SESSION REVIEW:

'Opening Urban Data: Platforms for Service Creation and Decision Making'

15th June 2011

Introduction: Explore and Learn With Your Peers

This collaborative effort—led by the World Bank, Cisco, and the Urban Age Institute—is designed to engage government officials, private sector organizations, international development agencies, academics and non-government organizations in a video-based, collaborative process of planning and critical thinking about *how best to design and develop 21st century sustainable urban areas and cities*. The aim is to convene a global community of peers, to debate and learn from some of the leading thinking around innovative urban projects.

Session Overview – 'Opening Urban Data'

The move towards open public sector information presents greater opportunities for **transparency, participation and engagement** across urban communities. The opportunities, challenges and directions this takes us, however are largely still to be discovered. **How can urban data be sourced** beyond public institutions? What policy decisions are required, and what mechanisms exist to **enable heterogeneous data flows**? What platforms are emerging to be able to enable better decision making? **Where can lower-income cities get involved**, and is this a luxury for established institutional frameworks only to grapple with? What **evidence of the drive to open data** movement are we seeing from small and large communities, and the benefits that are created?

Led by Industry Experts

These questions were explored from the perspectives from some leading thinkers on urban systems, from several city administrations, World Bank, Cisco and city networks.
Introductory remarks were made by three persons (each one for a maximum of five minutes):

- Emer Coleman – Director, Digital Projects, Greater London Authority
- Jay Newman – Advisor, World Bank and CitiStat, City of Baltimore
- Ger Baron – ICT Cluster Manager, Amsterdam Innovation Motor
- Aleem Walji – Innovation Practice Manager, World Bank Institute
- Dan Hoornweg – Lead Urban Specialist, World Bank
- Ed Borden – VP, Connected Environments Ltd and Pachube, New York
- Phil Ashlock – Program Manager, OpenPlans, New York

The moderator was Gordon Feller, Board Member, Urban Age Institute and Director, Urban Innovation, Cisco Systems

Over 20 attendees globally in 6 locations:
San Francisco, CA
Sao Paulo, Brazil
New York, NY
Washington, DC
London, UK
Amsterdam, Netherlands

Summary of the main discussion points:

Emer Coleman, Greater London Authority:
Three key lessons on 'data release' in London:

1. It's not about Data: Risk aversion, despite political will, administrations are resistant. What is fit's used wrongly
2. The state won't do this on its own. Success comes through three interconnecting circles/stakeholders: State, 'Digital Disruptors', Media.
    - Small advocacy group in the GLA
    - Open call out to developer community to engage
    - Transport and Crime were deemed the first data sets to chase
    - Don't get into a debate about formats, to avoid 'flame wars'. Just get the data, in any format released
    - Go ugly early. e.g.: no tight definitions of 'open data', eg linked
    - Engage digital disruptors to challenge resistance arguments from authorities, based on technical reasons
    - Conversation between the GLA team, developers and the media to blog on, lobby and challenge resistance
    - Use twitter as a powerful tool for real time conversations
3. Open Data is Emergent. So ROI justifications are pre-emptive at this stage. Many more data sets need to be released firstly. Many more data sets, other organizations to release their data, and to be linked.
    - Public Data Corporation plans may seem to be counter to the ethos being fostered here.
    - London Data Store is a different model to data.gov, and was launched in advance of the US federal site.
    - Public data sets 500. We would like train operating companies, mobile operators, utilities
    - Data privacy is a limiter to further data release thus far.

Questions and discussion from participants:

Any successful outcomes so far?
Response:
- Transport for London (TfL) data – Tube journeys, Jam cam, Cycle Hire
- Placr – transport modeling is proving a successful business, using the data store data.
- Visualisations – Inspirational uses, cycle hire, election results.
- Focus on real-time data
- Important also is the demographic, economic data for policy use.

Any publications/value cases on what the data means for the city?
Response:
- Ordnance Survey conducting research on Open Data
- No official reports as yet, but will have in the medium term.

What has been the most effective way of getting people onboard with this agenda?
Response:
- The real issue is not political will. However, it may be that maturity over time, as seen in the US sees a retrenchment in the commitment to open data
- It is a need to redraw the contract between what is to be a public official, and what it is to be a politician. This doesn't quite work currently. The issue is we are trying to bolt on 21st century tools to 19th century structures of governance. The shift from

broadcast, to communicative governance, to engagement . To use an analogy: we have public institutions populated by engineers, when what we need are poets.
- Data is the Trojan horse for this. It's not about the data!
- Emer's position is unique. She's on a short term contract, and is not a career civil servant. This liberates her to push the silos, and challenge convention. She is 'of the state, but not part of it'. Also she's senior enough to carry the Mayor's weight behind her.
- Example, TfL relationship has moved from adversarial, to collaborative, as she has continued over time to probe attitudes and perceived arguments against.

Jay Newman, World Bank:
CityStat came out the city of Baltimore, led by Mayor O'Malley. Applying basic common sense IT usage and management practices to city services and agencies. CityStat enabled the installation of the Call311 service, which wasn't able to be set up previously. This allowed transparency, adding participation, engagement, to request services in a transparent way, and knowing resolution times, through the use of ICT.

Discussion from participants:
- Political weight is very important. Example on the new UK government's attitudes to data. The government released the data during the Icelandic ash cloud issue. The London Mayor believes that government is not the answer.
- Illustrative of the need for political will, In DC, the policy of the CIO is, if that if information can be demanded by a FOIA then releases the data, before it costs more money.

Ger Baron, Amsterdam Innovation Motor:
The city of Amsterdam is very much focused on Open Data. The strategy of the city is to work on opening their data, for three main reasons:
1. Transparency: Good government practice, tracking expenditure.
2. Citizen engagement: and involvement in services.
3. Economic opportunity: to share insights.

Example: Weather forecasting service. Opening up the data has promoted the development of an industry of information providers, application developers, even though the original forecasting organization has shrunk itself.

The business case, in that example is there, the issue being that it does not reside in the same silo.
Apps for Amsterdam is in its Beta phase currently: www.appsforamsterdam.nl
Experiments with Cisco on visualizing environmental data: www.urbanecomap.org
This has promoted active discussions amongst functions, even though citizens themselves haven't been so engaged. People in the city have been surprised to see their source data repurposed, and used in a new way.

Now there is a discussion about how to get the data, and what is the data? But, resistance is strong in certain areas, despite political will. For example, data on where parking places are, was seen as 'controversial'. The effectiveness of commercial agreements with service providers become easier to monitor. The Alderman has insisted that this is exactly why such data should be released.

The city has organized an 'Ambassadors Network' where a community of app developers have organized a contest: www.appsforamsterdam.nl. Even with a small data set, community enthusiasm is tremendous. The event had over a thousand developers coming by to a hackfest.

The city is developing a data market, where all data can reside. More events will follow this year.

CIO has defined the requirement for every city service to come up with a business case on how open data can help his scope save money.

Timelines:

- Launch of data market
- Growth of data sets
- App development events later this year. In six months to have an organized network with an external 'ambassedors network', civil servants, app developers etc.
- Aim to be completely open across the city in two years.

Phil Ashlock, OpenPlans:
Experiences in NYC, working with Metropolitan Transport Authority (MTA), they had a litigious approach to reuse of data. OpenPlans acted as a liaison between the developer community, those who were getting sued, and the MTA. The ensuing brokered conversations led to the MTA dramatically changing its approach and policies to open data.
MTA now has an outreach strategy, and help facilitate the developer community to engage with their data.

Key point on the value proposition of open data: it's not just about those new applications which can emerge, but often about those applications that have already been built, that can be leveraged and used when data is open. A good example is Transit, where a lot of data has been standardized through the general transit feed specification (GTFS), an open standard, established by Portland TriMet and Google. City-Go-Round, highlights all the apps that work with this standard, and this can illustrate to new data providers, the value proposition, and how to open their data to this huge ecosystem.

This has informed OpenPlans in how to approach the civic issue/311 tracking space. CityStat was the first to establish the 311 service with CityStat. New York has popularized this. New apps like FixMyStreet, and SeeClickFix and in Amsterdam the Verbeterdebuurt does a similar thing. These applications were not integrated into existing 311, or CRM systems. At the same time, cities are developing their own applications, which are integrated to the city's systems. So OpenPlans saw a similar opportunity to provide the bridge between such developers, and the city functions to bring about an Open311 service.

The need for standard data specification is crucial to this. OpenPlans have just established the 'Georeport version2', which is being used in SF, Boston, Washington, Baltimore and commitments from NYC, Chicago and Toronto. This enables an ecosystem of tools for real-time CityStat type monitoring and reporting. The specification only covers the issue reporting requirements of 311, but not the information requests, which typically make up the main volume of calls to call centres. So, in NYC, the specification is being expanded to incorporate knowledge based capabilities.

Stela Goldenstein, State of Sao Paulo and Metropolis:
We need a mix of 1) cultural change, institutional reform in order to open data, plus, 2) the need for public will to engage. These two together are more important than technology itself. Inherent in these, they will demand technology to enable these objectives.

From the perspective of large cities, the challenge is the sharing of data between cities. Cities that share a territory, but not a government. For example, the city of Sao Paulo is an enormous conurbated city. The same in Bogota, Mexico City, Buenos Aires, New Delhi. Cities that have services in common, people who travel between different jurisdictions, and demand services in different areas. But the authorities to not share management or data.

Dscussion from participants:

- GTFS now allows you to compare services across cities. Then highlighting those cities that haven't signed up to the GTFS standard. Example from history is the creation of the standard metric system: The use of standard measurement is really about stretching the hand of governance. The agreement on a common metric system in Europe established a level of unification and identity.
- Resistance to information is all about control and resistance to standardization. Open311 and GTFS illustrate that the use of such standards can come subversively, before the government and institutions understand that they are ceding control. From the government side, if you can argue that it saves them money, they ride on it.
- There are a lot of small groups using open platforms and open data formats to illuminate transparency in their cities. Transparent Chennai is using OpenStreetMap (OSM) to locate public facilities, bus routes, to inform how efficient the government is using public resources.
- For public toilets, records were only available in written format in each district. This was then translated into geo data, once the data was shared. So, subversive data acquisition tactics seem to be needed in this space in the developing world.
- In the US, Manor, Texas is the poster child for the OpenGov movement. The CIO used open source applications to start cataloguing all government properties. QR codes were placed on every publicly owned object. An OSM planning challenge attracted participants from all over the world.
- For small communities, lacking huge resources, open platforms, and open data are very attractive. What you need is inspired leadership to take this approach forward.
- Urban Vision in India has implemented SeeClickFix in Bangalore. However, the government was not responding to the issues being raised. The need for the right governance model is vital to resolve the disconnect between institutions and citizens in this space.

Ed Borden, Pachube.com:
Pachube provides the tools for people to take hold of the Open Data, Internet of Things movement for themselves. The platform enables the sharing of real-time data, the discovery of new data. We and the community have created open source tools and applications, and we have also tagged onto the emerging open source hardware developments. One such device, the 'Nanode' is a hardware kit, which is a network device, that sensors can plug into.

We have published the Internet of Things bill of rights. It sets out what we feel is inevitable, even if we don't subscribe to this all yet. Essentially it says: people own the data about them, and data that is collected about them or their environment is theirs to utilize.

Pachube is building the marketplace for individuals. However, for this to really take off we have to create the environment where hardware companies and developers can all benefit economically. What is happening seems to be a lot of politics, but beyond this, when individuals have the tools, we see them moving way beyond what the government is providing.

In Japan, following the tsunami, and nuclear crisis, radiation data has been crowd-sourced, in real-time. Which is beyond the official, locked pdf, of indeterminate dates. This revealed better data and transparency of what was happening around evacuations.

In New York, using sensors, placed by a citizen at sewage overflow outlets, output data has been shared on Pachube, to inform citizens of the live situation with this civil infrastructure. The Dontflush.Me application enables to see for themselves to status, to influence their own behaviour patterns on water consumption. The city could provide this, but isn't. and is now being provided by other methods.

Discussion from participants:

- The significance of Pachube is in enabling crowd sourcing, tapping into other data sets is hugely significant to the development of public services; what and how it is being provided. OSM is a great example of a disruptive operating model.

Aleem Walji, World Bank Institute:
In the context of where is Open Data happening in the developing world, the opportunities etc, Kenya is opening an Open Data portal on July 5th. They are still exploring what data to place on this. The World Bank has been advising on how to proceed with this, and how to host, eg Socrata to get the data up quickly. In terms of data, Census data is the big data set, but also they are looking to open up all public expenditure data and mapped.

The opportunity is to show which data is of most interest, and can have immediate impact. The Wonkish data is one thing, but transport data, which they may not be so aware of, can engender political kudos, which then creates momentum for them to do more.

The suggestion is to take an African community and develop the learning, approaches and new ideas to, to promote further adoption in these countries.

The example of the Sao Paulo situation rings true. There is excitement around the technology, but the political will isn't there. So, quick wins such as transport data can be very helpful.

The World Bank has a huge amount of data, which was not in an open format, restrictions on re-use, and it wasn't free .

The Bank opened up its data, applying three principle. Data should be: 1) Free, 2) Easy to find, 3) No restrictions on use and re-use.

More traffic going to the data catalogue than to the home page of the World Bank. Millions of unique users in the first couple of months of opening up the data. The bank thinks about its clients, rather than the users of information.

The bank challenged people to do things with the data. We launched over 7000 indicators. And split these into sub sets.

The amount of applications that have been created out of this release and engagement is fascinating. Applications that we wouldn't have thought to create. One of the favourite examples is one that took the rainfall data, and took data weather data, and developed an application that told historical, projected future rainfall, and what crops can be grown in those circumstances.

World Bank data is however, highly aggregated. There is a need for disaggregated, granular information for contextual awareness on local communities. The challenge is how to do this disaggregation, by obtaining more granular data from official and crowd-sourced means. The hardest part is not generating the data, but the change process: the political, will, the systems etc. we are grappling with how you do that.

Discussion from participants:

- Not a lot is happening on Open Data in developing countries currently, so to tap into the examples we hear in this discussion are very powerful, to share with many other communities.
- I suspect the policy of Kenya has been driven hugely by the experience of Map Kibera. Developed using OSM, volunteers have mapped the densest slum in Western Africa. The Kenyan Constituency Development Fund (KCDF) investment activities in Kibera. The community took pictures of where developments were, geo located them, and triangulated to the official records of investments. Cases of completed projects, as reported, were exposed through mapped pictures to not exist in the way stated. I suspect this is driving the Kenyan government policy on Open Data.

Dan Hoornweg, The World Bank:
The value of cities like Sao Paulo can't be underestimated. Cities in developing countries need to see some comparable initiatives, beyond the likes of New York, London, San Francisco etc. they get the concept. But, to influence those cities in Africa, Asia, Latin America, the need for comparable initiatives is crucial. Particularly this is the case for India and China.

When it comes to data from cities, you have to crawl before you walk, before you can run with such real time data application, when you don't even have reliable data for waste, water supply, for infant mortality. The Global City Indicators Facility (GCIF) started with eight cities collecting 1100 data points every year. However, only two data sets across all of these, were comparable. The GCIF has defined an ISO standard level of simple data sets to enable comparability. We really need Sao Paulo, London, Amsterdam, to join GCIF to illustrate to other cities how to do this.

Through the Urban Knowledge Platform, and integrating urban metabolism perspectives, the World Bank has a secret set of data. Green data books, data on currencies, etc. We are thinking about doing a green data book on cities, maybe with the 200 biggest cities. In the World Bank, doing business is close to being cut off. Comparability rankings, partnering with companies and other agencies, would reveal this data.

We have been looking a GHG data, in order to enable financing. Working with Amman, they wanted to become the leader in carbon financing, and were able to provide the required data, prove their measurements sources in one weekend. This compares with The World Bank taking years to source correct data from other cities.

We should provide the data, for the sake of having good data. Rather than looking for the use, just having a common baseline of data is hugely valuable, which applications and services will fall from.

Discussion from participants:

- My work is at the aggregate level on the urban metabolism. I've been piecing together data points for sustainability, GHG calculations. I'd be keen to know how such micro level data can be combined for the total city view.
- The example of Amman, shows that data on that cities urban metabolism was compiled in one week. We'd like to see this rolled out to multiple cities.
- There is also the city to state and province relationship to consider. In sharing data streams, part of the problem is that there is a funding stream that goes from the state and goes back to the city. How do data streams relate to that and the federal government. Census data also drives the disbursement of funds. A larger regional, sub-regional planning framework across this larger constituency can feed from data release and sharing, beyond that which census data can reveal.
- A better connection between data producers and users of information will provide a solution. When we are open to discuss and collaborate, we can really define what data is useful and needed. It is not enough to organize data alone to define uses.
- Many cities worldwide have realized there is an economic case for being the greenest city. The build-up of urban data metabolisms can support the case for this claim to be the greenest city.

Thanks to all participants!—we will convene again on July 12th to discuss *'Work-Life Innovation: What do changes in our lives, work, social experiences mean for urban development?'*