STANDING COMMITTEE ON INFRASTRUCTURE AND COMMUNICATIONS

INQUIRY INTO SMART ICT

SUBMISSION BY SMART INFRASTRUCTURE FACILITY

Contact Details:

Ms Tania Brown
Chief Operating Officer
SMART Infrastructure Facility
University of Wollongong NSW 2522

Ph: 02 4298 1431
tania_brown@uow.edu.au

INTRODUCTION

The SMART Infrastructure Facility, at the University of Wollongong, is a multi-disciplinary applied infrastructure research centre. SMART stands for 'Simulation, Modelling, Analysis, Research and Teaching' and was established in 2011 following a $35 million grant from the Federal Government. SMART provides strategic advice, research and skills training across the whole infrastructure industry and in particular:

- National and whole of city planning and infrastructure management including scenario analysis, modelling and visualisation of infrastructure systems and human interaction
- Implications and consequences affecting liveability and the infrastructure systems as our cities grow and sustain higher population densities
- Large scale infrastructure investment, including better use of existing infrastructure through new technologies and sequencing of greenfield investments to best meet national priorities
- Enhancing infrastructure resilience to support continuity of service to the community and business
- SMART will seek to provide deeper knowledge and real time actionable intelligence of key interdependencies and connections in the infrastructure system.

SMART also features a Rail Logistics Centre that is committed to integrated multi-modal transport research, and will support the capabilities needed to underpin the next generation of passenger and freight logistics tasks in Australia.

In this submission we take the opportunity to address the following specific terms of reference;

- Considering the use of smart ICT in related fields, such as disaster planning and remediation; and
- Harmonising data formats and creating nationally consistent arrangements for data storage and access.

RECOMMENDATIONS

THE DEVELOPMENT OF OPEN SOURCE SOFTWARE (OSS) PLATFORMS ARE CRUCIAL FOR WIDE SPREAD DISSEMINATION OF REAL-TIME INFORMATION.

In the field of disaster management and remediation we believe a platform, that is open source and can harvest and display data in real-time, would be of invaluable assistance to decision makers and first responders. OSS should be designed for scalability and transferability with respect to the domain of application, the location, and the language of the users. Tools and platforms built for single-use applications are both costly and inefficient.

EMBRACE SOCIAL MEDIA PLATFORMS AS AN AVENUE FOR GATHERING CROWD SOURCING DATA THAT CAN INFORM DECISION MAKERS DURING EMERGING SITUATIONS.

Our research has shown that crowd-sourcing data via social media is an effective means of sourcing information to support decision making during disasters. An open source and open data approach is critical to establishing the trust and transparency needed between citizens and decision makers to make this happen.

ACTIVELY DEVELOP ROBUST SOCIAL MEDIA STRATEGIES AND PROTOCOLS FOR DISASTER RISK MANAGEMENT (DRM).

As the leading social media platform for real-time information sharing, Twitter offers a variety of functional elements that can be more thoroughly leveraged in the DRM sector. These functionalities include account verification, 'retweet validation' of citizen reports, Twitter Cards, programmatic reply functionalities, and the PowerTrack API Connection.

While Twitter has been used effectively to analyse disasters offline after the event, and to source data in an ad hoc manner during disaster events, decision makers and DRM agencies do not need to wait for disasters to afflict their cities before developing these elements to support their existing DRM information ecosystem. Building robust social media strategies as an element of preparedness is key to using social media during response, rescue, and recovery efforts.

ADAPT STANDARD METADATA FORMATS SUCH AS THE COMMON ALERTING PROTOCOL (CAP) FOR STORAGE OF SOCIAL MEDIA-SOURCED DATA.

DRM OSS for social media integration should be built with the aim of complimenting existing institutional frameworks and offer an open API for further integration into DRM information ecosystems; when possible, the storage of social media-sourced data should adapt to standard metadata formats such as the Common Alerting Protocol (CAP).

DRM OSS should be developed incrementally through an iterative, co-research process that involves the widest variety of stakeholders, including DRM agencies, government managers, scientific researchers, industry partners, and citizen-user groups. By developing modular OSS components through co-research, user feedback can be effectively integrated into prototyping and development processes, thereby optimizing user feedback as the driver of better design and functionality.

ADOPT SIMPLE METADATA STANDARDS THAT CAN EFFICIENTLY CAPTURE AND STORE INFORMATION INCLUDING DATASETS, SOFTWARE, PROJECTS AND OTHER RESOURCES.

By keeping standards simple, metadata that describes seemingly unrelated information can be linked to, queried and discovered on a number of different platforms. Well linked metadata opens up information to researchers from different disciplines. This allows these researchers to join in on and add to the conversation as simple standards reduce technical or disciple specific barriers to information. Well linked open datais an essential part of information management. These links form a web of connectedness that promotes consistency across the infrastructure landscape and provides multiple access points to these datasets, collections and resources.

SUBMISSION

CONSIDERING THE USE OF SMART ICT IN RELATED FIELDS, SUCH AS DISASTER PLANNING AND REMEDIATION. PETAJAKARTA.ORG - A CASE STUDY IN USING SOCIAL MEDIA TO RESPOND TO EXTREME WEATHER EVENTS.

Overview

PetaJakarta.org is a research project led by the SMART Infrastructure Facility, University of Wollongong in collaboration with the Jakarta Emergency Management Agency (BPBD DKI Jakarta) and Twitter Inc.

This research project, led by Co-Principal Investigators Dr Etienne Turpin and Dr Tomas Holderness, recently released a White Paper, outlining the results of the Joint Pilot Study which marks the first phase of this innovative research endeavour. A link to White Paper PetaJakarta.org: Assessing the Role of Social Media for Civic Co Management During Monsoon Flooding in Jakarta, Indonesia, can be found here http://petajakarta.org/banjir/en/research/#responsive

This research demonstrates the work SMART is undertaking in regard to disaster management and response times in Jakarta and shows that crowd-sourced data can be a valuable tool to support disaster risk management for governments and citizens.

This joint pilot study was conducted in Jakarta during the 2014/2015 monsoon season in collaboration with the BPBD DKI Jakarta, and Twitter Inc., and is a world-first collaboration between industry, academia and government. It is a web-based platform which runs on custom built open source software, called CogniCity, which turns the geo-tagged Tweets by Jakarta's citizens into a real time flood map.

The platform allows citizens to share flood information with social media peers while simultaneously providing BPBD DKI Jakarta with data to support decision making for disaster response.

Furthermore, the information shared via Twitter is visualised in real-time with an online map so citizens can warn each other about flood-affected areas and can safely navigate around the city. Information collected by the platform complements existing disaster response systems and helps BPBD DKI Jakarta respond faster to flood situations, which frequently occur across the CBD and outer city regions during the monsoon.

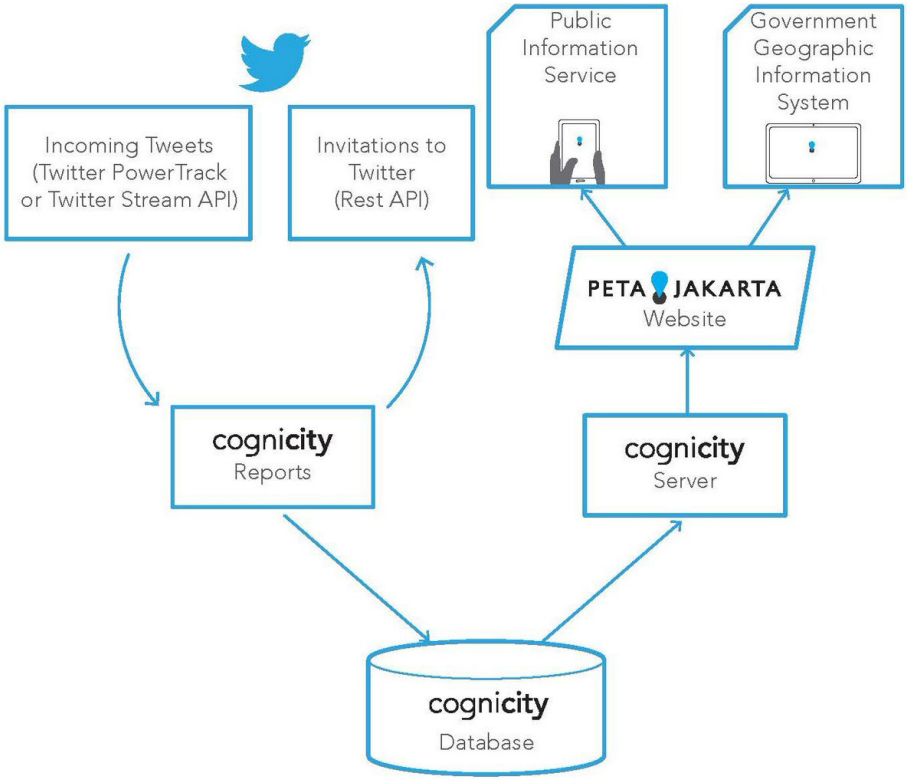

Fig. 01. CogniCity System Architecture

Since the project's official launch by Jakarta Governor Basuki Tjahaja Purnama in December 2014, thousands of people have reported flood problems to PetaJakarta.org via their mobile devices. At peak times, PetaJakarta.org handled more than 3,000 users per hour.

With the support of Twitter, two-way communication protocols to help solicit flood reporting from the public were also trialled; more than 89,000 programmatic invitations have been disseminated to Twitter users in Jakarta, resulting in more than 2.2 million Twitter impressions.

This data was used by BPBD DKI Jakarta to cross-validate formal reports of flooding from traditional data sources, supporting the creation of information for flood assessment, response

and management in real-time. This integration of information provided relevant, timely information to decision makers and assisted in response coordination during extreme weather events.

Why Jakarta?

The pilot project for CogniCity was located in Jakarta for several key reasons. Jakarta, and the surrounding conurbation of Jabodatabek, has the highest rate of urbanization in the world and comprises the second-largest contiguous settlement on earth. With a greater metropolitan area hosting 13 rivers, 1100 kilometers of canals, seasonal monsoon flood events, and over 28 million residents, Jakarta is a key case study for the development of improved risk management through new tools and open source software.

Open source software solutions are critical in Jakarta because cost-prohibitive private products remain unrealistic as solutions under current budget constraints. In Jakarta, risk information and coordination through open data protocols is critical to support decision-making about disaster response, emergency planning, and community resilience. Furthermore, rich suites of open and accessible geospatial risk data generate activity in NGOs and the private sector, especially for longer term disaster risk management planning tools, such as InaSAFE (developed by the Australia-Indonesia Facility for Disaster Reduction), and economic calculators such as JakSAFE (developed by the World Bank).

The development of social media and application-driven data collection via mobile devices allows for unprecedented data collection capacities; in order to be effective, these technologies require coordination through robust, enterprise-grade open source software.

BPBD DKI Jakarta is regularly faced with the difficult challenge of anticipating and responding to floods hazards and related extreme weather events in Jakarta. As a research partner, the organization allowed PetaJakarta.org invaluable access to their operations through a process of co-research and institutional ethnography. By carefully studying the operational procedures, concerns, and ambitions of the Agency, the project developed tools that could be integrated effectively into the existing structure of the organization and its various data flowlines, and transferred incrementally over the course of the research collaboration, thereby ensuring an efficient strategy for both

development and implementation.

Jakarta also has an extremely high proportion of mobile phone users who have embraced Twitter as their preferred social media platform. Because the PetaJakarta.org project relies on densely populated urban environments with high proportions of social media users, Jakarta provides an especially compelling environment for understanding the value of social media in a disaster risk management context.

In addition to BPBD DKI Jakarta using the system as a data source for operational response to flooding, the SmartCity Jakarta project used the PetaJakarta.org open API to access reports of flooding. The Smart City project was developed by Jakarta government's Office of Communication and Information Services (Dinas Komunikasi dan Informatika Provinsi DKI Jakarta) as a data-hub for the Jakarta government and citizens alike, including data on traffic, flooding and civic infrastructure.

SmartCity Jakarta is an example of public sector innovation wherein the government aims to provide better service delivery through digital applications; making this platform open, sharing the data collected, and fostering further innovations through an open API can all support additional improvements to service delivery.

Why Australia?

Given more than half of the world's population lives in the Asian region, and are experiencing high rates of urbanization, we need to work with our neighbours on developing smart solutions for mega cities.

Smart cities are mostly about smart people and Smart ICT is more than high tech and big data, especially in urban or disaster recovery contexts. Crowdsourcing relying on volunteering citizens - 'people as sensors' - and low/medium tech can be a powerful means of communication and coordination.

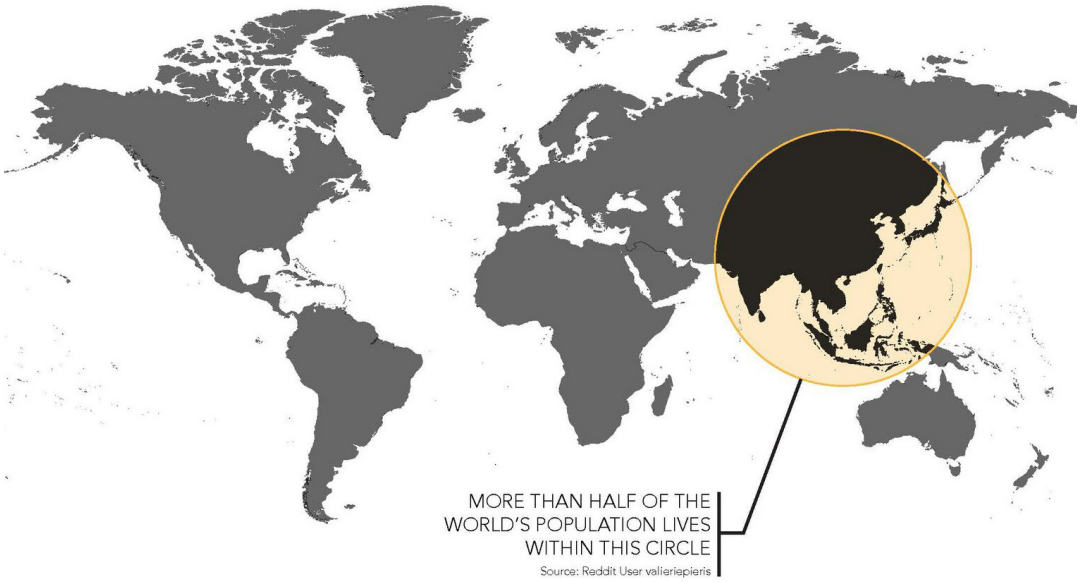

Fig. 02. Map showing Asia's population.

Why Twitter?

Twitter was selected as the ideal social media platform for the Joint Pilot Study because Jakarta has one of the highest concentrations of Twitter users in the world. Additionally, a Twitter #DataGrant gave the SMART research team unprecedented access to a suite of historical data about flooding in Jakarta, thus allowing the system to be developed and calibrated with large data sets, to ensure its operative functionality during the monsoon season in 2014-2015. By working in close collaboration with Twitter, the project benefitted from the advice and mentorship of a group of media experts, engineers, and advocates who were all critical to the overall success of the project.

This technology was assisted by the Government of Jakarta's embrace of social media as a means to communicate with residents; in fact, the Jakarta Emergency Management Agency has a strong mandate to improve communication with the public through social media.

Fig. 03. PetaJakarta Launch tweet sent by user @basuki_btp (Governor of Jakarta) to user @jokowi_do2 (President of Indonesia).

Further the open data and open source nature of PetaJakarta.org enabled BPBD to share information across government and agencies, and feed into other initiatives such as the Smart City Jakarta project.

Why PetaJakarta?

PetaJakarta.org enabled citizens of Jakarta to make independent decisions on safety and navigation in response to the flood in real-time, thereby helping increase the resilience of the city's residents to flooding and its attendant difficulties.

Critically, this outcome was achieved using the same data and map that was used by the government. Designing the platform to meet the needs of citizen-users and government agencies enabled and promoted civic co-management as a strategy for climate adaptation.

PetaJakarta.org has demonstrated social media's valuable niche within the disaster risk management information ecosystem, as an operational tool capable of providing decision support at the various spatial and temporal scales required by the different actors within city. It offered an innovative and inexpensive method for the crowdsourcing of time-critical situational information in disaster scenarios.

Fig. 04. An Example User Interaction with CogniCity; (1) unconfirmed report, (2) programmatic invitation from PetaJakarta.org, (3) confirmed report in reply including photo, (4) programmatic thank you message from PetaJakarta.org.

Outcomes

The Joint Pilot Study for the PetaJakarta.org project was operationally active from December 2014 to March 2015. During this time, the project enabled Jakarta's citizens to report the locations of flood events using the social media network Twitter, thereby contributing to a web-based, publicly accessible, real-time map of flood conditions at PetaJakarta.org.

This data was used by BPBD DKI Jakarta to cross-validate formal reports of flooding from traditional data sources, supporting the creation of information for flood assessment, response, and management in real-time.

The findings of the PetaJakarta.org Joint Pilot Study offer scientific evidence that prove the value and utility of social media as a mega-city methodology for crowd-sourcing relevant situational information to support decision-making and response coordination during extreme weather events or disaster management.

One of the challenges often facing decision makers using citizen reports is the ability to verify submitted information, particularly those harvested from social media. In the case of BPBD Jakarta, flood verification was previously undertaken manually, a time consuming and labour intensive task.

In disaster situations, where big crowd-sourcing of reports is the primary source of information, whether it is flooding or bushfire, the classification of reports as verified or not is of critical importance to ensure that decision makers and the public are only interpreting data which is directly relevant to the situation.

The large volume of data captured during the system testing, and corroborated by the Twitter #DataGrant, necessitated that PetaJakarta.org develop a new solution to collect verified reports of flooding directly from Jakarta's citizens, reducing the requirements for manual verification or intensive computation analysis through machine-learning methods. Only in this way was it possible to provide a situational overview for decision makers and the public in a real-time manner.

Through the PetaJakarta.org platform, the OSS CogniCity used the social media platform Twitter to

engage citizens and form a network of ground truth reports about the flood condition. In this 'people as sensors' paradigm, Twitter users within the system parameters (e.g. geographic location, discussing flooding) were invited to confirm whether flooding was taking place at their location. Users could then see the results of their contributions, as well as those of other citizens, visualised on a map linked to their network (Twitter), in the public domain and in real-time.

The map used by both the public and government agencies thus created a two-way communication interface between users, PetaJakarta.org, and the government. Importantly, users were offered the opportunity to report by text or media in an unrestricted manner (within the confines of Twitter platform) on the given situation.

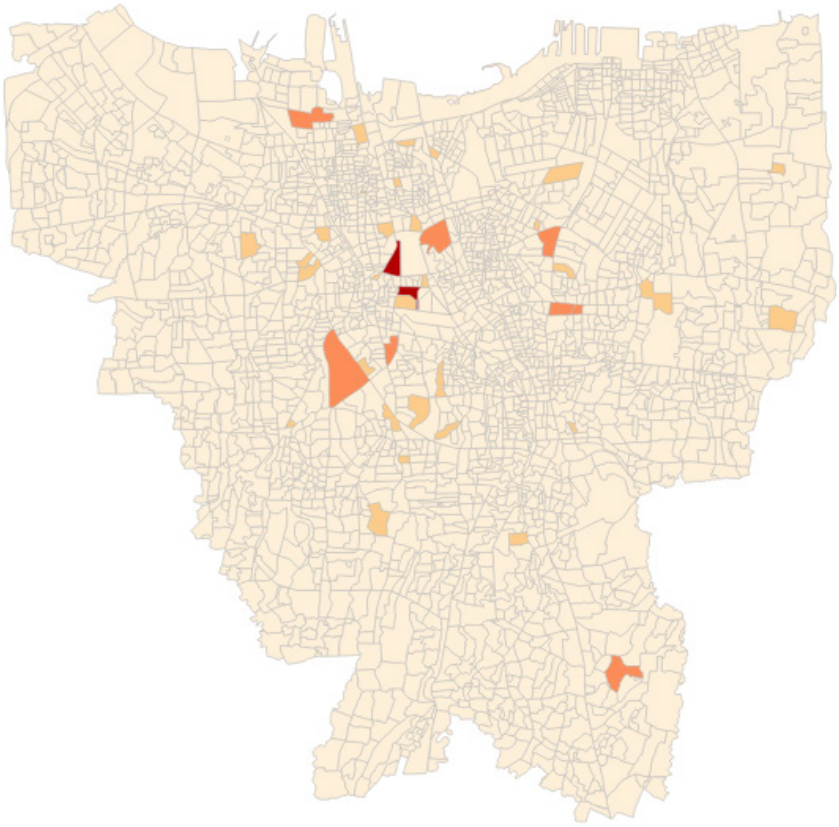

Fig. 05. The PetaJakarta.org map showing areas where flooding is occuring in real time

In this manner, PetaJakarta.org formed a self-evolving network of reporters, who through their engagement with the project, the platform, and each other, shared information and refined their reports during the course of monsoon season.

PetaJakarta.org became one of the first projects in the world to programmatically invite users to participate in a crowd-sourcing effort using Twitter. Whilst existing infrastructure exists within the Twitter platform to disseminate specific messages to a targeted audience (e.g. through promoted tweets), through this world-first collaboration with Twitter, PetaJakarta.org was able to send every user in Jakarta who mentioned the keyword 'flood' or 'banjir' during the monsoon season an invitation to confirm the flood situation via a Tweet. These reports were automatically added to the publically available PetaJakarta.org map.

Over the course of the 2014/2015 monsoon, PetaJakarata.org sent 89,000 invitations to citizens in Jakarta as a call to action to confirm flood conditions, gaining over two million Twitter impressions over the same period as a result.

A critical element of the PetaJakarta.org Joint Pilot Study was the sustained engagement with the Jakarta Emergency Management Agency (BPBD DKI Jakarta), who not only visited SMART at the University of Wollongong for an intensive planning workshop, but also hosted two groups of University of Wollongong Student Flood Support Teams (in November 2014 and January 2015, respectively) funded by the Australian Government's New Columbo Plan.

The collaboration with BPBD DKI Jakarta allowed the PetaJakarta.org platform to be developed in response to institutional goals, operational requirements, and the existing flowline for data management and information dissemination, thus avoiding an intrusive technological imposition and alleviating the difficulties of adopting and integrating external platforms within an already complex information ecosystem.

We believe PetaJakarta.org is a promising evolution within the disaster risk management information ecosystem and a new way for government agencies to engage with real-time emergency situations.

CogniCity, which is capable of processing 250 tweets per second, provides a new tool for civic co-management in disaster affected areas, and is the only fully developed and tested open-source platform that uses a geosocial intelligence approach to urban data collection, analysis and sharing.

SMART DATA MANAGEMENT - HARMONISING DATA FORMATS AND CREATING NATIONALLY CONSISTENT ARRANGEMENTS FOR DATA STORAGE AND ACCESS.

The SMART Infrastructure Facility with the collaboration of the CSIRO has developed a SMART metadata and data management system. The metadata system is built on the open source metadata software GeoNetwork and has been developed specifically to cater to the strengths of the SMART Infrastructure Facility.

The SMART Metadata System and its associated guidelines provide a central e-research platform where infrastructure planners, designers and researchers can access knowledge about infrastructure data from various sources.

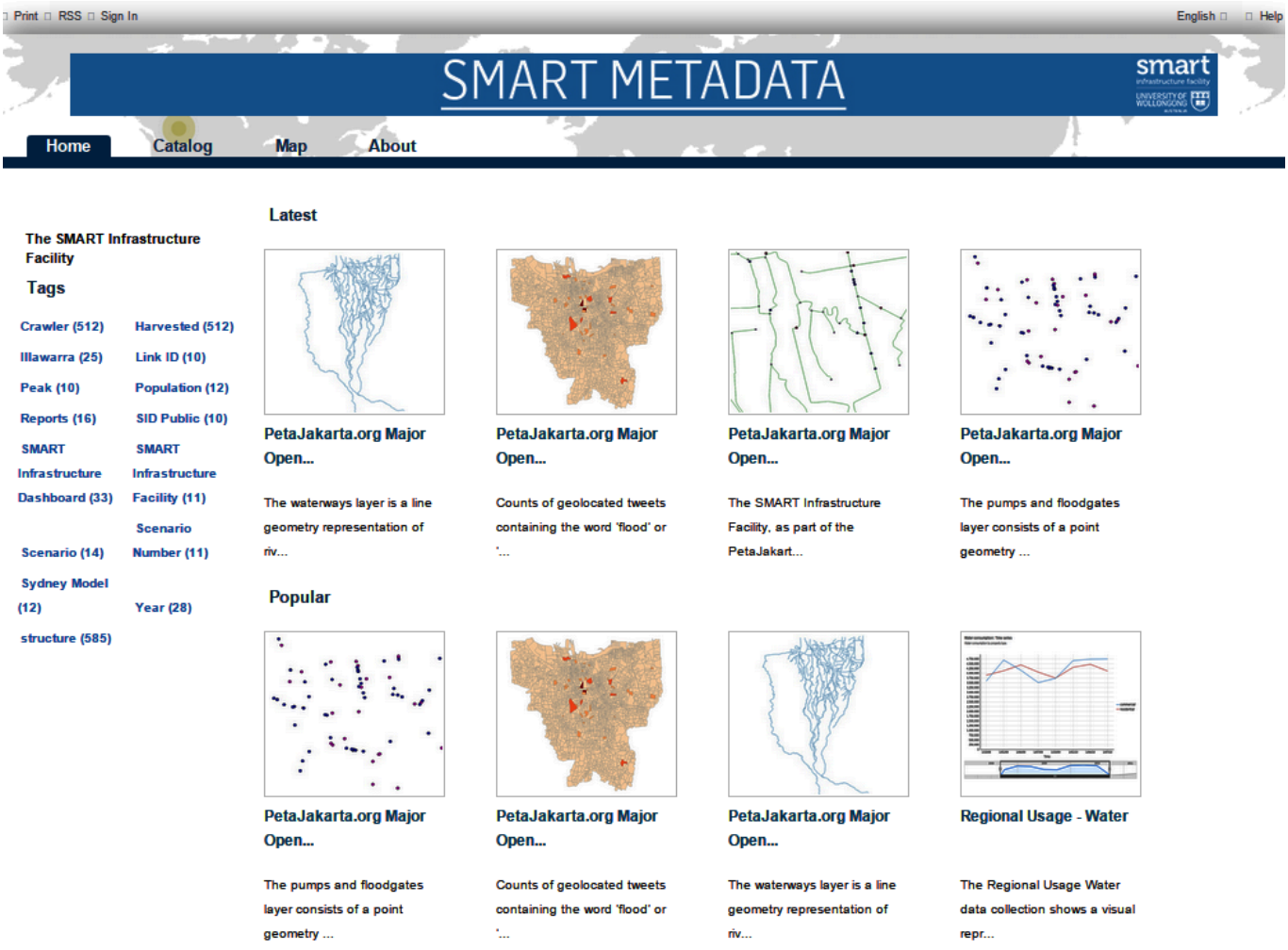

Fig. 06. The SMART Metadata System landing page.

Standardising Data

The system has been developed to catalogue datasets from disparate data providers, research outputs, Commercial and Academic research projects and is designed to be flexible enough so that any type of infrastructure information can be catalogued. SMART uses a subset of the ANZLIC Metadata Profile: AS/NZS ISO 19115:2005, these standards allow our stakeholders to find and use this information and data. This information had previously been unmanaged, hidden and unused.

The SMART metadata system has been configured to harvest research outputs from simulations and models from a number of different research projects. The ability to automate, harmonise and standardise research in this way is an example of how Academia can innovate and contribute within the ITC space.

The data climate of the SMART Infrastructure Facility has been harmonised so that Data Inputs, Outputs, Simulations, Modelling and Teaching have all been made consistent; this consistency facilitates good research outcomes for the Facility and the University.

Linked Data

Datasets published in the SMART Metadata System are linked across the web to provide greater exposure to these data sets. Linked Data lowers barriers to information and enables our information to be more useful, more connected and easily queried.

Published Infrastructure Datasets, like the *Twitter activity related to flooding in Jakarta, Indonesia* collection, link to a number of internal and external catalogues;

- UOW Research Online
- Research Data Australia
- National Library of Australia

These links form a web of connectedness that promotes consistency across the Infrastructure landscape and provides multiple access points to these datasets, collections and resources.

Petajakarta.org case study

As outlined, SMART has recently completed an ANDS funded Major Open Data collections project, PetaJakarta.org. The Major Open Data Collection makes available the archive of data collected by PetaJakarta.org during the 2014/2015 monsoon season. SMART is committed to the dissemination and promotion of its research findings as widely as possible.

These collections drive this agenda by being the first institutional example of data collections being made open, available and downloadable at the Institution level.

PetaJakarta.org is a cross-institution initiative supported by the University of Wollongong's Global Challenges program, SMART, the Engineering and Information Sciences Faculty, the Library and the Vice-Chancellors Unit.

The open distribution of this data will be invaluable to researchers in Australia, Indonesia and worldwide. Such data will demonstrate the potential of social media networks, through a 'people as sensors' paradigm, to inform decision makers on the response of megacity infrastructure to extreme events.

These data collections will form an in-valuable suite of open, well described and richly connected data relating infrastructure and social media use in the context of flooding in Jakarta. The data are all connected by their relevance to flooding, spatial location and temporal extents.

This data represents a unique insight to researchers seeking to understand the response of mega-city infrastructure to flooding in a developing nations context, where social media data offers a potential alternative to limited pre-existing formal flood data networks.