THE BROOKINGS INSTITUTION

SOCIAL PHYSICS:
HOW HUMAN SOCIAL NETWORKS SPREAD IDEAS

Washington, D.C.

Tuesday, February 18, 2014

PARTICIPANTS:

DARRELL WEST
Founding Director, Center for Technology Innovation
Vice President and Director, Government Studies
The Brookings Institution

ALEX “SANDY” PENTLAND
Director, Human Dynamics Laboratory
Director, Media Lab Entrepreneurship Program
Massachusetts Institute of Technology

CAMERON F. KERRY
Andrew H. Tisch Distinguished Visiting Fellow,
The Brookings Institution
Visiting Scholar, The MIT Media Lab

ANEESH CHOPRA
Co-Founder and Executive Vice President, Hunch Analytics
Senior Advisor, Technology Strategy, The Advisory Board Company

* * * * *

P R O C E E D I N G S

MR. KERRY: Well, good afternoon and welcome. I'm Cameron Kerry. I'm the Ann R. and Andrew H. Distinguished Visiting Fellow, and I'm pleased to welcome you and to discuss Sandy Pentland's book, *Social Physics: How New Ideas -- How Social Networks Spread New Ideas*.

We'll have a brief discussion ahead of time, and then we'll welcome any questions or comments that you have.

We'll also have a Twitter feed with the #TechCTI if you want to post comments on this, and we'll also take questions from the virtual and live audiences.

Let me, at this point, remind you to please turn off cell phones, which can interfere with the communications that we're using here to record this so that it will be on the Brookings web site afterward.

So, to set the stage a little bit, social physics is a quantitative science -- a quantitative science that develops mathematical equations to look at information flow and data flow, and to understand how that moves from person to person, how it moves through social networks, and how that movement shapes norms, shapes ideas, shapes human behavior, shapes creative output.

This science provides insight into human behavior, organization behavior, to decision-making across disciplines -- health, finance, politics, city planning -- and enables the tracking of social interaction in ways that can help public policy, can help individual decision-making, can help organizations.

And we're fortunate to have with us today one of the fathers of this science. Sandy Pentland is the Toshiba professor of media arts and sciences at MIT. He directs the Human Dynamics Lab at the MIT Media Lab.

His background is in quantitative science and, like the best polymaths,

his writing deals with economics, with philosophy, with human behavior and psychology.

And he has really led an extraordinary range of work, from the World Economic Forum's work on big data to advisory boards for Nissan, for Motorola, mobility for Telefonica, and the research and work of a great many start-ups. In fact, one of our other panelists here, Aneesh Chopra, who I will introduce in a moment, said on arriving, you know, all the people that I've been working with are people who study with Sandy Pentland.

So it's not surprising that *Forbes* magazine named him one of seven most significant data scientists in the world. He was recently also named to the advisory board of the *Harvard Business Review*, but I'm sure it is a coincidence that the *Harvard Business Review* recently said that data science is one of the sexiest jobs in the world.

We are fortunate to have, to discuss social physics and Sandy's book, a terrific panel of experts.

I mentioned Aneesh Chopra. Aneesh is now the co-founder and executive vice president of Hunch Analytics. It's a start-up that's focusing on enhancing the productivity of public and regulatory sectors of the economy, using data analytics. That is where he was a leader in the federal government, as the United States first U.S. chief technology officer and assistant to the president and deputy director of the Office of Science and Technology Policy, and really led the Obama administration's work on open innovation and open data sets.

He is continuing that work with a forthcoming book on the *Innovative State: How Technologies Can Transform Government* -- so, something that is very much on point to the work that we are talking about today.

And we'll also be joined by Brookings' own vice president, Darrell West, who heads the Governance Studies program and is the founding director of the Center

for Technology Innovation here at the Brookings Institution.

Darrell's research and work has focused on technology, on mass media, on campaign reform and on technology in government, and his writing has been prolific. He has written 18 books but perhaps, most notably, won the Don K. Price Award for the best book on technology -- *Digital Government*. He has really been a leading researcher use of e-government in his work, both here at Brookings and before joining Brookings as a professor at Brown University.

So Sandy will make a brief presentation, we'll then have a discussion among the panel, and we'll then open to all of you for your questions.

So, look forward to the discussion and to your questions.

Sandy.

(Applause.)

MR. PENTLAND: Thank you for the very kind introduction.

Let's see. Are we going to have a screen here?

So social physics is a very old name. It's over two centuries old. It's a name that was invented by the father of sociology, and he had it that ideas were the things that shaped culture and shaped our society.

But, like most of the social science, political science, medical science that has happened in the last centuries since then, we haven't had a lot of data. It's mostly been freshmen filling out surveys, or people stopping folks on the corner and asking questions.

Even the big studies, like the Framingham Heart Study, where they talked to tens of thousands of people for decades, really only asked one question per month per person. So there's a lot of context that was missing.

And so what has happened is that as big data -- things off of cell phones,

off of badges, off of credit cards -- become available, it's a little bit scary for society, but it's a gold mine for social science.

And that's the thing that I thought it was worthwhile calling this "Social Physics." It's what happens when social science meets big data.

For the very first time, you can actually get enough data off enough people, continuously, fine grain enough, that you can tell which theories are in action at what point; you can build mathematical predictive models of it.

One very senior sociologist said to me, this is like the transition between alchemy and chemistry.

In alchemy, people were really pretty smart. Newton was an alchemist. But, while they knew about many things, they didn't have organizing principles quite simply because they didn't really have enough systematic data.

And the same is happening in the social sciences today.

What I want to do is not sort of take you through a whole tour of that because we don't have enough time.

And I should also apologize that I have a 4:30 flight, by accident. So about 3:15 I'm going to become very -- sorry, 4:15, I'm going to become very -- 3:15, I'm going to become very itchy.

PARTICIPANT: At 4:15, you would get really nervous.

MR. PENTLAND: 4:15, I'll get like super nervous, yes. Okay.

But let me show you a couple of the things that I think are most relevant to government.

So this is an example of something that is in Germany. It's in a bank. It's an organization that is creating a new campaign around a new mortgage product.

And the little red dots, where you can't see the labels -- those are

different departments in this bank.

And what you're seeing here in the blue is all the e-mail that goes back and forth between the departments, and the red is something you've never seen before, which is all the face-to-face communications.

Now there's no content in here. This is all metadata -- you know, thinking of the NSA and things like that. But what you can see is which groups are talking to each other.

As they go through this planning exercise, nobody talks to customer service. And, when they release the product, it's a disaster, and they have all-day meetings with customer service to try and figure out what to do.

So, in our looking at dozens and dozens of organizations making things like this, what we found is there is a couple of patterns that really dominate organizations' productivity and creative output.

And they're very simple patterns. They're patterns that your grandmother could have told you, but then your grandmother might have told you many other ideas, too.

What we find is this big data tell us certain things are really, in fact, the case.

One is if you look at the pattern of communication within this organization, there is something we call engagement, which is, do the people within a group talk to each other? If I talk to you and I talk to you, do you two talk to each other?

It's a mathematical quantity that you can actually measure, and the productivity of the group is directly proportional to that. It typically accounts for 30 percent of the variation in productivity. That's huge. It's larger than almost any other factor you can name.

Another factor is what we call exploration, which is people talking to folks that they're not supposed to talk to according to the org chart. It's talking to the sales guys; it's talking to the execs; it's talking to the janitor, to get different perspectives on what's going on in your group. It's bringing new ideas in which then are harvested into new habits by this engagement.

So that is something that's quite amazing. Independent of content, independent of the people, independent of the task, the pattern of communication is the strongest determinant of productivity and creative output in organization after organization, ranging from drug discovery units in big pharma through call centers through banks.

This won an award as Paper of the Year from Harvard Business Review, which I suppose is why they wanted me to be on their board. But it also won an award from the American Academy of Management, and I think it's the first time that the Academy and Harvard Business Review have ever agreed on something. So maybe it actually has some good truth to it.

So this is what we find when we get big data and social sciences together. The same idea turns out to be true in social networks.

So this is a social network. I'm sorry for the sort of quality of this.

It's 1.6 million people, and every place there's a dot, that means somebody is following somebody else on the vertical axis.

So it's the same people on both axes. If A is following B, you get a dot up there.

And there are areas in there without much following at all. So it's very isolated people.

And there are areas where everybody is having a big party, and they're

all following each other.

But this is an interesting social network because you can tell how good people's decisions are because what they're doing is they're day traders. These 1.6 million people are buying and selling euros and dollars, and you can see how much money they make every day.

The answer is that the isolated people don't do very well at all. They're essentially market-neutral.

The people who are having the big party in the echo chamber -- they don't do very well either, and occasionally they get wiped out. That's financial bubbles.

The people in the middle make 30 percent more money than everybody else.

Now this is financial trading, but you can imagine your organization that you work in. You're part of a social network. Part of it is online. Part of it is face-to-face. This principle of having diverse social learning going on, this notion of an engagement and exploration, is the same thing that we see in these things as going on in your organization.

And you can ask, could we do something to make the idea flow in your organization better, to make the decision-making improve by this large amount? And I think the answer is probably yes.

So that's sort of the basic idea, but let me show you some things that are outside of the organization. For instance, this is something that a former student of mine, Nathan Eagle, first noticed.

These are all the UK councils. So these are neighborhood governments.

And along the vertical axis is socioeconomic percentile. So this is a combination of infant mortality, crime rate, GDP and longevity. It's how good the

neighborhood is.

And along the bottom is basically the same thing of exploration and engagement.

So, if you tell me the call pattern, or you tell me the mobility pattern -- so do they get out of their neighborhood? Do they meet with each other in the neighborhood? -- I can tell you how many babies die. And, I can tell you very accurately how many babies die. I can tell you the crime rate. I can tell you the GDP.

People and our cultures, our social fabric, need this flow of ideas and engagement in order to be healthy.

So you can now do something that's like the census. But, you can do it from telephone data. You can do it every day. You can do it essentially for free. That's got to be transformative in terms of transparency.

Here's a map that we made for the Ivory Coast. I helped convince a carrier -- telephone carrier, Orange -- to release all of their data for the country in an aggregate form that protected the privacy of individuals.

But we were able to make a similar sort of map. We could see places where there was extreme poverty even though this is a country that recently had a civil war and the government can't go in many places.

And, when we released this data, we invited researchers from all over the world and come together and look what else we could do with this data.

One research group found a way to change the bus lines in their capital city to reduce transportation time by 10 percent. That's huge in terms of pollution, in terms of lost work and stress.

Another person found a way to improve their public health system by almost 20 percent.

And the key was they never knew where the people were going or where they lived, and just by looking at activity and cell phone towers -- that the people went from this tower to that tower -- you could finally look at the mixing of people. And you can read out poverty, but you can also read out a lot of social systems and improve them rather dramatically.

So here's the other side of that.

We did a survey of some 300 cities -- 150 in the U.S., 150 in Europe. And, if you tell me the density of the city and you tell me about the transportation infrastructure, I can tell you the pattern of interaction between people because it turns out that people behave in a very sort of lawful, standard way regardless of the city.

It really is how much opportunity they have to interact with each other, and that gives you a fairly precise mathematical sense of how often people meet.

But here's the catch, or the punch line; if you tell me the density and the average commute time, I can tell you the GDP extremely precisely. And that is what is shown there. The infrastructure determines the GDP.

I can also tell you the crime rate, the patenting rate and other sorts of phenomena, just from the very simple thing of the frequency with which different communities interact with each other.

So something, again, that your grandmother might have told you is that ideas banging together creates innovation. And innovation determines the GDP and other characteristics of society seems to be true now that we have this big data.

Notice that I did not talk about classes and access to means of production. I did not talk about specialization. I did not even talk about education -- not that these things don't affect these other quantities like GDP and crime, but they seem to be definitely secondary effects.

The main thing is getting different communities to bang together. The diversity we saw in social networks gave greater GDP, greater return on investment, greater GDP in cities.

So how do we use all of this?

Well, that's the big data thing. So we're moving from research now to how do we actually live.

And, as I'm sure many people have heard, big data has many different ways to mislead you because there are lots of false correlations when you get that much data that's that rich. And so the question is, what to do?

Also, just because I have these relationships, we have to think about the ethical dimension of: Do we want to live using these types of relationships? Do we want to change the way we govern ourselves?

And so what I've been doing is setting up what we call living labs. For instance, this one is one in Trento. It's in Italy. It's an autonomous region.

And what happened there is the telephone company and the municipal government and the people agreed to try and live under a different privacy rule than the rest of Italy, than the rest of the EU -- one that would promote greater sharing of ideas by reducing the risk of sharing ideas. It does this through a method that's called a personal data store where you control information about you, not a company, not the government.

What we're trying to do is find out -- does this indeed raise the innovation level within the city, and is it actually safe? Is it something that people want? -- because it's only by trying things out on the ground that we're going to figure out how to go forward in this new world.

We don't really know -- even though we're discovering these new scientific facts, we don't really know how to translate that into governance. And I think

what we need to do is we need to do more on-the-ground experiments where people opt in to trying new ways of living, and we learn from each other.

So, currently, I've done this in Italy. We've done it in Ivory Coast. We're doing it in Paraguay. We've done it in Cambridge, Massachusetts. And we're making learnings about what works and doesn't work on real people in real situations.

And I'd like to see lots of cities begin doing this so that we can figure out how to use this big data and some of these new scientific insights to build a much more innovative, supportive and safe society.

And I'm going to stop there.

(Interruption)

MR. KERRY: Well, Sandy, thank you. You've laid out, I think, some of the principles of social physics, some of the ways that it can be used to help fashion policies.

Let me ask Aneesh to talk about, from your experience, both what you're doing now and what you did in the government, what you see as the uses of social physics -- you know, how do we take what Sandy has put together and make it work for public policy for government?

MR. CHOPRA: Well, thank you very much.

And I want to admit in the beginning that I'm a card-carrying member of the fan club for Professor Pentland. I do think there's a profound set of opportunities for how we solve problems today in the public sector, embracing and engaging in a lot of the concepts behind social physics.

Perhaps I might share my response to Cam's question in three ways: One is by providing some context for how open data is part of our emerging policymaking tool kit; second, to dive deeper on one aspect of open data that's having some particular

impact around this personal data and the control that individuals have on its use; and then, last but not least, how it actually can solve problems if we can piece together these concepts with other policy tools that are emerging.

So let me just take a couple minutes on that to set the stage.

First, in terms of context, the good news is that Washington, in almost a bipartisan fashion, is really heading in the direction towards opening up data held by, or regulated by, the government for the betterment of society. President Obama, on day one, issued an open government memorandum, and he asked all of the agencies to begin liberating data assets that the government held.

Previously, they might have been available if you FOIA or you requested this, but the question was, can we flip the default and actually make it available and do so in a format that would allow entrepreneurs and innovators to take it and make use of it.

One small example. We have a database of all the low-cost health insurance clinics in America. But to visit that database you have to go to HHS.gov and find the clinic locator, and not very many Americans do that.

But we made that data set available, and an entrepreneur that built an iPhone app that basically said don't go to the ER if you don't need to and find alternative places to get care, in the context of that app, having access to that data was useful.

So I triaged the app.

When it came to its conclusion that you don't need to go to the emergency room, you might want to look for a local clinic, it now populates a list of all the low-cost clinics. And without any cost to the taxpayer and without any real burden on the entrepreneur's side, they generated over 100,000 referrals to low-cost clinics just by taking that data set and incorporating it.

So that's an example of how open data is starting to take hold in the

policymaking arena.

I say it's bipartisan because both sides of the aisle have embraced the idea. Even the House Republicans have now made open data a pledge, and they're releasing data that they have on congressional activity that would be useful for entrepreneurs.

So the point of view is that there seems to be some emerging consensus that such liberation will be useful. It could be useful to solve problems or, just in general terms, useful to create economic wealth.

McKinsey recently estimated that open data, not just in the government's context but at large, could unleash three to five trillion dollars of economic value, annually.

The second piece, though, is the particular benefits of the very sensitive data, the personal data that Sandy had described in his talk. And you're starting to see an emerging culture where about a decade ago, with privacy policy, that answer was always no.

Can I do something with health care? No.

Can I do something with education data? No.

In fact, even I was in the governor's cabinet in Virginia, we were trying to help understand the business case for universal pre-K, and we needed to show that kids that did enter into the universal pre-K did better on third grade reading scores.

But you can't link the two data sets, and with FERPA -- the federal records privacy act for educational purposes -- you literally couldn't link the data bases on enrollment statistics in pre-K with educational performance. So states have been trying to create workarounds.

But what's happening now in this second bucket is that individuals are

actually taking possession of their own data, and it's their choice if they want to make it available to someone else.

In my prior life, we started a movement toward these colored buttons. If you wanted access to your own health care data, push a blue button, and you can download it from the place where it's stored today and take and share it with whomever you want.

The VA, for example, has over a million veterans now push that blue button to download it and to share it with caregivers to make sure that they know all the information about the patient before they treat them.

And education has moved on with its version of a "my data" button, and energy has got a green button, and you are going to see more and more of this take hold.

I bring this up because we're starting to see a whole economy be born around managing and sharing back all of that very sensitive data but, hopefully, for better and for good.

I'll give you one example of a project that we're hoping to see deliver some results on the education and skills front.

So you can imagine there are a lot of folks who are struggling to find employment and there are others who have reentered the skills market to get more training to land jobs. Well, who's analyzing, and what are we doing to understand what was the path chosen by Person A, who maybe looks just like me but went and got a degree from this program and has a job, versus Person B, who went and chose a different path and hasn't been as successful?

Well, we're not analyzing that data today, but we're starting to see a culture emerge to do such a thing.

But I want to emphasize the final point, which is where Cam started the

conversation -- how can this be used to solve a lot for policymaking and problem-solving?

If you read the book -- and I hope all of you will -- the professor highlights the story of a start-up called Ginger.io, a start-up that I had a chance to get to know because I was a judge on behalf of a company called Sanofi who had this competition called the Data Design Diabetes Challenge.

And the idea was, how can people with diabetes live better lives, and how can the care delivery system better support those who suffer from Type 2 diabetes?

And this award-winning team, which came out of his lab and now actually has been venture-backed and is making its way in the commercial market -- it has this potential, in theory, to predict when someone is likely to be going south, if you will, on their health condition. Before you pick up the phone to call the doctor -- to say, I'm feeling ill; can I come in for a checkup? -- the system, conceivably, could alert the physician and say, hey, you might want to reach out to Susie; she might need some health care.

Now that's all technically interesting. But, how does it close the loop?

In this country's debate about solving the health care cost crisis, we have now created under the Affordable Care Act the legal model to allow us to change the way the health care delivery system gets paid.

So, as we speak, if I built this cool iPhone data analyzing app and I've done exactly what I just outlined -- I can predict that you're going to get sick before you even know you are -- I don't get paid in today's -- there's no insurance billing code for predicting me falling ill before I can.

But, under the Affordable Care Act, there's now an Innovation Center at Medicare. In fact, Brookings is the home of a lot of the research on what we should do to think about payment in a different way.

And, under the Affordable Care Act's Innovation Center, there's now the legal mechanism to test, to validate and, eventually, to scale new ideas that could be reimbursed if the net effect of it is that it lowers costs and improves quality or outcomes.

So where we're seeing the marriage, Cam, and to conclude my reaction to what we've just heard, as entrepreneurs and innovators start coming up with these tools that could help solve some of these societal issues, it is imperative that we look to change the business model, or the policy, around what we do to solve that particular problem, to create an on-ramp for these ideas to be tested.

Does such a thing -- is it worth $50 bucks a pop to have someone get a prediction algorithm going? And, if so, how do we measure that, and how do we ensure there's a feedback loop to see it succeed?

I think this is going to happen in health care. It's going to happen in education and skills. And it will happen in areas like energy efficiency -- just to get started.

And that's my hope and my sense for where this is going.

MR. KERRY: Good. So, Aneesh, this is a chance to plug your book.

MR. CHOPRA: Yes.

MR. KERRY: Are there points of overlap between Sandy's book and what you're writing?

MR. CHOPRA: Well, yes. So I have a book coming out in May called "Innovative State," and I basically presume that for policymakers there are now four new tools in the tool kits you can deploy if you want to fix something.

So, in the old days, it was increase the funding for something or to regulate more in an area, and that those tools now can be complemented by the following:

One, you can encourage more openness of data, which I think is the direct overlap.

A related overlap is encourage the industry to adopt standards so that more information can be liquid across private organizations in regulated sectors -- in this case, energy data. So, if I have the right to access my real-time energy data from the power company, but I want to give it to the cool nest thing that I bought on my wall to change the thermostat, there should be a standard by which that information can flow, so the government can play a role in standards.

Issuing challenges and prizes to find folks like the Ginger.ios of the world, to see if their idea could make a difference in health care.

And then, finally, inevitably, there will be areas where the policymakers or existing agencies won't know what to do with this stuff. So the ability of agencies to adopt lean government start-ups who can go after these tightly defined areas where we can make use of these tools in ways that we hadn't thought of, like the FDA Innovation Pathway, and we can go on and on.

So the point is, yes, there's a huge amount that policymakers can be doing. It is a direct linkage, I believe, to the facts that are on the ground about the value of this stuff.

MR. KERRY: So, Darrell, one of the focuses of Governance Studies is on how to improve government, and it strikes me that social physics is something that's ripe for study in government.

I mean, I've looked at Sandy's diagram of that German organization with the department that nobody talked to and was sort of thinking back to the Department of Commerce. Okay, which was our customer service office that needed to get relocated?

How does this bear on governance?

MR. WEST: Well, there clearly are lots of opportunities for improvement in government. I really like his book. We're selling copies of it in the back.

So, if you haven't already gotten a copy, please stop by and get it. It's a book that's really filled with tremendous insights, and I think you'll really enjoy it.

When you look at the book, a lot of the examples that he cites, and that he cited in his presentation to open up, are in communications networks in terms of how you can study the networks, see which departments of organizations are communicating, which ones are not, and then how that affects the ultimate performance.

I think in those types of sectors we've seen tremendous use of data analytics. It's been very successful. Companies have demonstrated the ability to do this in finance and in other areas.

I think government represents a much bigger challenge because one of the things Sandy talks about in his book is this idea that he calls the living labs, where: You basically go out. You try innovative things. You're basically experimenting. You're monitoring the results. There's a quick feedback loop, and there's rapid learning involved that allows you to make adjustments as you're going along.

That's the goal. That's what we want to get to, but the problem in many government agencies is we're not there. Almost none of those processes are there. We don't really treat government as a learning lab.

I think Aneesh, when he was in government, started to move the government more in this direction, and that was very noble. And there were certainly some successes that came out of that, but there is still a lot more work to do.

We don't really experiment. We haven't set up the rapid learning networks.

You know, there is risk of failure in government. I mean, government

bureaucrats know that a failure gets you on the front page of the newspaper; ten successes probably don't. And so there are kind of skewed incentives against risk-taking in the public sector.

So we've done quite a bit of work at Brookings in our Center for Technology Innovation on data analytics applied to education and health care. These are the two last sectors of our economy to undergo this technology revolution. They're both very labor-intensive. People are concerned about cost. They tend to be more regulated than a lot of other areas. So it's no secret that change has been much more difficult in these areas.

When you start thinking about the opportunities for data analytics in health care, there are tremendous opportunities. And we have certainly seen at the local level examples of cities where data-sharing networks on health care have formed between physicians and hospitals, and sometimes insurers have joined in. So we're starting to create the data-sharing networks that enable data analytics, that will give us a better handle on service delivery, on performance metrics and on cost savings and what we can do to do a better job in that area.

But health care is a very challenging example in terms of data analytics because so much of the information is still considered proprietary. Companies understand that information has value, and so it's harder in those types of sectors to create those data-sharing networks and then do the data analytics that will enable the learning that he talks about.

So I think the model that he has outlined is very promising. I think education and health care are big challenges in terms of implementing those ideas.

MR. KERRY: So, for me, one of the things that this underscores is really how central to policy in this area is the privacy debate and getting privacy right.

Sandy talked about using personal data stores, part of the discussion he has in the book of what he calls a New Deal on data -- empowering individuals. That is a critical part of unleashing the power of data.

In this world, Google famously said that its mission is to organize the world's information and to make it available. In a real sense, every organization, and potentially every individual, has that mission today. That can be enormously informative, as my co-panelists have described. It can be enormously empowering to individuals.

I mean, think if you could capture for yourself all of the information that is flowing from this device. We are now in this quantified-self world where people are using Nike Fuel Bands and devices like that to try to capture information about themselves. This is a powerful tool that does most of the same things.

We are all generating enormous amounts of data. That can be a tremendous public good, as Sandy has discussed. But we can't do that without dealing the protection of individuals, protecting individual choice, individual freedom of action.

If we get that right, if we can move to a more dynamic framework, away from a stale debate that focuses on collecting or not collecting, on the pure framework of consent and control, and focus on technologies that will protect the uses of data, that will protect the -- you know, that will apply, as Sandy's team has done with the research in Trento, powerful technological tools to protect the privacy of that information. We have the potential to unleash this enormous engine to really harness the power of knowledge in so many different ways and, at the same time, to protect individuals.

Any other comments from the panel before we go out to the audience?

MR. PENTLAND: I can see hands up already.

MR. KERRY: Good. We welcome your questions. Please wait for a microphone to reach you.

So we'll take the gentleman at the back of the room on my right, over there, if you could stay and just say your name.

SPEAKER: Sure. Scott Talan, American University.

I'm a big fan of what you're talking about and have been using social media as a graded assignment since 2005-06. So I'm not a skeptic, but I hear a lot cheerleading.

So I'm just wondering; when is all this going to get there, and why isn't it already there?

MR. KERRY: Aneesh, I'll start with you.

MR. CHOPRA: Yeah. Well, I'll -- let me begin by just talking about scale and impact in the public sector.

What we've done in the first four years of the Obama administration, and it has continued, is a lot of experimentation and permission to try. So, when you dig, I could fill the lecture hall of case studies of one-offs that you might not have ever heard of because it's just not had the kind of national scale because these ideas haven't bubbled up.

I think what's most -- to connect all these dots, what's going to happen is the following:

There are now pathways for new ideas to be tested, for them to be validated and then for them to scale, only in health and education and maybe a few other areas like in social development.

So, in education, the i3 grant program which was created in the Recovery Act -- it's now the fifth anniversary of the Recovery Act. So we've got five years.

Any idea in education you can pose to get an initial test case, and if you

propose one that happens to harvest some of the data in a manner that Sandy outlines, you're welcome to. It doesn't require it. It's not limited to just data-driven ideas. But at least you can go forward. The school district could get some grant money to get some early evidence.

If you're looking to say, well, did it just happen in a one-off? Can I get some data to prove it can be replicated in half a dozen other communities? -- there's money for that.

And then, if I've gotten it done in a half a dozen communities, but I want to scale it up across the country, there are resources for that.

So the i3 program is where I would look to see whether these kinds of ideas in education will scale.

Similarly, the Innovation Center in Medicare and Medicaid -- CMMI is what it's referred to -- though not as likely to find sort of social media or data exhaust-fueled innovations yet, the law allows it.

And if we shine a light on it, and professors and entrepreneurs start to participate -- I remember my most frustrating day was like in November of 2011, and we put up this blog post: Billion-dollar health care innovators challenge. Anyone with a crazy idea on how we should reform payment in health care to reward better things and improve quality and lower cost -- and you could hear the crickets chirp.

All we got mostly -- and I don't mean to those who applied. We basically got NIH holdover losers, in a sense, meaning people that applied to the NIH for grant funding, didn't get it and sort of submitted their next thing.

You didn't see a lot of this inspired stuff. Maybe those folks didn't see the blog post, for example, which is very likely.

So that's -- you know, I don't mean to say it's all coming and it's

awesome, but there are channels you could look at to see this stuff bubble up.

MR. WEST: And, Scott, let me just add one quick footnote to what Aneesh said.

The problem that I see is that it's not a technology problem. It's an organization problem, that in a lot of these areas -- and I think education and health care are good illustrations of this -- we're not really set up organizationally to facilitate the data-sharing and the data analytics.

It's not the problem in terms of social media or how we're configuring the technology.

It's been very difficult to get the organizations to basically volunteer their information because here we have a very decentralized and fragmented system. You know, each particular niche has its own information. They control that. They can use it for their business model.

It's been very hard to break down that fragmentation because to create those data-sharing networks that will enable exactly the type of thing that Sandy is talking about you have to get those organizations to be able to cooperate. That has been one of the big barriers that I see.

MR. KERRY: And, if I could add, I think part of the organizational challenge is that all of this is very new. By some measures, 90 percent of the world's information has been created in the last 3 years. So we are really just at the beginning of this next wave of an information revolution.

And that is a challenge to institutional leaders, to CIOs, to understand how to capture that information, what to do with it all.

And it's a challenge to policymakers. I met with a member of Congress not much over a year ago, who was in charge of some of these technology issues, and I

said something about the internet of things, and I got a blank look. What is -- what do you mean, the internet of things?

MR. PENTLAND: But this wasn't your brother.

MR. KERRY: It was not, no. It was not.

(Laughter.)

MR. KERRY: But that is clearly changing.

I think one of the byproducts of the Snowden revelations is certain to shine a light on the scope and the scale of data collection that is possible and what it's possible to learn. People understand a lot more metadata today than they did nine months ago.

MR. PENTLAND: Can I add one thing?

Which is, the two organizations which I find really interesting are the G8, which is not the one you think of -- it's the digital innovation centers of some of our largest cities. There are eight cities that have agreed to work together to try out different innovations and just share with each other. It's an informal network.

And the other is there is an amazing state government network where things that are invented in Kansas get taken up by Indiana. It's quite amazing actually. All these guys know each other and are busy trying to learn from each other.

I find that sort of peer-to-peer experimentation to be really interesting because it lets you try a lot more things in a shorter period of time.

MR. KERRY: Okay, we've got a question up here in the front.

Sir, hang on for the microphone.

And may I remind people? You can Tweet at #TechCTI.

SPEAKER: My name is Eskil Ullberg from George Mason University.

The question for the author of the book today here -- it was interesting to

hear about the Trento experiment, I think you called it, in Italy. What was the response? If you could elaborate, what was the response to these sorts of privacy concerns that you were trying to introduce, that the information is more private?

That's question one.

And, two is do you think it's possible at all to keep data and the users of data private, which seems to be a premise in the discussion?

MR. PENTLAND: So the experiment in Trento has been met with a very positive response by the participants. They have indeed felt safer and shared more.

The software and the commercial structure, because it's a contract law structure, is patterned after the SWIFT Network. So the SWIFT network is the interbank transfer network which handles some $3 trillion a day and, as far as we know, has never been broken, which is an interesting record when you compare it to all other networks.

And what it does is it allows for sharing of data under a peer-to-peer contract and for auditing of that. And so it actually has been very successful at operating in all sorts of questionable jurisdictions and with questionable actors for many, many years.

We believe that scaling that to individuals, to have that same level of security, will change the risk-reward for sharing information. You can never make it completely safe, but you can make it much more safe than it is today, which should promote a little more sharing.

You also have to be able to give people the ability to audit it -- to say, did it really do what I said it was going to do -- and the ability to retract. And those are the sorts of things that technically and legally, in terms of contract law, not new legislation, are very much within scope.

In fact, we have an NGO which has built the software and built the legal

contracts. They're open-source. They're available for everybody. And people are beginning to experiment with them.

But I'll emphasize that it's experimentation that is the proof, not that I say it's good.

SPEAKER: Thank you.

MR. KERRY: Let's take a question back there on the left.

SPEAKER: Thank you very much. Paul Selker, Spark Street Digital.

My question is about causation, primarily for Professor Pentland.

I was encouraged to hear your talk about living labs because it seems like it approximates, or moves towards, a randomized clinical trial almost, which is the gold standard of evidence and the best way to establish causation.

And I guess I would wonder, beyond sort of people's tolerance for sharing information, what you've been able to find, especially when you start talking about the relationship of socioeconomic status to information-sharing or innovation or innovative capacity, whether you've been able to get a little more information on which way causation really runs or if there are other examples of trials or experiments that tackle that question, which seems to me to be very thorny.

MR. PENTLAND: Right. So we've been able to look at these questions in small groups and in small organizations, where we actually do interventions and predict outcome, and we can quantitatively predict outcomes.

So we can do things like randomized trials.

We can actually do quite a bit better. Randomized trials are not the best thing to do. They're a good thing, but if you can actually do many independent, quantitative predictions, that's actually rather better.

What we haven't got is that at the scale of an entire city or an entire

society. That's the hope for things like living labs.

What we do have is we have things at the scale of a community, where we can look at the causation of these things as natural experiments, where we segment things into different populations and we ask -- if they become more poor, does the behavioral signal change? If they become rich, does the signal change?

And the answer is there seems to be a basic motivation in people for exploration, and that's expressed more when they feel more comfortable. So, when they feel more wealthy, when they feel more safe, their behavior changes to be more exploratory. And that's the signal that we're looking at.

Does that actually turn into a more wealthy society? We have data over several years which argue that it does at least somewhat.

I mean, our squares are like 0.3. They're not huge. But there does seem to be some causation there, and I can talk more about it offline.

MR. KERRY: We have a question up here in front.

SPEAKER: Hi. I'm Jessica Floam with the Medill News Service.

Aneesh, you mentioned government standardization and the role that it can play in data-sharing. I was wondering if you could go into the specifics of that.

And then, Mr. Pentland, if you could maybe talk about specifics again of what data you hope can be shared -- I mean you mentioned Nike wrist bands and cell phones. In an age where two weeks ago, we had hearings of Target and Neiman Marcus and these places where there are pretty serious data breaches happening. How can people feel safe and know that their data is not going to be misused?

MR. CHOPRA: I'll start very briefly.

So, if you recall, I think it was 2000 and -- by the way, the timeliness of your question; the White House just issued a request for comment on updates to the

policy framework which is called the Circular A-119, which governs the role of government and standards, so just a quick primer on that for why and how this has had impact.

Prior to the 90s, the concern in government was that, frankly, the DoD would set their own specifications that were completely different from what the commercially, widely available stuff had been available. And so it increases cost, creates this complexity, and in the case of trade, many foreign countries might pick standards that are different than what are normally ascribed to how we do WI-FI or whatever the case may be.

So you have this sort of bizarre circumstance where you kind of want to have government staying away from standards activities. You'd rather let these voluntary industry consensus bodies kind of work out the Sony-Betamax-VHS kind of debates and not get into this as a top-down governmental thing, which is great. That's our philosophy. That's the American way of life. That's our strategy.

But it became an issue as these digital assets started coming to market, where there's a societal, or policy, goal that's at stake in whether these things all interconnect.

So just think for a moment in the energy sector. If every electric car company built their own plug and socket design that had to interconnect with every power company in the United States, just imagine the kind of inefficiency that would be wrought.

So there the question was the government didn't want to create the standard, like "Tablets from on high," but could play the role of convener, and maybe an impatient one, to bring folks together to say, hey, maybe we should get these disparate assets in the energy sector to talk to each other.

So, off has been built the Smart Grid Interoperability Panel, designed

and built by the private sector with some nudging in government.

How this comes into play with the personal data assets is that very early in the first term of the Obama Administration a gathering of folks in health care said, look, we should start sharing data, and one of the biggest sources of data is the VA.

So the VA said, look, we're willing to participate as one of you. We're all data suppliers. We're just going to make the commitment that if a veteran wants to download safely and securely a copy of their own health data, well, we'll just make it available to them and they can push this button.

But, by the way, we would love the community to help us define what should that data structure look like. And so, with your input, let's all do this together.

And so without a law mandating that everybody must do this, we've now got over 150 million whose health care data is held by a private entity that has pledged voluntarily to give you back your data. That's the blue button pledge.

So, just 10 days ago, all of the pharmacies -- CVS, Rite Aid, blah, blah, blah, Safeway. They've all said they're going to join the movement precisely so that people can -- and now folks that want to live in living labs or whatever, they now have a standardized way where they can access their health records and then do with it whatever they choose -- store it somewhere, share it with others.

And that's an example of the role of government in that standard activity, which was to be sort of an early adopter but also a convener and nudger to get these things on to the finish line.

MR. PENTLAND: So, to answer the other part of your question about security and trust, there are these networks that have been in operation for a long time that are very secure and are trustworthy, like the SWIFT Network or the Visa interbanking network, but typically, they've been very expensive to run, both in terms of labor -- you

need a lot of lawyers -- and the computation research.

But that's not true anymore. You can now build them really as inexpensive consumer things, and I think that's a very hopeful sort of a sign.

A lot of the things like the Target and so forth are people trying to save pennies. I know that adds up to a lot of pennies if you get all those terminals together, but we can do quite a bit better.

The other thing is that you should not have the attitude that this is an absolute thing. We're in sort of a war, and we're going to lose some of the battles. What we want to do is win by far the majority of them so that there will always be some risk in sharing. But we want to minimize that in relationship to the reward for sharing, and we want to make that balance between risk and reward something that's visible to you.

So, like people who went to Target had no idea that there was any risk that they would get hacked or maybe a tiny, tiny thing, but in fact it was a substantial risk.

So we have to let people know what the risk is; we have to minimize it; we have to make visible what the reward is, so people can have an informed decision about whether to share or not.

MR. KERRY: Ma'am, why don't you go ahead?

SPEAKER: Thanks.

Is this on?

MR. KERRY: Yes.

SPEAKER: Oh, thank you. Thank you, gentlemen.

My name is Lorelei Kelly. I'm at the Open Technology Institute.

I've been researching the impact of technology and transparency in Congress and how we could decentralize functions out into the states for the last couple of years, and what I've found is that the sort of expectations for direct democracy are

running up against the needs of the institution for authentic expertise. Really, not everybody can participate and create the best outcome for legislation. It's people with real substantial knowledge.

And what I've found is Congress itself isn't as sort of sordid and corrupt as it looks on House of Cards as it is obsolete and incapacitated. I mean, it can't handle the information demands of any modern democracy. So, right now, it needs the crowd less than it needs the curators, you could say, in technology speak.

And, as you know -- especially, you, Aneesh, who work in state government now -- a lot of reputation value is based on real relationships, and that means you have to show up in person and do the work. And I'm not seeing this.

To me, I see it as the difference between sort of transparency and accountability. And we haven't been very clear about transparency and what that means. And then accountability is the governance piece of it, which requires a new social contract almost.

Could you comment on that? How are we going to build a compromise?

MR. CHOPRA: This is a tough one. But let me go the regulatory state versus Congress to give you a flavor for maybe I might have a slightly different view about the value of the crowd versus the curator, and I'll just give you an example of the Dodd-Frank bill. You know, Wall Street reform.

So, if I were to ask you what was like the big enchilada in Wall Street reform, the one that got a lot of attention certainly was the Volcker Rule. This was the idea that banks should essentially separate functions and was very politically charged in terms of what is the line that should be drawn in terms of what is considered proprietary trading and what is considered the functions of the bank.

And, despite all of the attention, if you looked at the number of people

who issued comments during the debate over the Volcker in the executive branch -- that is, the law was already passed, but the Treasury put forward rules for public comment -- I don't have the final tally, but I believe it was something like 20 people commented on the Volcker Rule, I believe. And all probably were these curators who kind of amalgamated opinions about what should or shouldn't be done, and you could see people's perceptions about that.

Also, in the Dodd-Frank law, we birthed a brand new agency, the Consumer Financial Protection Bureau.

And a week or two after Elizabeth Warren was asked by the President to sort of set this thing up, we met, and I asked her this very basic question, would you like to sort of build this as a 21st century agency or sort of be the last vestige of the old guard?

And, of course, that was a framed question. She's awesome. So she's like, I want to embrace the future.

So, in the contrast with the curator-Dodd-Frank-y thing, or the Volcker Rule, she had her own rule. She had to create a simplified mortgage disclosure form within two years of the law.

Now she could have waited two years to then like issue the proposed rule and have all the lobbyists comment on the rule. Instead, two years before they even had the authority to issue the rule, she crowd-sourced the mortgage disclosure form. You can go to the web site, Know Before You Owe, and you could log in, and you could vote.

Seventeen thousand Americans -- talk about big data. It tracked where they looked on the form, where they didn't look, where they had questions, what didn't they.

And all of that data -- they had seven rounds of public comment. That allowed President Obama to hold up a copy of a simplified disclosure form that the American people essentially wrote, and the rule hadn't even been proposed in the Rulemaking.gov web site because it hadn't been official.

So I think these participatory channels are still hugely important. And not every one of the 17,000 people was a "curator" or expert, but that product, I think, was an extremely high quality example of what's possible.

So just drawing the contrast, I think the opportunity to still tap into the expertise of the American people is early days but shows great promise. I'm not going to give up on the crowd versus the curator just yet.

MR. PENTLAND: Let me add one thing to that, which is that in the things that I showed, but the things in the book also, we often get millions of people trying to make decisions. And what we find is the primary thing that determines the quality of the decision has to do with the diversity of information sources.

In things like you were just talking about, you can actually measure the diversity because you can see who looked at what without looking at people who are all looking at each other, so it's an echo chamber, or where they're actually looking at a very diverse set of opinions. And what you can now do is you can begin to make measures about who has actually had the most diverse set of opinions, who is likely to make the best decisions.

And there are actually some math and statistics behind it that say, these are likely to be the ones that are the best, and these ones don't have either enough input or they're living in an echo chamber.

What you do with that -- there's a lot of things you can do, but it's interesting that you can begin doing these sort of new algorithms to determine when you

have enough input and it's diverse enough to be able to have a good decision.

MR. CHOPRA: That's great. That's a great idea.

MR. KERRY: And Aneesh talked about the role of government in standard-setting as a convener, as a mediator among stakeholders. It's something that's been critical in the standards development role. It's something that the National Telecommunications and Information Administration applied in developing mobile transparency rules.

I think that is something that in this multilayered, complex world, where you're facing a lot of technology issues, is something that can be a model for decision-making in a place where, as Aneesh illustrated, you can bring to bear technology to help expand the range of opinion and the stakeholders who can be part of the discussion.

Let's take a question back here on the left.

SPEAKER: James Sang.

On the topic of echo chambers --

MR. KERRY: Back here on the left, over here.

SPEAKER: Oh, sorry.

MR. KERRY: I'm sorry, my left.

SPEAKER: Kristine Quinio from the National Academy of Social Insurance.

How would you describe the relationship between the predictive power of social physics and the practical application of behavioral economics with, say, improving retirement security?

MR. PENTLAND: The way I look at it is the following: Danny Conowin had this insight that people have two ways of thinking. There's sort of the habitual way which hauls fast, and then there's the sort of reasoning it through, and the interaction of

those two gives people biases.

And what you can do is you can set up situations that nudge people one way or another, and that's sort of the heart of behavioral economics -- that there are these biases and you can do things.

But, if you wanted to go a little bit further and you had a lot of big data, you could actually write down the equations.

So behavioral economics talks about, oh, there's this bias and this bias and this bias. You can weave those together into a holistic description of how people make decisions.

So, for instance, in these financial things, you can see the behavioral economic biases, but you can also write down equations that predict what people will actually do. And, when you can do that, then you can begin to offer incentives, craft incentives, which will make everybody behave much more to their interest.

In the sorts of things we've done, we see rather huge improvements in efficiency. When you understand the mental processes of people much better, you can set up a situation which causes them to make much better choices, better than just looking at the biases.

I don't know if that got the question.

MR. KERRY: Yes. I mean, the research that really caught my eye in that regard is Sandy showed the chart with the trading statistics combined with social network, which showed the echo chamber effect up in the upper right-hand quadrant, which indicated that there was a bubble effect taking place.

Well, imagine if the Fed -- imagine if the SEC had data like that in 2006 and 2007 to inform fiscal policy or other market policies.

So, the gentleman back there on my right, whom I cut off earlier.

SPEAKER: A question for Professor Pentland. Following up on echo effects, echo chambers, you pointed out that figures as a sweet spot.

You pointed out that most of your data show that innovation and productivity increases with increasing communications, but each of those points for your communities or something represents a distribution of outcomes.

As you move up the curve with increasing communications, does the distribution on each of those points change very much?

This is obviously getting to a winner-take-all type of question in the wonderfully connected future.

MR. PENTLAND: Well, there are two things.

One is that as you get greater -- the question is about flow of ideas and how many independent, different strategies, ideas and techniques are you exposed to, to choose from.

On the curve with the sweet spot, what happened is people got to the point where the flow of ideas was such that it crowded out new ideas, and that's the makings of a bubble.

So I think that's -- in organizations, it's possible also to get too much communication, where people are just hearing the same things again and again.

So there needs to be a bias, a mediation between those two things. And it's different in different situations, exactly where the sweet spot is.

The distribution is actually a lot tighter than it looks like from that graph, but I think that's a much more detailed thing.

SPEAKER: Do those points change very much as the (inaudible) from left to right?

MR. PENTLAND: Well, they go like that. I think there's a little more

detail than people want to listen to.

In that graph, each point is the result of 1.6 million people trading for 1 day.

If I did it just for one asset class, like dollar-euro, it would be a thin, skinny line like this.

This would be echo chamber.

This would be people working alone.

And, in the middle, you would have the sweet spot.

You can actually pick out the people who will make the best decision based on their communication pattern, and you can do that several times a minute. And, if you do that, it's actually a good trading strategy.

(Laughter.)

MR. KERRY: So we had somebody in the middle here on my left.

(Pause.)

MR. KERRY: I guess, why don't we move to the front of the room here?

SPEAKER: Hi. I'm Fred Altman.

From the conversation today -- with the NSA collection of all this data, there are lots and lots of problems with it. From the sounds of the discussion today, the thing that the NSA should do is release it and have everybody have access to it. Is that a good idea?

(Laughter.)

MR. KERRY: No. I thought Edward Snowden already did that.

MR. PENTLAND: So one of the things that you can do, though, as we're talking about these personal data stores, is giving individuals much more control over their data. If you actually did that, it would give the NSA some headaches. They would

still be able to break into things and get the data they need, but they would raise a lot more oversight in doing that, which might actually be a good thing.

I'm in conversations with some of the people who are thinking about the policy for homeland security, which is headed in that direction because the NSA isn't really the one you have to worry about; it's the people in Romania who want to take down the power grid and do other things like that, using the very same data.

So, if you'd like to have a resilient society, you need to imagine that you're constantly under attack by lots of people and you need to make sure that there's no one place where all the data are put and you also need to make sure that no THADD or dictator or authoritarian element takes over all that data.

Probably the best sort of thing you can do is to make it very distributed and give people a lot more control over their data so that it becomes harder to use that as a weapon in various ways.

MR. KERRY: Okay, I see Professor Pentland starting to scratch himself over here. So let me put one last broad question.

Our republic is founded on these very orderly principles of Newtonian physics. What you've outlined is something far more complex.

Does social physics sort of give a pathway to sort out some of the complexity, the dysfunction, of our current system of governance, or is it time to go from alchemy to chemistry or some new model?

MR. PENTLAND: Well, let me put my foot in it first.

MR. KERRY: Okay.

MR. PENTLAND: So one of the things that has continually sort of irritated me is when I go to a place like Davos all of the reasoning is based on 17th Century science.

So the sorts of things you see -- economic reasoning, Locke, things like that -- are things that treat people as individuals.

And, of course, we are individuals, but we're also part of a social fabric. And our wants, our needs, our reactions are as much a function of the people around us and the people we interact with as it is the thing between our ears.

That is something that our current ways of thinking about society and governance don't take into account, which is, of course, the same thing that causes bubbles, fads and polarization.

So I think that there are ways of governing ourselves that avoid a lot of these. We were just talking about, how do you make decisions and how do you crowd-source ideas, and how do you integrate that input from everybody into a real wisdom?

And we're getting to the point where we can begin doing this for real. We need to try it out. We need to be experimental. But there is a possibility of having government and organizations make far better decisions than they do today by thinking about the social fabric and not as a collection of isolated individuals.

MR. KERRY: Well, with great respect for our terrific panelists and the audience, I'm going to let Sandy Pentland have the last word and send him off to Chicago.

Thank you very much, Sandy.

Thank you, Aneesh.

Thank you, Darrell.

(Applause.)

* * * * *

CERTIFICATE OF NOTARY PUBLIC

I, Carleton J. Anderson, III do hereby certify that the forgoing electronic file when originally transmitted was reduced to text at my direction; that said transcript is a true record of the proceedings therein referenced; that I am neither counsel for, related to, nor employed by any of the parties to the action in which these proceedings were taken; and, furthermore, that I am neither a relative or employee of any attorney or counsel employed by the parties hereto, nor financially or otherwise interested in the outcome of this action.

Carleton J. Anderson, III

(Signature and Seal on File)

Notary Public in and for the Commonwealth of Virginia

Commission No. 351998

Expires: November 30, 2016