OECD-Canada Technology Foresight Forum
Session 2b
Research 2.0: e-Science and new ways of interaction in the science community

Ottawa, ON

--- Upon commencing on Wednesday, October 3, 2007 at 11:25 a.m.

MR. VICKERY: Welcome everybody. I assume everybody has come back from coffee.

My name is Graham Vickery. I work for the OECD Secretariat.

If you have got any questions, any problems, any issues to raise about understanding the emcee who is in the other room, please contact us.

I would like now to hand over this session to Walter Stewart, who is going to be our very able Chair, and I will give some small amount of directions about a meeting we are going to be having over lunchtime at the end of this session.

So Walter, please.

MR. STEWART: Good morning, ladies and gentlemen. Bonjour, mesdames et messieurs.

C'est avec plaisir que je vous accueille à ce volet Recherche 2.0 : La cyberscience et les nouveaux modes d'interaction dans la collectivité scientifique.

Nous avons quatre conférenciers/conférencières. Je vais vous les introduire par nom et titre seulement.

Vous pouvez lire toutes les biographies sur le site Web de ce Congrès ou dans ce petit livre qui est disponible sur la table dans la salle principale.

Je vous demande de noter la description du volet sur l'aperçu que vous avez reçu quand vous avez enregistré ce matin.

Il y a trois questions que nous avons demandées à notre intervenant de s'adresser.

Après les présentations, il y aura une occasion pour vous, vous qui sont dans la salle et aussi vous qui participez sur le Web, de demander vos questions.

I would also draw your attention to the translation devices -- perhaps I should have done that first -- draw your attention to the translation devices.

We may well have -- the presentations will be in English. We may well have questions though in French and so please feel

free to ask your questions in whatever language suits you.

With that, I am going to introduce our first panellist.

Our first panellist is Andrew Herbert from Microsoft Research in Cambridge, U.K.

MR. HERBERT: Thank you.

Hopefully, my slides will appear momentarily in front of me.

So what I want to take as my theme is perhaps a little more broadly than just the Web itself but actually it is a look at the impact that computing and computer science as a whole is having on the other sciences and some of the consequences of that.

So the key, I think, is that the sciences -- and I use that quite broadly, physical sciences, life sciences, engineering -- are all increasingly relying on advanced ideas from computer science essentially to reduce the time to scientific insight.

In the past, perhaps we thought of science as being divided into theoretical science,

primarily the domain of mathematics, and experimental science, the world of the test tube and the accelerator.

In between those two now sits a third strand of science, which is computational science, that is, the world of simulation, data mining, visualization, pattern recognition, machine learning and many other techniques, and the advances in those techniques are primarily coming from the computer science community.

And indeed, I think one of the questions that needs to be addressed is how are we going to produce people in the scientific community who have the right balance of skills between the core scientific disciplines -- the biologists, the chemists -- and who are able to work with the most recent computer science techniques and indeed contribute to and advance those?

This is a perennial problem. Many scientists learn their computing in the first or the second year of their bachelor's programs. They lock into the operating system and the

programming language of that time and know a little bit about bashing data in files. Using advanced database techniques, using computational grids and so forth are all very new and exciting things for them.

So I think the key points are that the computers are enabling scientists to share massive amounts of data. For some of the sciences it is the massive amount of data.

If you think of physics, when the Large Hadron Collider at CERN comes on stream, that is going to be generating petabytes of data in which the physicists are searching for very rare events.

In other disciplines like the life sciences, actually it is lots of very small databases that have to be that have to be connected together as we are trying to join up different parts of biological knowledge: two different problems but both equally complex and both dependent on networking large amounts of resource.

I think you will hear more from

the other speakers about the way in which Web 2.0 technologies are being used by scientists to create virtual organizations, linking scientists in different laboratories together, combining their resources, whether they are competition resources, data resources, access to facilities, in many new ways.

As a consequence of that, it's revolutionizing the way we think about scientific publication. If scientific work is ongoing and being conducted through blogs and online experiments, online meetings, why do we need conferences, why do we need printed journals? There are interesting questions in that online world about the provenance of data, the tracking of data, the archiving of it, ownership and very deep issues.

So for me I think the computing ingredients that come into the picture through technologies like sensor networks, we can bring real world data into the theoretical models. We can link those to our computer models, simulations and so forth. We can store huge amounts of

persistent distributed data and so we can bring the experiment, the models and the data together simultaneously by using computer techniques to alternate scientific workflows. Using the technologies of data mining, which have come out of the world of the enterprise and business, finance and so forth, we are able to perhaps even think about automating some of the aspects of generating scientific insights.

I don't think scientists will go away. Computers never succeeded in making people go away. But what computers have done is let people focus on their core skills and competencies and the computers have done the drudgery behind the scenes for us. They have made us more productive and I think the same is happening in science.

What people are good at is interpretation and insight.

I want to give you one example of an area of research that colleagues in my laboratory in Cambridge are actively engaged in, which is a fascinating crossover between computer

science and biology.

We have been working in an area called computational systems biology, and this essentially is looking at the biology of cells and how they interact in organisms. The approach has been to treat cells as if they were abstract computers, which is where the computer scientist gets interested. And as a computer scientist, we have developed many tools to help us model computers, model complex software systems to do things like prove software is correct, to understand how one computer relates to another, to decide if particular models of computation are equivalent.

And we are now starting to transfer those ideas to help the biologists who have many of the same problems. And indeed to a computer person -- and my background is in hardware and operating systems -- when a biologist explains how a cell works, it sounds as though it's three little abstract machines connected together.

Cells are all about membranes

dividing the various parts of the cells, and cells themselves. The membranes are about confinement, storing things, and indeed the bulk transport of things around the organism. Those are computing words.

There is the protein machine driven by the amino acids, which is where metabolism takes place. Food is consumed and turned into energy. Things are propelled around the system. Signals are processed. That very much feels like a processing element, if you like.

And then there are the genes, whose role in biology is clearly the regulatory system to keep all the pieces working together, and those things signal to each other. The genes are perhaps like programs.

So as a computer scientist, there are many of our words that we can bring to describe the biological system of the cell and many of our techniques and modelling complex systems that we can perhaps offer to the biologists to help them develop fuller models of what is going on in their field.

The biologists have a challenge. Physics and chemistry advance because of mathematics. Once you could model through applied mathematics, the world of physics, we can make predictions in the mathematical models and then go and verify the experiments.

Chemistry made huge strides when models of atomic structures, of molecular structures, could be represented mathematically once we had the equation and other mathematical tools.

Biologists don't have that mathematical framework. They are still fundamentally doing zoology and botany; collecting things, squashing them, sticking them in albums, trying to deduce conclusions by looking at what they have discovered. And they have no way of writing it down apart from little cartoons and writing statements in simple English.

So perhaps some of the formality and notations that we have developed in computer science can help them.

That's the track down which some

of my colleagues have been going. We have been taking ideas from abstract models of software systems where we have essentially mathematical notations for describing dynamic systems and how they interact.

The particular one that we use is called the pie calculus. It's the notation that people interested in theoretical computer science use to explain different programs to each other and to understand what is really going on in those languages. When the vendors are arguing about Java being better than C-Sharp or the other way around, the theoretical guys can say no, that's all just syntax. These are the fundamental ideas and the fundamental explanations.

In those mathematical models, we often start by drawing simple graphical pictures as a way of capturing knowledge. What we have been able to do is give the biologists a formal graphical notation. There are some examples on the screen. Simple biological entities, arrows representing ways in which they send outputs to each other or respond to signals. And then in the

names of those things, we can capture their formal behaviour.

Once we have a formal graphical way of describing things, which is the information capture part of the process, we can then transform that into something which looks a little bit like programming language. And once we have that programming language, if it's something which is indeed truly compositional, we can start describing circuits or organisms, in the biological word, by combining those libraries together.

And then with those programs, perhaps we can run simulations looking farther into the future. Perhaps we can turn those programs into essentially manufacturing steps to build organisms entirely as instructions and look at their behaviour.

So some particular work that we have been involved in has been looking to see if we can use those models to actually explain real pieces of important biology. One of our early results has been looking at some parts of the

human immune system.

On the right-hand side of the screen, you can see the kind of pictures you would see in a standard biology textbook describing the process by which a receptor looks for hostile cells in the system, traps them, absorbs them, breaks them up into components and then ejects them from the system.

Today biologists do that by essentially drawing cartoons. The pictures are too small to take you through all the details. But that's essentially the level of formality.

How can that picture explain to you the general concepts?

There is no information in there about how long it takes that reaction to occur. There is no information in there to let you think about how you might generalize that particular mechanism to tackle other kinds of receptors, and so forth.

What we have done, working with the biologists and listening to them and their explanations as they unpick some of the

biochemistry, is to turn those cartoon diagrams into the things in the middle, which are the graphical representations of the various parts of that immune system, and then from those generating the programs, if you like, in our formal notation. And then with those programs we can start to run simulations.

The graphs on the left are the computer simulations showing how the concentrations of various of the biochemical parts of the system change as the reactions take place.

We have made very good progress in that. We can simulate biological systems as simulations match the behaviour of the real system. There are a number of cases where we have actually helped the biologists explain what are some of the key signals that are actually driving the process. Biological systems are immensely complex in understanding which of the key elements is a particular difficulty for them.

So we have helped them understand their science better. And indeed we have made some early steps, when others in the field are

doing similar things, in building custom biological structures at the gene and cell level that have a behaviour that we want to impose.

We have build the biological equivalent to the computer multi-vibrator, the thing that flashes off and on, and we have built a biological system that can do that.

So those are exciting early stage results.

The question is: Where is all this taking us if we look into a long way forwards?

First of all, I think there are interesting opportunities here in modelling the effect that drugs might have on the personal gene machine, and so pharmaceutical companies are very interested in this line of research.

If you combine that with some of the work that is going on in biosensing in the field, that gives us perhaps the ability to be monitoring our own personal gene machine in real time.

With some of the work going on

nano-materials, particularly nano-materials based on engineering with DNA, we have the possibility of modelling the affect of drugs on our system, modelling our own system, creating drugs that are optimized for that individual system and that leads us to a vision of personal healthcare, something a guy called Leroy Hood has talked about a great deal.

Healthcare that is predictive, and so we are responding to things before they become a problem, that is preventative, that is removing bad things from the system, that is pre-emptive, striking before it is too late and indeed which you as a person may participate. Because if all this is happening with software technologies, the opportunity for you to be involved in the negotiating of a doctor is very important. So that is one direction it might go.

Another is thinking about engineering bioenergy systems and predicting a model in those. So that is just one area where I think computing is having a huge effect on a particular science. There are several others and

that I have been closely interested in and talking to people about, understanding the human brain. The human brain is not a computer, there is no notion of software in the human brain. But certainly at the level of neurons and synapses a lot of what we understand from machine learning or patent matching seems to be what is going on and so that is helping us with some of our interpretation.

Using computers to model global epidemics. As a Brit, we are quite concerned about this. We have got two diseases rampaging our country at the moment, Bluetongue and Foot and Mouth. We kind of stopped worrying about cells and that is in someone else's backyard.

An indeed the work that is going on, and with the physicists, trying to understand the origins, workings and indeed the ultimate demise of the universe, you can't experiment with the beginning or the end of the universe, that has got to be done with computers.

So I have tried I think to open up the way in which computing is changing the way

science is done, accelerating the pace of science. If you want to follow-up in more detail on some of these things where I have wet your appetite, with colleagues we have published a report called 20/20 Science, that is trying to address many of those things. You can download it and you are very welcome to do so and I would be happy to have a further discussion about it.

Thank you.

MR. STEWART: Thank you, Andrew.

I would commend that report to you, it is truly excellent and not to be missed if you haven't read it.

I would now like to introduce Bill St. Arnaud, who is the Senior Director of Advanced Networks for CANARIE.

MR. ST. ARNAUD: Thank you, Walter.

One thing I am going to talk about in my brief a few minutes here is how the impact of these web2 technologies you will be hearing about not only will affect how scientists do science but how it will allow a greater community

to participate in those research and scientific activities. I think that is going to have a very profound affect on scientific policy and other educational policies and so forth.

So you have been hearing all the talks, the web2.20 tools mashups, blogs, wikis service from the architectures are transforming all different walks of life. You have been hearing about Business2.0 enterprise, Battlefield2.0, the U.S. Military has a major program and using these technologies in a variety of fields. Microsoft has just introduced a program called Telco2.0, all sorts of mashups for telephone companies and network services and so forth.

So these same tools which are transforming all sort of walks of life are, as you have heard from our speakers and I am sure following speakers, are going to transform science and research for scientists and researchers themselves, but also for a much larger community.

And this has been labelled a citizen's science, it will allow a faster transfer

of knowledge, you know, as opposed to waiting for the papers and journals, we are seeing now the transfer of science and knowledge so much quicker coming from academia and the research community through blogs and wikis and so forth. And now that is the major medium now for new knowledge and new information that has been past around the world.

And it is also democratizing science. Increasingly as we see scientific data being digitized, therefore it becomes immediately more accessible, assuming you solve the DRM issues. And so it is not only accessible to all the scientists, but it is also accessible to members of the public. And the public can then take the same data and run their own models and do their own analysis, and this is going to have a significant policy impact.

Let us just give you one simple example. You may have heard of a few weeks ago, a blogger had taken some of this C02 data and discovered that in fact the warmest period in earth's history in the last 100,000 years was not

the last 10 years, which is the common assumption, but actually happened in the 1930s, because he had done some analysis and comparative analysis and corrections and so forth.

Now, this has been debated but this is a good example of how one individual, one blogger can get access to this data and do a different analysis interpretation which, of course, has significant policy implications and so on and so forth.

But now there is all sorts of activities by students and members of the public involved in doing these types of things in astronomy, in high energy physics, climate science and all sorts of things. And Intel, for example, just released a product called Mashups for the Masses, which is a set of tools that really enhances capability for individuals to grab datasets from different areas, mash them together and create new results and new interpretations of the original datasets.

So you have heard of mashups mostly coming from the Google world of, you know,

taking geographical data and mashing it up with real estate data and violence and so for so you can get maps showing where the most houses are, where are the lowest crime rates and so forth. But now people are using these mashups to merge together different data sets from all sorts of different fields.

And so also in the past this computational science was largely restricted to those who had big high performance computers and the big databases and storage facilities to do this type of analysis. But with tools like EC2 and S2 from Amazon and other companies those types of resource now are available to the average user or to students at a very low cost so they can do this type of computational science themselves, take these same datasets, run large models using either peer to peer networks and so forth or the newest tools like from Amazon and other service providers.

So the key elements of course for precipitative web free science are the distributed databases, instrumentation and computational

facilities, extensive virtualization. What we are seeing is these ad hoc developments of what we call virtual organizations built around these types of structures, not only between scientists themselves but between communities of users interested in these very scientific activities and using workflows and mashups and so on and so forth.

And so what is happening is this real democratization of science being made available by these web 2.0 tools. Bioinformatics community, for example, there is a group of researchers and members of the public who are developing a whole bunch of mashup tools and service architectures using Amazon S2 and EC2 to provide non-researchers tools to do a lot of these bioinformatics analyses, genome analyses and so on and so forth. And so these are types of things that are starting to happen out there at a grass-roots sense rather than from a formal research environment.

So here is some very quick examples of these types of activities. A big one

is of course crowd-sourcing. There is a group of researchers now who are using crowd-sourcing tools to allow the large community of humans to really identify new research techniques and new scientific evidence.

So, for example, there was a gold company here in Canada put out a prize, sent out to the large internet saying our geologists think the gold vein is here, we invite the community to analyze that same geophysical data and come up with their own interpretation where the best gold veins are. And surprise, the community came up with the better answers than the professional geophysicists.

And so now -- I wanted to bring out my computer except it broke down -- there is now a research community in the United States dedicated to this, using crowd-sourcing, to use the large collective knowledge of the human population to identify these new trends and new ideas.

Another good example is this Project Neptune many of you may have heard of.

This is was a joint U.S./Canadian project. Fortunately, Canada got its funding first and we are the first to deploy. But this is a large undersea fibre network on the ocean floor off the west coast of Canada and the United States and this is now being deployed as we speak. And this is going to have all sorts of undersea instruments, cameras, robotic devices, sensors on the ocean floor to measure all sorts of geophysical and oceanographic phenomenon and so forth.

And of course, you can't send a researcher to the ocean floor. This is all going to be remotely accessible. And this data will not only be accessible to the scientists who participate in this project, but it is designed from day one that this data will also be accessible to students and to the public at large. And this talk of virtual aquariums, we already put down high definition TV cameras, you can watch these smokers, you can see the various biota that exist around that and it is available to anybody on the website right now to look at this type of

activity.

Another good example is from Canada, our forests are very important to us, as well as snow. But this is a large distributed grid being built by our government research department linking up sensors on the forest floors, databases, satellite data and data from a variety of sources to measure the health of Canada's forests. But one of its primary objectives is to measure Canada's compliance the Kyoto Agreement.

The Kyoto Agreement is dead now, of course, but Canada signed on under the assumption that our forests are big sinks for carbon dioxide. But that was an assumption, we really do not know how well our forests absorb carbon dioxide and so we hope that the data from this will allow us to justify driving our SUV's over the next ten years.

But, again, this data is all going to be made available on the public, so the public can also interpret this data. So, it's just not going to be some high priests of science who say

yes or no, but also the same information will be available to any community to re-interpret and re-examine this type of data.

Another great example is the ALTA Cosmic Ray project. This was started at the University of Alberta with the high energy physics community. It involves fifty schools now, or probably more, across North America, who are looking at very deep space high energy cosmic rays, and the students participate in this activity. The data is gathered through web services and collected at the University of Alberta and the students are involved in the analysis and interpretation and so forth, to really understand the very cosmological origins of these very deep space high energy x-rays. And it’s a great project for students to work with real science and scientists on trying to analyse and interpret this type of data.

Another on in New York is the Meteo Grid project. This is to allow the democratization of weather forecasting, something that’s very important to a lot of people. Today,

weather forecasting is very much big, central computers, you know, that grind out surveys every four hours and so forth. But now what they're doing is producing metadata sets which are distributive (inaudible) very centres on using peer-to-peer networks and so forth, so schools and communities can do their own very localized forecasts on a much smaller grid than what is possible from these big central government sites. So, again, it's like this example of how data can now be migrated to various groups who can then use that data, mash it up with their own local information from their own local sensors, and come up with a very detailed forecast for their very specific area.

And, of course, the Sloan Digital SkySurvey. This is the late Jim Grey who was very instrumental behind this. Again, this is a site of astronomical data. Many of these services were built by students. Again, it is available to scientists and students and the public at large.

Now, because of this type of service, most of the large supernovae are being

discovered by members of the public as opposed to professional astronomers. So, by using various techniques of scanning all the images and scanning the data, it's the public who are making these discoveries as opposed to professional astronomers.

So, that's just -- the one last one is the Faulkes telescope. This is an eccentric billionaire in England who has funded this project. He was told it was going to be a few million dollars and he believed the researchers. It turned out to be close to one hundred million, I think. Anyway, this is two telescopes he's built, one in Hawaii, and in Australia and they are professional telescopes used by professional astronomers. But, also that information and data is being made accessible to students and schools in England and a couple here in Canada. And the beauty of it is, because when it's nighttime in Hawaii of course it's daytime for the schools in eastern Canada, at least, and in the UK, and the students -- there's all sorts of activities. The students work with astronomers

looking at real data and phenomena with these telescopes. And, again, there is this type of collaborative ad hoc virtual organization that's possible and the extension of science into a much larger community.

So, that, I hope will stimulate some of your thinking of the potential of what these Web 2.0 and participative technologies will enable, at least in the scientific community, as well as many other walks of life.

Thank you.

MR. STEWART: Thank you, Bill.

Le prochain intervenant sera Diana Rhoten, directrice de programme, Office de Cyberinfrastructure, National Science Foundation.

Diana.

MS RHOTEN: So I'm going to talk a little bit about some of the learning and knowledge production affordances of Web 2.0 for science. I just wanted to start with this clip.

--- Video presentation

MS RHOTEN: So in addition to providing some interesting statistics about some

of the usage of Web 2.0 services, particularly by kids, it is important to note that that clip which is available on-line was created by a high school teacher and then it was edited by myself using tools available on-line and then using creative comments licensing. I can present it to you mashed up, mixed up, re-mixed by me. So, just to use some of the tools of Web 2.0 for the purposes of the presentation.

So, what's the calculation for science? Technological capacity is increasing; we've all heard that. Moore's Law tells us that scientific complexity is increasing -- we've heard that Andrew -- requiring at the same time increase specialization and increased collaboration and integration.

If you multiply that by the fact that the generation of scientists coming into science are coming through a digital society we have what's potentially Science 2.0.

We need to think about the next generation implications for science by thinking about what the next generation expects from its

use on-line and its expectations about use.

What are some of the characteristics of Science 2.0? This is courtesy of Ian Foster whom some of you may know. It's been adapted by me. We see that there's changes in the nature and the size of scientific data. I'm not going to go through each one of these. But changes in the unit and venue of scientific communication.

As Bill mentioned wikis, blogs, project websites, become very much the outlet for both scientific data, scientific finding, scientific publications. But, we're also moving beyond just publications to simulations, visualizations, creating new databases. These are all new products that are coming out of Science 2.0.

It's also changes in the location and the structure of the social aspect of science. Science used to be done in co-located environments. Research centre was very much -- at least in the United States, and the investment of NSF in the 1980's.

We're now really looking at distributed science. We have scientists sitting in Canada, sitting in China, sitting in Australia, working on the same problem. This new forum requires new social norms as well as organizational forums.

We also see venues of scientific interaction changing. We've gone from community co's (inaudible) to science gateways, campus and national grids, to science on the Internet.

We've talked about science as a computation. From computational science to science as computation.

We also see this bleeding into all fields. We're moving from just the physical sciences to all the sciences, including the social sciences as well as humanities.

So, in Science 2.0 we really are looking at distributed knowledge production and learning. And I've created here, this is a chart borrowed from Dan Atkins. You can see we've gone from same time-same place, to same time-different place, different time-different place. Much of

the aspect of Science 2.0 is happening in a virtual environment.

I wanted to talk to you a little bit about some of the virtual environments or virtual exemplars of Science 2.0 that come out of NSF or are supported in part by NSF.

So BIRN is the Biomedical Infomatics Research Network. This is a geographically distributed virtual community that shares resources, including instruments to examine medical images and create diagnoses. This is an example of a closed virtual environment in the sense that this is really still left to, to use both terms, the high priestess of science.

So it is a distributed network but it's a very high level science. It's limited to professional researchers.

Whereas if you look down at nano hub, the science gateway, this was created in 2001. This is a good computing base but web-enabled portal that enables anyone to access scientific tools, do research demonstration and even run simulations.

So just by getting a user name, a log-in and authentication code, anyone, including myself, including yourselves, can run simulations around anything related to nano technology. They have a variety of different workshops, lectures, curricula and simulation tools available.

Just a few stats on them.

In the last year they have had over 25,000 users from 172 countries. They have had 5,730 users run 230,000 simulations. So you are seeing a real draw to this portal.

Eighty research publications actually now cite nano hub.

Bill mentioned that what I have there, the Sloan Virtual Observatory, the Sloan Digital Sky Survey, this is an example, as he mentioned, of citizen science and crowd sourcing. Just to give you some stats on the use there, 200 million Web hits in the last five years; 930,000 distinct users versus 10,000 astronomers.

So again to the point of citizen science and the democratization of science, you really see the general public being drawn to these

opportunities and committing their time and their energies to participate in science.

In the bottom right-hand corner, we have the example of Second Life. You heard Jim talk about Second Life this morning.

Science is coming to Second Life as education is coming to Second Life. There are learning affordances within Second Life, as well as communication affordances, in the sense that you can actually share, create objects, learn about objects, recreate objects and manipulate objects in a 3D environment.

As of my last check, there are approximately 160 universities now in Second Life.

Their activities range from giving courses and running conferences or lectures to trying to actually create new objects that can teach science in totally new ways.

There is a new area within Second Life also dedicated to science, called SciLands. Nature Publishing has an island within science. So we are seeing a flood of activity there, which is interesting.

The image I have on this website is of NOAA and its 3D visualization of a live national weather map. So this is a constant real time data flow that you can go and see what is happening with the weather across the country.

The last example I wanted to give you is called Sci Vee, what we call Science Vee. We actually recently, just very recently, funded this.

Sci Vee allows authors to upload an article that they have already published. It has to be an open access article obviously. They then can create a video or podcast presentation that they then synchronize with their publication so that you can view the publication at the same time as you view the author talking about the content of the publication. Sci Vee calls this a pubcast.

It's a new venture. We are seeing a lot of traffic there already. I think what is important about science YouTube in general -- and we'll talk a little bit about open access -- is that 15 per cent, only 15 per cent, of all

research publications at best estimate right now are open access and yet we are seeing a very high citation impact advantage for those publications that are going open access.

So this presents a real question of incentive versus some of the conflicts or some of the constraints with the publishing industry, which we can look at.

So while there are real opportunities for Science 2.0 -- and I think we are seeing them emerge and they will continue to emerge -- I think Second Life, for one example of an virtual environment, has laid the territory for some really exciting terrain. I think we'll see some increasingly complex and potentially proactive virtual worlds coming into place in the next 12 months that will really contribute to the learning and science potential, knowledge production potential.

But while there is real potential, there are real challenges. So I just want to go through some of what we see; why 2.0 hasn't had the effect on science that it has had in business

and industry to date and some of the obstacles that I see as explaining that.

Pax Informatica. So there are thousands of databases of valuable information, each of them with different conditions, different formats, different privileges, different goals. We have a very significant interoperability question which prevents some of the collaborative aspects of what Science 2.0 should look like.

Cognitive overload. The amount of new scientific information at a minimum is doubling every two years. Beyond the technical problems of interoperability, there are the social and psychological problems associated with trying to locate, sift, manage and qualify the number of papers available on your sub-specialty, let alone the specialties of those with whom you are working in a collaborative environment.

How do we manage this information is a major question.

Also the collective action problem. While we see this digital culture coming up, as particularly with the younger generation, I

think right now within science there is also a culture of secrecy, competition. Incentives and reward structures don't lend themselves very strongly to Science 2.0 in many fields.

A single author publication counts very differently than a co-author publication, let alone the creation of a new simulation or a visualization, all of which are incredibly important components of Science 2.0. How do we motivate the scientists within the fields and the institutions within the fields to recognize these as contributions?

Quality control. We've talked a little bit about the democratization of science. We have closed and open systems, questions about authenticity, validity of data, and how do you balance that with access in terms of scientific production.

Legal limits. I won't go into any detail, but the current generation has grown up with a variety of data sharing and format sharing and information file sharing formats. We are running up against questions of intellectual

property.

What is the right balance between open source and proprietary management of data, data findings, results, publications and so forth?

I think we've seen some very interesting experiments with Science Commons and CAMBIA. I think we need more. I think we don't know what the right balance is at a hybrid model of open source and IPAIN. Innocuna, Exhumina, these are all meant to be provocative little terms. But by 2023, when first graders now will just be about 23 years old, it's only going to take $1,000 computer to exceed the capabilities of the human brain.

So how do we think about this? What's the role in this? How do we think about computational thinking and the role of the human in that process and train them appropriately for that environment?

I just want to close with one of our new announcements from NSF that just came out on Friday. It's called Cyber Enabled Discovery and Innovation. I can provide more information

about it, but for the sake of time I'll just give you this quick summary.

This is a cross-foundation initiative. It's five years. This year the minimum will be $26 million for FY 08. Its aim is to transfer from science through innovations and advances in computational thinking. And by that, we mean computational tools, algorithms, concepts, methods and practices. There are three themes within the solicitation I've written in there for you from data to knowledge, understanding complexity and building virtual organizations.

The intent of this solicitation and the work that we hope will emerge as a result of the solicitation I hope will answer some of these and help us overcome some of these challenges that we see to Science 2.0.

MR. STEWART: Thank you, Diana.Notre dernier conférencier sera Mario Campolargo, chef de l'unité "GEANT et infrastructure émergente" de la Commission européenne.

MR. CAMPOLARGO: Thank you very

much for this opportunity.

Being the last speaker, it's difficult to say something that has not been said before. Hopefully, at least, I don't contradict very much of the very good presentations that have come before.

In fact, when we lead the programs in the European Commission in the area of science, we are trying to put forward this new vision that has been so very well explained here by my predecessor speakers.

Obviously we have global challenges and this implies a global approach. Some of those global challenges that we have seen before have a very high societal impact. The data deluge is something that is very present in all our day-to-day business and science. The replacement of wet labs by virtual labs has been very well explained by Andrew in the beginning.

So this all requires an improvement in the scentific process.

The aspects of cross disciplinarity become very, very important. When

we talk with the engineers in electricity companies and they were used to using a strong simulation in super computers to try to understand how the nuclear power plants could be influenced by a number of factors, now they are very much aware, for example, it was nothing that they were a few years ago. And for that they need information data and models from other communities, not exactly the ones that they used to deal with.

And all of that, like actually I was just mentioning at the last slide, I mean that raises the question of working together. I mean, collaboration is really a fundamental aspect for addressing the new challenges of science. We believe that it is fundamental to build this science through collaboration and research communities where research is having identified common goals and being able to put forward complimentary or shared information tools and knowledge. Being aware obviously of the research protocols, how you value each one of the -- well, the just example of publications is rewarding, is

one particular example of that. And when they are served by efficient means of collaboration then they can start building these virtual organizations.

And those research communities, it is not really to put them working in this context, although it may be thought that this would be very easy, more easy than with citizens. And as we have seen some from of the example here, it is probably not the case.

There are other aspects when we talk about collaboration, this type of virtual research, again very well displayed here before, is sometimes called a science with different names. But that falls on this path from the original empirical through the experimental and theoretical and computational science and that today really is basically using huge amounts of data, abstracts and model simulation, etc.

There is another aspect of the virtualization, that is the aspect that in fact the users, being them citizens or in this particular case researchers, they can work

together no matter where they are, unhindered by time or institutional boundaries.

And finally, obviously because we have globalization of our challenges, there is very very important aspects of global dimension, win/win situation is in this case very fundamental, especially when you try to find those international collaborations.

When we think about a virtual community I am not trying to put anything more formal or less formal. We see that those virtual communities around the world work together. And when we think from a funding authority, that is the case of the European Commission in relation to research in Europe, then we have to see where you can promote some economies of scale.

And those three areas seem to be relatively stable to promote some gains of efficiency and economies of scale by promoting the support of interoperability between different virtual communities and then allowing these researchers to focus on the top part of their preoccupations in their domain of research rather

than having at their disposal those facilities that are common to other disciplines.

In some sense, it is like if we look into three particular perspectives, the idea that we will be able to link all the facilities and the researchers around Europe or the world, be able to promote the sharing of the Federation on Computing instruments and applications and mimicking this also in terms of data and as we see the importance of data being more and more in the scientific process this aspect as acquired in Europe, a particularly relevant area.

All having in mind that what we want is to promote this virtual collaboration, is to promote this emergence of these research communities that can work with each other to solve these goals that otherwise would be very difficult, if not impossible, to address.

In some sense, we try to bring together the finest minds, some in sharing, federating all the best scientific resources and being able to do science in a different way by building those global virtual communities.

One first attempt is to ensure that you have a global dimension. This is what we have when we look into research networks, into the underlying ability to link all the facilities. This is particularly important for areas that we have not been analyzing here very much today, but it is like biodiversity, do require simply collaboration between different parts of the world.

In Europe and maybe in the States we have lots of good museums, lots of databases, lots of simulation models, but we don't have the pessimism, we have the people that can go and collect the information on that. So here are areas where this global dimension acquires (inaudible) a very big importance.

When we think about this ability to share and federate computational power, for example, or instrumentation, here is an example of a multi-science grid developed in Europe, it is now in its second generation, we have now more than 240 sites around the world. And interestingly enough, obviously this EG particular

project started from the concrete needs of the (inaudible) physics community, but is now developing into other areas. And you will see that you have more than 200 virtual organizations created within this multi-science grid.

I mean, not all the organizations have the same power or relevance or amount of researchers or amount of collaboration, but it is very interesting to see how these virtual organizations are being created dynamically.

Furthermore, it is interesting to see that virtual organizations, you know, tend to progressively get into more specialization and generate other virtual organizations thus allowing very much interaction with scientists.

And what you see here is just information that was collected a few days ago or a few months ago. And actually, I took this picture from a particular example, that is an example that is, although in the scientific domain as a far-reaching implication, this was a collaboration between Taiwan biologists and European biologists in a drug (inaudible), in

particular for Avian Flu. So what you see is a real time display of the cooperation done within EG for that particular purpose

When we think about data and all my colleagues all like very much the importance of data and we see this cycle of the relevance of data. I mean not just, as has been very well mentioned, not just the traditional way of publishing through the paper metaphor, but really looking into the aspects of making all data from simulations or captured from instruments available to the wide public.

Then we have to pay particular attention to a number of continuum that are important for us. First, the preservation and creation of data for the next generations of scientists. But also, like I emphasized before, the data that is particularly relevant or was particularly relevant for one particular community, they took care in generating and creating that one, but is now important for a number of other disciplines.

The same applies from one to

multiple organizations, from data to publication, from research to education, and for citizens in general, it has been very much highlighted in the previous presentations.

In our approach we look therefore into making sure that Europe has a number of stable genetic infrastructures. Those are just the code names for some of the products that are relevant in this context. But also it is very important that you look into the use of communities. You need to work with the scientific communities to help them to use the infrastructures available and not just to use them, but becoming major actors in influencing the way they develop.

And this is the case, for example, for the ability that European radio astronomers have today, to link and collect information from all over the world or the ability to put in place large testing infrastructures for ICT communities looking into what could be new architectures for the future internet.

The same applies when we think

that neurologists that have been collecting a massive amount of data through scans of the brain and are looking, for example, to illnesses like Alzheimer's, the need that they will face now to compare this information, to have simulation models that can run over their images and derive some indicators.

I mean, the same could apply when we think about nuclear fusion and, as you know, Europe and the world moved into creating this new gigantic initiative called ITER, but you need not just to build it, you need to simulate it in advance to predict this behaviour and cell projects like Euforia are looking to those aspects.

When we look into the particular aspects of data, obviously today already we have references to a combination of satellite data, to ENC2 data, to censors, that's the objective, for example, of GENESI-DR lead by the European Space Agency that tries to combine this information making it available to the public, but also making sure that the interoperability aspect, that have

been very much referred here, will be taken care. And the same would apply for virtual centres for the astronomers or for the biologists.

So what you see here is just the way the scientists in Europe are using these opportunities opened by the internet that is really like one of the titles and one of the objectives of the conference in Seoul as fuelling creativity.

I think that we are basically experiencing it with our scientists in Europe and around the world but, as we can see, the impact of the infrastructures is not just in terms of the science in stricto senso but you see very much in the line of the examples that have been given here before, a lot of impacts outside science.

And just an aspect that is quite interesting, the civil protection in a number of countries in Europe is now trying to use this type of technologies to measure information from meteorology, fires, monitoring the sensors in forests, et cetera.

This empowerment of users is very

important, not just as consumers but also as producers, but I think that we are all aware of the social changes and the sociological implications of all that.

An experience that we will also launch a very soon is this interface between more formal, more advanced green infrastructures for science, way more citizen grids and making sure that they can interact in a transparent way.

The access to information, the trust, the simplicity, the services, the way that we can use becomes very, very important.

But there is one question: Do we really need to rethink from the architectural point of view the Internet? That may be the case.

In Europe we launched, actually like NSF has done similarly with initiatives like FIND and others, we launched the FIRE initiative in Europe that looks into the future Internet research and experimentation, looking into the aspects of making a simulation models and then try them in large scale to see which mechanisms could be put in place to satisfy the needs that we are

observing coming from these communities.

Obviously all the aspects that are so important for all of us, like values that we all have in our society, fundamentally is to invest in people. We have seen through the examples in research that if there is not a huge investment in training and education, those services, those systems do not go out of the cocoon where they have been initially launched. So if we want them to spread we need to really invest in people.

Overall, all those put together, will contribute to this knowledge society that we are all, from politicians to citizens and researchers, contributing to this knowledge society.

Thank you very much.

MR. STEWART: I would like to thank all the panellists for taking the time they were allotted and not running over and preventing my having to get out the big hook.

We did start late. We finished on time, so to speak, but we started late, but I have

been assured that if we run over by 10 minutes to facilitate the discussion we will still get to lunch before the other group.

C'est votre tour. It's your turn. Your questions. Vos questions.

There are microphones here. If you would go to the microphones to ask your questions I should be grateful.

Sir, you are the closest to a microphone.

QUESTION: Right.

MR. STEWART: Would you indicate who you are before you give us your question. If your question is specifically for one panel member or another, would you also so direct it?

Thank you.

MR. LEVITT: Sure. My name is Karl Levitt and, like Diana, I'm from National Science Foundation and part of the FIND effort, but I don't want ask a question about that.

So a question about DRMs. Bill quickly brought it up and then didn't go further. Diana went much further with it.

So this is a tussle and I'm just wondering what we can do about it.

In the previous session it was mentioned that, well, if you try to use cryptography or cryptographic sealing -- some kid, I can't remember what country he said, let's say Finland as an example, some kid will break it, okay, and let the whole world know about it.

So then he mentioned a hybrid model and that was the most intriguing thing and I'm wondering what we can do, because I think that's what we need here. Okay. Yes, you really want to protect the next Beethoven, but we also want to allow the free exchange of data and we don't want the networks to impede this particular availability with data.

MR. STEWART: Questions? Comments from the panel?

MS RHOTEN: I think there's lots of openness and proprietary intellectual property rights at different phases of the research cycle as well. I think there is data sharing and then there is publication sharing, and so forth and so

on.

As I mentioned, I think Science Commons is doing some very interesting work. They have done interesting work with creative comments in terms of optional licensing approaches with copyright, and so forth. They also have some new projects under way in terms of material sharing and where the rights intervene at that point.

CAMBIA is a foundation in Australia. They are doing some interesting rights around open patenting, so patenting processes and then making them open as a way of protecting some of the opportunities of these data going forward, genetic data, and so forth.

I don't think we have the answers right now in terms of what is the right balance and I think it is a question that a lot of people within NSF are asking, a lot of people without of NSF are asking.

I had the opportunity to interview someone the other day about a new virtual world platform. This particular person was absolutely adamant that open standards were destroying the

potential quality of what could happen within a virtual world, both in terms of the content being produced by users, but also the activities being conducted by users.

I don't think that's necessarily the case. I do think that there is -- if we look at the literature we know there is incentive structure-building around the protection of rights, but I think we need to find that right balance.

MR. ST. ARNAUD: Just to add a comment, in the time of Beethoven and Newton and many other famous people there virtually was no copyright or digital rights protection or any type of protection and they still became famous and well-known.

I think in the 1860s the whole copyright issue came about and that was to protect the property of the publishers. We have to remember who is the beneficiary of these types of technologies.

MR. HERBERT: Can I pick up a comment there?

Of course Newton wrote his results in an impenetrable form of mathematics so that other people couldn't steal it from him.

--- Laughter

MR. HERBERT: So if you haven't got systems like copyright and patents people will find ways to be secretive.

Working for Microsoft I live in this tussled space and there are some observations I can make. There aren't magic answers.

Running a research lab we published our research in the open literature because we want to subject it to peer review. It's the best way I can get our research calibrated.

We also patent quite a lot. We use the American patent system because we can patent after we have disclosed in the scientific community. That's an interesting discussion in the European context where the systems are quite different.

In the world of commerce, companies like Microsoft have to decide, sometimes

we are told, where the interoperability points are, where the benefits of the community of revealing proprietary information is in the interest of the market and you as a company. Sometimes that's done by regulation, sometimes it's driven by commercial and market pressures.

There aren't magic answers. I think we do live in a very mixed economy of ideas. We need to make sure our systems are open to that mixture and finding the right approaches.

The manager of intellectual property in the pharmaceutical area is very different to the manager of intellectual property in the software industry for example.

In the case of Cambridge University, which I know very well, they got in a terrible muddle by trying to impose one discipline's model on another and had all of their academics get very upset about those issues.

So I think there aren't magic answers. It is balancing commercial interest with scientific collaboration and openness.

I think there is a lot we can do

it information-sharing in addition to worrying about issues of information rights management.

Also information standards, what metadata we put on the information so you know what it is that express issues about usage policies, where it came from, who has been tinkering with it.

And the thing which I think will make information rights management a very big challenge is: How am I going to decrypt a data file 100 years from now when the person who knows the secret is dead?

MR. STEWART: Sir...?

QUESTION: Thank you.

I'm Richard Hawkins from the University of Calgary Innovation Lab and Complexity Science Group.

First let me put my credentials on the table here because I'm going to say something that is critical later.

I'm an economist. I work with physicists and biologists doing much of the kinds of work that Andrew was talking about earlier. We

are in the process of trying to build in fact an international network that uses all of these bells and whistles as much as we possibly can.

So I'm not a Luddite here, I'm not sceptical about the use of technology in any way shape or form.

But this is an OECD policy forum in the main thing that we have to consider with all of these possibilities is that we have finite resources and we have to determine where we put them.

My colleagues in the physics area know about black holes more than I do, but I know about economic black holes and I'm afraid that I see some dangers that this might become one of them if we're not careful. I think maybe if we could bring the conversation down from 37,000 feet to maybe a couple of hundred feet, to the level of the scientists actually in the laboratory and what really goes on there, I think this might help us here.

In the first place, I don't think you meant to put it this way, but science

influences computing as well as computing influences science. I mean, in our group, we don't actually go to the computer until we have done science. The computer allows us to do calculations that people predicted we could do in the 1920s, but we never had the gear. We can do it now.

And that very often influences the way our colleagues in computer science think about computing, think about these environments. So we shouldn't think in terms of the computer as being some kind of determinist element in science. I think that would be completely wrong and it might put some of this in perspective.

But the other thing is, you know, democratization, it is easy to say but, you know, I work in this multidisciplinary group but I am never going to really understand particle physics I am sorry. And they are not really ever going to understand the Markoff universe in economics either.

So, you know, there are lines of demarcation. We can participate together, we can

learn together, we can learn to do science in different ways, which is what we are trying to do. But participating is not the same as doing and I think we need to be very very clear on that. And so the justification for building large networks because we are going to include everyone in science, I am sorry it is just not going to happen above a certain level.

Also, I think we have to be careful about these claims of increasing the speed of science. Certainly, it has, this can be verified easily and empirically, but increases in data do not necessarily indicate increases in quality.

I will give you one example, the amount of money spent on cancer research has increased nearly exponentially in the last 20 years, the death rate has gone up proportionately, it has not gone down. So obviously, there are some breakthroughs that need to be made and we haven't gotten there yet.

So I am 100 per cent in favour of doing this, but I would caution that we need to

think more closely about what this is going to cost versus all of the other things that we do not yet resource adequately. I am very fortunate and my own research is well funded, but I have colleagues who are more eminent in their fields than I am in mine who get by on peanuts.

We have to think about those kinds of things. Maybe there is a way that we can orient this to make that environment more productive for them, I don't know. But I think building the infrastructure and seeing what it can do and then making all of these claims is probably not going to lead us to the result that we need. So well done, I am all in favour, but I would just offer this slightly subdued message.

Thank you.

MR. STEWART: I am going to give the panel an opportunity to comment on that in a moment. But I know that we also have some feedback coming from the people who are participating online. So I am going to ask for that feedback.

QUESTION: Actually, we are seeing

some good traffic on the blog. This is not actually a question from the blog though, it is wearing my science library hat. So we heard a lot about how computer science and computer networks are transforming and impacting science publication. Representing a science library, where does the panel see the role of the library, the science library, the research library in the e-science workflow?

And, you know, feel free to be as critical as you wish. If you don't see a role, that is certainly a valid answer. From my perspective, I think there is lots of roles the library could play, particularly the issues of data curation and data archiving were mentioned. The issues of very long-term access to data, being able to access data 100 years later when you have gone through multiple generations of format change. I am interested in your feedback on the role of the science library.

MR. STEWART: Okay. So I am going to invite the panel to comment on that question and on the previous set of comments. I would say

in terms of the previous set of comments, the Science 2020 paper that was spoken about earlier in this panel very clearly talks about the relationship in the way that science influences computing and vice versa. And in that paper that circular nature of the relationship is discussed quite fully. So if you haven't read that paper, again, I commend it to you.

But panel, comment on the questions about appropriate use of resources and the question of the science library. Your comments please.

MR. CAMPOLARGO: Thank you.

I will not be exhaustive, because obviously the question is very interesting. I will (inaudible) and I want my colleagues to..

But I just want to make a reflection. For example, in the case of Europe where we have a number of cancers, not all of them at the same development. Obviously, all what we said here, I think that does not imply that we don't need to invest in physical infrastructures and in instrumentations and things like this.

But when you look to Europe, there is a diversity of speeds and amount of resources that can dedicate then the emergence of networks and of all this ability to share infrastructure is very fundamental. It is not just something that you invent. With out it, you really put aside a lot of huge human resources that are an immense capital for a content like Europe.

And you can extrapolate this, not just in theoretical terms, but you can really put this in very good perspective when you look globally. I mean, you know, unfortunately for our good friends in Africa nobody is more expert in Malaria or in AIDS than themselves. I mean, if you look into some other sciences, I mean, you can't simply afford to duplicate exercises like building large atomic colliders, for example.

So I mean, I don't think that any of us will advocate that, you know, computational science drives the way we do science. But I mean, there are a number of areas where we could not simply do it without. I mean, when we think about prediction there is no way of predicting diseases,

propagations or things like this without anything. So, you know, it is not a panacea for everything, but it is a very huge contribution that we have to do.

Now, monitoring, trying to understand the effect on the scientific process, trying to understand that if the investments that we make in networks, in grids do have an impact on the way science is done is very important and we are not yet with the tools that we need. I mean, as an interesting exercise that has been published just a few days ago about what we need if we call are you ready, being ready, the recession, the development indicator.

So we are trying to work with a number of indicators to try and to know if the investments that we do in one particular country in deploying high-speed networks for researchers or putting more grids or more computational power, etc. has an implication on the way we do it, those are difficult, but they are fundamental processes that need to be put in place.

And I think that the question, I

mean, not just by the explicit areas that it implied, but inducing also a question of, you know, what is behind all of that is very important, but I think those are just some elements of (inaudible).

MS RHOTEN: With regard to the computer science driving domain sciences, I hope I didn't imply that just to speak to the cyber and navel discovery and innovation solicitation that has just come out. Very critically, that solicitation is designed to create teams and expects and requires teams to be composed of domain scientists, computational scientists at large, interpreted broadly. With a very specific goal being that this not be computer science driven, that it be domain science and computer science meeting together so that the solicitation will actually create new infrastructures that serve the scientists to do the science that they need to do to transform science in their domains. We are very very adamant about that.

And I can tell you, having sat on the committee who drafted that solicitation, this

is a very very big commitment, a very big important aspect of the commitment of NSF to this. I think we have learned from previous mistakes where science and technology, the capacities, haven't actually been at the same point in their development. Think what a really unique historical moment where the domain sciences can work in strong partnership with the computational sciences to shape something that is transformative.

I completely appreciate your comment about how much do we invest in infrastructure? We don't right now I think it is fair to say. We don't have the metrics, we don't have data to know the impact of the infrastructure investments we are making on scientific production and innovation.

I am a sociologist, I come at this from a very different perspective. I study scientific collaboration. I can tell you in my data scientists still collaborate primarily via email. They don't collaborate that much by wikis, by blogs yet.

But I think part of what I was trying to introduce in my presentation is that my scientists in my dataset right now are 40 and older. The next generation of scientists aren't going to be limited to email, they are incredibly literate in all of these digital technologies and we need to think about how they can drive and will drive the way science is practiced.

But let me just pose this back to you as just an exercise.

So if we think about the learning affordances of virtual organizations -- not the scientific production, let's look at learning -- and you had to make a policy decision between spending $11 million or $100 million on rehabbing all of the laboratories in your high schools versus creating 17-25 virtual laboratories to which high schools could access shared tools, online data, real scientific data, how would you weigh that policy decision?

I think it's a good question for us to be asking.

MR. STEWART: Does anyone else on

the panel wish to comment on that?

MR. HERBERT: I would very much like to agree with the comments others have made. Obviously computing itself is driven by science. It's physics that gives us Moore's Law. It's Moore's Law that gives us the faster computers that let us run more software. I fully accept that.

I think I'm concerned we are over-focusing in this conversation on infrastructure. Infrastructure is important and grand projects are always very exciting to do and big investments and particularly represent funding agencies. Fortunately, I don't represent a funding agency.

There are actually some nice examples of where e-science has taken away the need for infrasture.

The most recent Boeing aircraft was completely designed without the need for a wind tunnel. So that's one piece of infrastructure that went away. It was done using computers.

The theme I was trying to

communicate -- and perhaps I didn't do it so well -- is I think what computing is bringing is productivity to science by taking on some of the drudgery and making more information accessible in the same way that computing has accelerated the pace of administration in the office, accelerated the pace of business through things like e-commerce and so forth -- that I think is a space I wanted to explore -- and has dropped the cost of computing while doing so.

Computers have taken over those roles because they are cheaper than infrastructure and people that we had before.

It's about getting to the science faster.

It's certainly not a land grab by the computer scientists to get all the scientists' budgets or even do their work for them.

What I was trying to postulate is we have a growing need for a class of person in science who is happy to work at the intersection between disciplines and who has some very good competence in working on dynamic multi-scale

models, building and manipulating those, pumping data through them, computing with them and helping the people understand the physics, the economics to bring their ideas together in a formalized way and manipulate them.

To the libraries question and back to infrastructure, I think what used to be super computer centres are increasingly going to be super data centres, and that brings them into a relationship with the libraries. And the person who asked the question I think nicely identified the roles.

Ironically, at the start of computing, Tom Watson, the founder of IBM, said we'll only need five computers. That statement is sort of true. We've always had five big computers, five big super computers or five big infrastructures and we'll still want five big infrastructures because some problems are that big. But most science is done by scientists in small laboratories, in small groups using the PC under their desk, which these days has the horsepower of many of the high performance

computers of five years ago.

MR. STEWART: Very good.

Jack, you are keeping us from lunch. So I'm going to ask you if you can make your question a policy question that can be read into the record but we are not going to have time for the panel to respond.

You have been standing there for some time.

QUESTION: Thank you, Walter.

Jack Smith, National Science Advisor's Office of Canada.

Diana opened the door on sociology and I would like to pose the question briefly for the record: What is the frontier for the social sciences as part of this endeavour in the future?

MR. STEWART: Thank you, Jack.

I'm afraid I have a terrible confession to make. They have actually broken already. I apologize that you are going to the back of the line. It's my fault for not checking sooner.

In order to conclude this session,

Graham has an announcement.

Thank you for your attention and thank you again to the panellists.

--- Applause

MR. VICKERY: Thank you very much, panellists.

I have two reminders.

One is there is going to be a presentation by IBM over lunch. It is now going to be ten minutes after the allotted time. So it will be at ten past 1:00.

And the people on the panel and some of you in the audience were going to have a little break-out session beginning at the same time as the IBM speech -- I apologize to IBM -- in Room 304 upstairs; just a break-out session to actually discuss how we might want to follow up on e-science for the Ministerial in Seoul next year: what we might want to do preparing for that meeting and what we might want to do afterwards.

The people know who they are who are going to go to that break-out session. It is room 304.

The lifts are down at the very end of the corridor, up on the third floor, for those people who are joining us.

So that will start at ten past 1:00.

--- Whereupon the session concluded at 1235