industry.

Second, ANA recognizes that the advertising and marketing industry is in need of comprehensive federal standards delineating reasonable data practices across the spectrum of all marketing data uses. To further this end, ANA has helped launch the Privacy for America coalition, a group dedicated to advancing a new federal data privacy paradigm that provides clear rules of the road for businesses and defines acceptable and unacceptable data uses and activities. We do not think a sector by sector approach any longer is adequate or appropriate. ANA and its members, including those who have been labeled “data brokers”, look forward to working with the Congress to craft a federal data privacy framework that clarifies beneficial and harmful data uses for all businesses that touch consumer information. We believe this more encompassing approach is the most productive way to handle data privacy and security issues rather than the existing sector by sector or category by category regulatory regime.

Before answering the questions that the Committee submitted to us, it is important to emphasize again that, to the extent ANA represents data brokers, it does so with respect to these entities’ marketing and other activities that fall outside of the Fair Credit Reporting Act (“FCRA”) activities. Some ANA members may have divisions, affiliates, or business lines that operate as consumer reporting agencies regulated under the FCRA, but ANA does not represent these specific types of activities or efforts. ANA, therefore, does not have the relevant information to provide to the Committee regarding the practices of consumer reporting agencies that collect and use information for these FCRA purposes.

Questions and Responses

- 1) **Do data brokers have any information bearing on an individual’s (or group of individuals’) creditworthiness, credit standing, credit capacity, character, general reputation, personal characteristics or mode of living that is used (either by the data broker or any unaffiliated third party) to establish eligibility for, or in the marketing of, a product or service related to (1) credit, (2) insurance, (3) employment, or (4) housing?**

¹ See, e.g., FTC, *Data Brokers, A Call for Transparency and Accountability* (May 2014), located at <https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf>; Senate Committee on Commerce, Science, and Transportation Office of Oversight and Investigations, *Staff Report for Chairman Rockefeller, A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes* (Dec. 18, 2013), located at <https://www.commerce.senate.gov/public/cache/files/0d2b3642-6221-4888-a631-08f2f255b577/AE5D72CBE7F44F5BFC846BECE22C875B.12.18.13-senate-commerce-committee-report-on-data-broker-industry.pdf>

Data brokers may use this type information to assist their clients in marketing products and services and providing relevant offerings to individuals that best suit their needs and desires. However, data brokers are prohibited by industry self-regulatory programs, that for covered entities is also enforceable by the FTC, from using such data for any of the listed or other eligibility purposes.

The ANA Guidelines for Ethical Business Practice expressly prohibit entities subject to the Guidelines from using marketing data for eligibility for employment, credit, and health care treatment, and for insurance and underwriting.² The Digital Advertising Alliance's ("DAA") self-regulatory program, which was founded by a number of industry leaders, including ANA, sets forth a similar set of enforceable principles regarding the use of data for eligibility purposes.³ As further described in Question 5 below, ANA and the Privacy for America coalition envision a federal data privacy paradigm that would leverage the successes of self-regulation and craft a comprehensive, non-sector specific standard that defines appropriate and inappropriate data uses and practices.

2) How do data brokers ensure that information bearing on an individual's (or a group of individuals') creditworthiness, credit standing, credit capacity, character, general reputation, personal characteristics or mode of living is not used in violation of the Fair Credit Reporting Act?

As noted above, ANA does not represent companies or organizations acting as consumer reporting agencies and regulated by the FCRA. Rather, ANA data broker members focus on data marketing activities, which are wholly distinct from FCRA-covered activities. ANA represents marketing solutions providers, and these entities use certain mechanisms to separate information that is subject to the FCRA from information that can be used for marketing purposes. These organizations may implement a range of technical and organizational measures, such as firewalls and internal use limitations, to keep FCRA data separate from data that can be used for marketing. This tactic of maintaining FCRA data and non-FCRA data in silos assists data brokers in meeting their FCRA obligations.

Data brokers also may apply certain contractual strategies to help restrict the use of marketing data they provide to companies by stipulating that the data cannot be used in violation of the FCRA. These contractual terms deter data recipients from using data in ways that would abridge relevant federal laws.

When data is used for marketing purposes (and not FCRA-purposes), the potential consequences of inaccurate information are generally limited to irrelevant advertising. Because

² The Guidelines were formerly known as the Data & Marketing Association Guidelines for Ethical Business Practice before ANA merged with the Data & Marketing Association. Association of National Advertisers, Guidelines for Ethical Business Practice, Part I, § 5 (2018), located at <https://thedma.org/accountability/ethics-and-compliance/dma-ethical-guidelines/>.

³ See, e.g., DAA, *Self-Regulatory Principles for Multi-Site Data* 4-5 (Nov. 2011), located at <https://digitaladvertisingalliance.org/principles>.

these potential consumer consequences are limited, the FTC has noted that it is not necessary to provide consumers with the ability to access and correct marketing data or to take special steps to ensure the accuracy of such data.⁴

3) To what extent are data brokers covered by the European Union’s General Data Protection Regulation (GDPR) and how has the data broker industry reacted since GDPR has become effective, including changes made to data privacy practices and policies?

The European Union’s (“EU”) General Data Protection Regulation (“GDPR”) covers any data broker that has an establishment in the EU, or if a particular data broker is not established in the EU, the GDPR still applies to the entity if its data processing activities are related to offering goods and services to EU data subjects or if the entity “monitors” such data subjects’ behavior, as long as the behavior of the data subjects takes place within the EU.⁵

Data brokers are not uniquely regulated under the GDPR; if a data broker is acting as a data “controller,” as that term is defined in the regulation, it is subject to the same requirements as other businesses acting in that role. The GDPR has instituted certain requirements with which all data controllers must comply, such as additional privacy policy disclosures, transparency requirements, and the need to obtain the data subject’s consent, where required by law.

Notably, the GDPR requires controllers to disclose “from which source the personal data originate, and if applicable, whether it came from publicly accessible sources” in privacy policies.⁶ Additionally and of particular relevance to data brokers, all data controllers must provide certain disclosures to data subjects when personal information was not obtained from the data subject originally.⁷ Such disclosures include: “the purposes of the processing for which the personal data are intended as well as the legal basis for processing”; “the recipients or categories of recipients of the personal data”; and “the contact details of the [data broker’s] data protection officer.”⁸ On the whole, it is our impression that covered data brokers have reacted to the GDPR by updating their business practices and consumer-facing disclosures to account for the law’s requirements.

⁴ Federal Trade Commission, *Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers* 30, 65 (March 2012), located at <https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf>.

⁵ GDPR, Article 3.

⁶ GDPR, Article 14(2)(f).

⁷ GDPR, Article 14.

⁸ GDPR, Article 14(1).

4) What safeguards do data brokers put in place to protect individuals' data?

Data brokers employ a variety of administrative, technical, and physical safeguards to protect individuals' data from unauthorized access and use. Tactics may range from encryption mechanisms and network segmentation to facility access controls and workstation security mechanisms. A number of states, such as Massachusetts and Nevada, have enacted data security laws and regulations that impose specific requirements on organizations that maintain consumer data.⁹ Data brokers operating in such states must comply with those laws by instituting measures to reasonably secure the consumer data they hold.

Data brokers also may anonymize or pseudonymize individuals' information so that the data a broker maintains is not linked to an individual's personally identifiable information. Data brokers may do so by "hashing" the consumer data they collect, assigning it an arbitrary alphanumeric code that does not provide any insight into information beneath the random code. This practice is privacy enhancing because it disassociates the underlying information from the individual consumer, thereby allowing businesses to analyze information in the aggregate to help make decisions that benefit consumers and the economy alike. ANA, however, does not have company specific information on these various privacy approaches.

5) What steps should Congress take to:

- a) Ensure individuals are more informed about the collection, sharing, or use of their data by data brokers;**
- b) Give individuals access to the data collected about them by data brokers; and**
- c) Clearly provide individuals the opportunity to correct inaccuracies in their data held, used or shared by data brokers, or to opt out of data brokers sharing their data with others for use in marketing, including opting out seamlessly across data brokers holding the same or similar data?**

In ANA's view, Congress should pass a federal law that provides all Americans with strong and effective data privacy protections and defines acceptable and unacceptable data practices for all businesses that deal with consumer data, including data brokers. Through our work with Privacy for America, we have advocated for such a federal paradigm to set forth comprehensive, clear, and enforceable privacy rules for businesses with respect to their data practices and authorize strict penalties for violations.¹⁰ ANA believes consumers should have strong privacy protections wherever they are located. Geographic location should not determine the level of privacy protection for American citizens.

⁹ Mass. Gen. Laws ch. 93H; 201 Mass. Code Regs. 17; Nev. Rev. Stat. § 603A.

¹⁰ Privacy for America, *Creating a Strong New Paradigm for Privacy and Responsible Data Use*, located at <https://www.privacyforamerica.com/overview/>.

Currently, laws and practices in the United States place the onus on consumers to read numerous and extensive privacy policies and make choices regarding the use of their data by each company that has access to it—including those with whom they do not even interact. This structure places a burden on consumers to be constantly vigilant evaluators of businesses' privacy practices and to take action to limit practices they do not want to apply to them or their data. The GDPR, with its opt-in privacy approach, creates similar, or even broader, requirements that already have been criticized for creating privacy "notice fatigue" for consumers. Privacy for America's new paradigm would shift the burden away from consumers to enhancing only appropriate business practices by establishing a strong national privacy standard backed by enforcement and stiff penalties for those who do not comply.

Furthermore, Privacy for America's new paradigm would: (1) prohibit or limit the use of data for eligibility, discriminatory, or fraudulent purposes, and limit the use of "sensitive data"—data including medical, biometric, financial, and geolocation information as well as email communications and private recordings; (2) strengthen privacy oversight and enforcement by creating a new Data Protection Bureau within the FTC and providing it with additional privacy staff, resources, rulemaking authority, and jurisdiction; (3) help ensure responsible advertising practices by imposing significant restrictions on data used for advertising, banning certain data from being used altogether, and allowing consumers to identify their preferences regarding what advertising they want to receive or not receive; and (4) require strong data security protections to guard against data breaches.¹¹ ANA looks forward to working with the Congress to identify a clear, nationwide standard that protects consumer data by defining prohibited and acceptable business data practices and uses.

ANA will be glad to discuss in more detail these issues at any time. Please feel free to contact me or our Group EVP Dan Jaffe at djaffe@ana.net or at 202.296.2359

Respectfully submitted,



Bob Liodice
Chief Executive Officer
Association of National Advertisers

¹¹ Privacy for America, *New "Privacy for America" Coalition Calls for Strong Data Privacy Protections for All Americans* (April 8, 2019), located at <https://www.privacyforamerica.com/new-privacy-for-america-coalition/>.



June 11, 2019

The Honorable Mike Crapo
Chairman

The Honorable Sherrod Brown
Ranking Member

US Senate Committee on Banking, Housing, and Urban Affairs
534 Dirksen Senate Office Building
Washington, D.C. 20510

RE: Hearing on Data Brokers and the Impact on Financial Data Privacy, Credit, Insurance, Employment and Housing

Dear Chairman Crapo and Ranking Member Brown

I write to you to clarify a point made during today's hearing regarding CoreLogic and to respectfully request that this letter be included in the record.

We consulted multiple times with majority committee staff about how our knowledge, science and expertise might contribute to this hearing. As requested during these prior consultations we held multiple dates open in order to ensure that we would be available to testify if our contribution was determined to be of value.

Through these consultations it became clear that the issues that were a focus of today's hearing were not ones related to our company, which is why we were not invited to testify.

Sincerely,

A handwritten signature in black ink, appearing to read "Stuart K. Pratt".

Stuart K. Pratt
Global Head
Public Policy & Industry Relations

Cc: Members of the Senate Banking Committee



Jim Nussle
President & CEO

Phone: 202-508-6745
jnussle@cuna.coop

99 M Street SE
Suite 300
Washington, DC 20003-3799

June 10, 2019

The Honorable Mike Crapo
Chairman
Committee on Banking, Housing
and Urban Affairs
United States Senate
Washington, DC 20510

The Honorable Sherrod Brown
Ranking Member
Committee on Banking, Housing,
and Urban Affairs
United States Senate
Washington, DC 20510

Dear Chairman Crapo and Ranking Member Brown:

On behalf of American's credit unions, I am writing to express our views ahead of the hearing titled "Data Brokers and the Impact on Financial Data Privacy, Credit, Insurance, Employment and Housing." The Credit Union National Association (CUNA) represents America's credit unions and their 115 million members.

Safeguarding consumers' money and personal information is the bedrock of the financial services industry and has been for a long time. In order to meet requirements of many different laws and regulations, financial services companies store and collect many different types of consumer information. The Gram-Leach-Bliley Act (GLBA) sets forth data security and privacy laws for credit unions and other financial institutions. Other sectors are also subject to Federal requirements such as the Health Insurance Portability and Accountability Act (HIPAA) for healthcare providers.

HIPAA and GLBA have been in place for 20 years and reflect the importance of privacy and data security requirements for service businesses that must collect and store information in order to provide necessary services. Since these sector specific laws were passed by Congress, much has changed in the economy. There are many more businesses now that collect, aggregate, and sell Americans' most personal information; even traditional businesses like retailers have found value in collecting and analyzing their customers' data.

Since Americans' personal information has become so valuable in the aggregate to businesses and criminals worldwide, the time has come for new Federal protections regulating the use and security of data held by all businesses and entities. Europe's General Data Protection Regulation (GDPR) and California's California Consumer Privacy Act (CCPA) show that foreign governments and states are not willing to sit on the sidelines and neither should Congress. Action is required to ensure that all Americans can enjoy robust protection of their most important personal data from misuse and theft.

The current gaps in data protection and privacy laws hurt consumers and businesses as information is misused by criminals and other actors with malicious intent. Financial institutions are at the vanguard for misuse of stolen data. Although data security is a major issue for credit unions, we realize the problem is much bigger than the financial services industry with robust privacy and data security requirements for all industries becoming increasingly necessary.

The cornerstone of any new privacy requirements should be robust data security requirements for business and other entities that collect consumers' personal information. The current patchwork of laws is complex even at the Federal level. For example, federally regulated depository institutions are subject to data security requirements promulgated by each entity's prudential regulator and subject to privacy requirements promulgated

cuna.org

by the Consumer Financial Protection Bureau (CFPB) even though GLBA is the implementing law for both. Companies such as Equifax follow the Federal Trade Commission's (FTC) Safeguards rule and the FTC further uses UDAAP to enforce data security and privacy requirements for entities not subject to specific requirements. Layering additional state laws onto these rules creates complex challenges for compliance which is challenging for the largest of businesses and nearly impossible for smaller businesses.

Although GLBA has served the financial services industry well, Congress must work with the Administration and industry to finally address consumer data privacy in a meaningful way. To that end,

- Any new privacy law should cover both privacy and data security. There cannot be privacy of data without protection from loss due to breach or other types of theft.
- The law should cover all institutions, not just tech companies, credit-rating agencies, and other narrow sectors of the economy. Any company that collects, uses or shares personal data or information has the opportunity to misuse the data or lose the data through breach.
- Data security requirements should be based upon protection of data to prevent theft and misuse. Notification or disclosure after the fact are important but are not the stopping point for adequate protection. By the time a breach is disclosed, harm could already have befallen hundreds of thousands, if not millions, of individuals, so robust protection is paramount for any new requirements.
- A law should provide mechanisms to address the harms that result from privacy violations and security violations, including data breach. Increasingly courts are recognizing rights of action for individuals and companies (including credit unions). However, individuals and companies should be afforded a private right of action to hold those that violate the law accountable, and regulators should have the ability to take action against entities that violate the law.
- Any new law should preempt state requirements to simplify compliance and create equal expectation and protection for all consumers. Just like moving away from the sector specific approach, the goal should be to create a national standard for all to follow.

On behalf of America's credit unions and their 115 million members, thank you for holding this important hearing.

Sincerely,



Jim Nussle
President & CEO



3138 10th Street North
Arlington, VA 22201-2149
703.522.4770 | 800.336.4644
f: 703.524.1082
nafcu@nafcu.org | nafcu.org

National Association of Federally-Insured Credit Unions

June 11, 2019

The Honorable Michael Crapo
Chairman
Committee on Banking, Housing
& Urban Affairs
United States Senate
Washington, DC 20510

The Honorable Sherrod Brown
Ranking Member
Committee on Banking, Housing
& Urban Affairs
United States Senate
Washington, DC 20510

Re: Today's Hearing: Data Brokers and the Impact on Financial Data Privacy, Credit, Insurance, Employment and Housing

Dear Chairman Crapo and Ranking Member Brown:

I write to you today on behalf of the National Association of Federally-Insured Credit Unions (NAFCU) in conjunction with today's hearing, entitled "Data Brokers and the Impact on Financial Data Privacy, Credit, Insurance, Employment and Housing." NAFCU advocates for all federally-insured not-for-profit credit unions that, in turn, serve over 117 million consumers with personal and small business financial service products. NAFCU and our members welcome the Committee taking this next step in examining consumer privacy and data security standards by holding this hearing.

As NAFCU wrote to the Committee on May 6, 2019, we believe there is an urgent need for a national data security standard for those who collect and store consumer information. While depository institutions have had a national standard on data security since the passage of the *Gramm-Leach-Bliley Act* (GLBA) over two decades ago, other entities who handle consumer financial data do not have such a national standard. Along those same lines, we also believe that there is a need for a uniform national consumer data privacy standard as opposed to a patchwork of standards stemming from different state data privacy laws. We hope today's hearing can be another step toward achieving these goals.

NAFCU looks forward to working with the Committee to address these concerns with consumer privacy and data security. We are also pleased to work with those in industry to try to find common ground on a comprehensive proposal. We would urge you to work collaboratively with other interested Committees in the Senate to find a package that can advance and receive bipartisan support.

On behalf of our nation's credit unions and their more than 117 million members, we thank you for your attention to this important matter. Should you have any questions or require any additional information, please contact me or Janelle Relfe, NAFCU's Associate Director of Legislative Affairs, at 703-842-2237 or jrelfe@nafcu.org.

Sincerely,

Brad Thaler
Vice President of Legislative Affairs

cc: Members of the Senate Banking Committee