Open Research Online

Mining a MOOC to examine international views of the "Smart City"

Conference or Workshop Item

oro.open.ac.uk

Mining a MOOC to examine international views of the "Smart City"

Daniel Gooch, Lorraine Hudson, Matthew Barker, Annika Wolff, Marian Petre
School of Computing and Communications
The Open University
Milton Keynes, UK
{daniel.gooch, l.e.hudson, matthew.barker, annika.wolff, m.petre}@open.ac.uk

Abstract—**Increasing numbers of cities are focussed on using technology to become "Smart". Many of these Smart City programmes are starting to go beyond a technological focus to also explore the value of a more inclusive approach that values the input of citizens. However, the insights gained from working with citizens are typically focused around a single town or city. In this paper we explore whether it is possible to understand people's opinions and views on the Smart City topics of Open Data, privacy and leadership by examining comments left on a Smart City MOOC that has been delivered internationally. In doing so we start to explore whether MOOCs can provide a lens for examining views on different facets of the Smart City agenda from a global audience, albeit limited to the demographic of the typical MOOC user.**

Keywords— *Smart Cities; Citizen Innovation; Digital Civics, MOOC*

I. INTRODUCTION

While every city has certain unique issues, many face the same broad challenges such as rapid urbanisation, climate change and increasing pressure on city services such as transport and healthcare [6]. Smart Cities are one approach to addressing these issues. While there is no consensus of what a "Smart City" is [24], one of the broadest definitions is provided by Caragliu et al. [8], who argue that a city is Smart when "*investments in human and social capital and traditional (transport) and modern (ICT) communication infrastructure fuel sustainable economic growth and a high quality of life, with a wise management of natural resources, through participatory governance*". Such a definition highlights the shift from a purely technology-driven approach to Smart Cities [9] to a people-centric view in which citizens are seen as collaborators rather than users, as innovators rather than consumers [26].

The Smart City community is currently exploring how to tackle key issues around governance models [25], privacy [20], open data [16], standardisation [7], financing [12] and the role of businesses and other stakeholders [24]. Currently each Smart City project addresses these concerns in a particular way, given the context and constraints in which they are operating. While the HCI field has engaged with civic leaders to understand how these Smart City issues are dealt with in practice [5] [23], we also have an obligation to explore the views and opinions of the general public.

Typically citizens have been involved in Smart Cities through specific initiatives that deal with pragmatic concerns: from using open data to produce services [1] to being data collectors [29], from reporting problems that need to be fixed [4] to being involved in hackathons [14], from being a stakeholder in a Living Lab [11] to crowdsourcing ideas to change local communities [26]. While each of these initiatives is of interest, they all share the same focus on practicalities rather than exploring the underlying issues the concept of Smart Cities involves. We want to take a broader perspective and understand what people's views and critiques are with regard to the entire Smart City agenda beyond a specific initiative in a fixed location.

We have thus investigated whether the comments left on a Smart City MOOC can be analysed to understand the public's view from a range of different countries. Smart Cities is a topic of global interest and, as such, a MOOC is an ideal mechanism for gathering a wide range of citizen perspectives from international learners in different locales, albeit at the cost of not focussing on a single city, losing detail and context. Furthermore, each of the correspondents will have learnt about different approaches to Smart Cities, allowing them to offer their own informed opinions. In this paper we present an initial analysis of MOOC comments and argue that this initial work indicated this is an area of research worth pursuing.

II. THE SMART CITY MOOC

Through our relationship with a MOOC platform organisation, we had access to the data and participants of a MOOC which provides a short introductory course on key topics around Smart Cities, presented in English. The design of the MOOC follows the linear 'X-MOOC' style of presentation, with knowledge transmitted from instructor to student, but with a strong pedagogical focus on social learning through commenting, discussion and provoking conversations [17].

The MOOC is designed to be studied over 6 weeks taking 18 hours of learning time. Each week is composed of distinct teaching elements, called steps, which include articles, videos, audio clips, activities or discussion steps. The 6 weeks are titled:

1. Introduction to smart cities

2. Smart citizens

3. Infrastructure, technology and data

4. Enterprise and innovation

5. Leadership and strategy

6. Measurement and learning

Due to their online presence, MOOCs can attract a large number of international learners. If they have left comments, this would allow us to examine views not only within a single city but across multiple countries. One of the reasons we chose to analyse the selected MOOC was that it uses case studies from around the world, including Milton Keynes, New York, Dubai, Reykjavik, Ajmer, Rio de Janeiro and Songdo, building diversity into the course content. This may have made it more relevant and meaningful to global learners. For more details on the MOOC, see [15].

We are reporting on comments left by learners from three different presentations of the MOOC. The first presentation ran from September 28th 2015 and 8,005 people were enrolled. The second presentation ran from January 18th 2016 and 6,438 people were enrolled. The third presentation ran from April 11th 2016 and 4,972 people were enrolled.

Through data collected by the MOOC platform we can examine the nature of the learners enrolled on the course. Table 1 shows the number of Joiners (people enrolled on the course), Learners (joiners who start to use the course), Fully participating learners (completed at least 50% of the available steps) and Social learners (posted at least one comment) for each presentation.

TABLE I.   NUMBER OF DIFFERENT LEARNERS FOR EACH PRESENTATION OF THE MOOC

Type of Learner	Presentation		
	One	Two	Three
Joiners	8005	6438	4927
Leaners	3692	3070	2598
Fully participating learners	626	528	410
Social learners	727	549	475

It is common for MOOCs to lose a high percentage of their initial joiners as many people's investment in a free course, which demands time and effort, is not sufficient for them to start their studies [3]. The average number of fully participating learners across the 3 presentations was 17%.

III. METHODOLOGY

We chose to focus on three key topics within the Smart City MOOC - Open Data, Privacy and Leadership (presented in weeks four, three and five respectively). These three topics were selected for a number of reasons. Each topic is an area of ongoing research, allowing us to compare the views expressed by learners in the MOOC against views gathered through other methodologies. Furthermore, each topic is being actively researched as there is no agreed upon approach to open data, privacy or leadership within Smart Cities, suggesting that different citizens from different countries may express different views, particularly based on cultural norms. Finally, the topics are of direct relevance to citizens' lived experiences and are areas of the Smart City agenda in which citizens' views matter – a city cannot institute a privacy policy that the majority of its citizens are against.

Within these topics we selected comments from the designated discussion steps, where learners are encouraged to post their thoughts and opinions about the topic. Table 2 shows the breakdown of these social learners for each discussion. Each learner may comment on more than one topic or in multiple presentations but we do not have the data to identify such instances.

TABLE II.   NUMBER OF SOCIAL LEARNERS FOR EACH TOPIC FOR EACH PRESENTATION OF THE MOOC

Social Learners	Presentation		
	One	Two	Three
Open Data	166	136	113
Privacy	188	166	129
Leadership	126	97	80

A. Demographics of the MOOC Learners

The MOOC platform gathers demographic information through an optional demographic survey that started after the first presentation of the MOOC. However, this is linked to the learner rather than to a course, meaning that we have demographic information from all three presentations. Across the three topics and three presentations, we have demographic information on 67 social learners.

The surveys highlight the range of people commenting on the MOOC. Around 66% were male (44 respondents), 33% female (22 respondents), with 1 stating "other". Table 3 shows the range of ages of the respondents.

TABLE III.   NUMBER OF LEARNERS FOR EACH AGE RANGE

Age Range	Number of Respondents
<18	1
26-35	14
36-45	8
46-55	16
56-65	11
>65	15
Unknown	2

Our respondents were typically well educated, with 88% holding a university degree. The majority were employed (57%, 38 respondents) or retired (27%, 18 respondents) with the remainder either students (3 respondents), looking for work (3), not working (2) or unknown (3).

Perhaps most interesting is the range of countries represented in our sample. Our respondents came from 23 different countries with at least one learner in every continent. While nearly 50% of respondents came from the UK (49%, 33 respondents), Spain (4 respondents), France (3), Nigeria (2), India (2) and Mexico (2) were also represented.

Our complete data set comprises 619 comments left by 281 unique learners. The average comment was 60 words long.

B. Data Analysis

We used an inductive open coding approach to examine the meanings embedded within the comments made by the MOOC learners in the three discussion threads [10], informed by our interest in the top-level concepts of "Open Data", "Privacy" and "Leadership". The comments were subjected to a line-by-line analysis in which concepts were identified and labelled within the data. These codes were subsequently combined into emerging themes. No codes or themes existed prior to the analysis; they were created through constant comparison of the data and the application of labels to the text. Each of the themes was checked for inter-rater agreement with a second coder, highlighting almost perfect agreement (all Kappas > 0.903, at p<0.001). All disagreements between coders was adjusted through conversation between the two independent coders.

These themes help us examine how useful the comments are in examining Smart City concerns. During the analysis and interpretation of the comments, all of the authors took part in extensive discussions to ensure that we were being led mainly by the data, informed by an understanding of the context and content of the MOOC. As an individual comment reflected multiple views, we refer to the number of comments that reflected a theme rather than the number of learners. Under our agreement with the MOOC platform, permission for quoting comments have been obtained from the commenter and are directly attributed.

IV. RESULTS AND ANALYSIS

A. Open Data

The learners engaged with the topic of Open Data in week 4 of the MOOC. Learners were asked to "*take a look on the NYC Open Data platform and identify any datasets relevant to the city problem you chose in Week 2. This will be for New York rather than for your own city but it will give you an idea of what data may be made available. Is there something of interest to you here? Do you find the platform easy to navigate?*"

218 comments were left by 150 learners. The main theme of the comments was how challenging the platform was to use – 34 comments (from at least 3 countries across 2 continents) described how hard it was to navigate the platform and the challenge involved in finding and extracting the data they want to use: "*It's quite overwhelming. I've done a few searches and exports but am struggling to find datasets that would help me with one of the issues I've identified*" [Emma Doyle]. In comparison, only 16 comments (from at least 3 countries

across 2 continents) argued that the platform was easy to navigate.

The other main view expressed was the challenges involved in actually using the data. 21 comments discussed the lack of data regarding specific issues; 19 noted that raw data was difficult to work with; 12 argued that graphical representations of the data would be easier to interpret and 13 discussed the limitations of the data formats on the platform: "*There is a dataset about air quality in the city. I think there are too many sources of information opened in different formats. If you want to use several sources you have to build something by your own*". Similarly, 9 comments noted the importance of context in understanding what the raw data means; 9 noted that much of the data appeared out of date and 8 argued that the value of the data really emerges when you have the ability to merge multiple data sets, something the platform doesn't currently offer: "*I found the tree data and green spaces useful. My city Liverpool [UK] has an open data portal but as far as I can see the data input appears to have stopped about 2014. It was poorly supported*" [Pauline Fairclough].

The views expressed by the MOOC learners correspond with the current research into open data systems which argues that current systems are challenging to use for citizens and app developers alike [16] [19].

B. Privacy

Privacy was the second discussion point, discussed in week 3 of the MOOC. Learners were asked: "*Are you concerned about increasing numbers of city sensors and their impact on your privacy? Or perhaps you think the benefits outweigh the risks? Do you know of any good examples of cities addressing the issues of privacy and security?*".

This topic sharply divided opinion. 250 comments were left by 174 learners. 79 of the comments (from at least 5 countries across 2 continents) expressed reservations about their privacy: "*I am a little bit scared about privacy because I think that companies and governments could use this data for their interests*" [Montserrat Sans Boza] while 78 comments (from at least 3 countries in 1 continent) were not concerned: "*I am not concerned with the number of city sensors as long as they are there for a defined purpose, used by appropriate and defined city authorities who are audited, checked and held accountable. Enhanced safety, security and incident response being the benefits, although there are no guarantees*" [Graham Stephens].

For those worried about privacy, 24 comments noted that technical systems are not infallible while 37 comments (from at least 4 countries in 2 continents) were focussed on overreaching powers by governmental or commercial organisations: "*the idea that I can be monitored, located and commercialised makes me very cautious to the whole idea of smart cities. I want to help improve my city but not to the cost of my own privacy*" [John Memtsas].

In contrast, 54 comments (from at least 6 countries across 4 continents) were not worried about privacy and had made a judgement that the benefit the systems bring outweigh any of the costs: "*well, I agree with benefits outweigh the risks… 'El*

que nada hace, nada teme' [means] 'It who does nothing, fears nothing'" [Jose Urrea].

Two additional topics were discussed. The first focussed around the different types of information which may be collected and the more stringent privacy controls that should be available to deeply personal information (31 comments from at least 6 countries in 4 continents). The second was more concerned with auditing, which included explaining how the data was collected, how it was audited, how one gains access to it, how one can check the information stored about oneself and the level of governmental oversight provided: "*good practise would be that everyone collecting data should have a clear statement about what they collect, how personalized it is and who it is available to. This should be subject to audit and challenge*" [Tim Elliott].

The views expressed by our learners are closely related to those found in previous privacy research both within and beyond research in Smart Cities [2], [22], [27], highlighting the complexity of decision making when it comes to the disclosure of "private" information.

C. Who should lead our smart cities?

The final discussion point, leadership, was discussed in week 5 of the MOOC. Learners were asked: "*You've already heard that leadership has been a factor in establishing successful smart city partnerships. Can you think of positive or negative examples of city leadership? Thriving cities exhibit creativity and innovation. Would you say that leadership is a vital driver of these factors, or a hindrance? Who do you think should lead our smart cities – politicians, city government managers, businesses or community leaders? Do you know of an inspiring example of smart city leadership or any interesting partnership approaches?*"

151 comments were left by 129 learners. The overwhelming view was that leadership had to be independent and responsive to the needs of the city: "*the city needs a clear vision and somebody to guide all efforts towards reaching it. But of course the leader should represent others, most of all the citizens, not just himself*" [Daniela Miscov]. 64 comments (from at least 6 countries in 4 continents) argued that for such important decisions, politicians were suitable only if they were responsive to citizens' and the city's needs. However, the learners also felt that only city politicians had the democratic mandate, and the financial control, to make their city smart: "*In terms of who should be the leaders of smart cities, if not the person at the head of the city than what is the purpose of the mayor?*" [Amelia Nicola] with a further 36 (from at least 7 countries in 4 continents) arguing that political will was not enough, that creating a Smart City required input from a range of different stakeholders, and that partnerships were the only viable option "*to become a Smart City is a medium/long term process, so we must prevent this process from political changes. The SC process should be shared among a huge majority of agents, citizens, etc.*" [Carlos Ochoa].

However, a small proportion of the comments reflected alternative views. 15 comments argued for governance by experts, a technocracy. 11 commenters argued that only citizens and community leaders would have the knowledge needed to shape Smart Cities in the right direction: "*Community Leaders, as they are closer to their communities, that way we can get the communities to be heard and thus get appropriate solutions for each community and the city as a whole*" [Sergio Herrrera].

This was due to one of the underlying scepticisms amongst our learners, questioning organisations motives. 12 commenters argued that the financial backing of politicians, their lack of transparency, accountability link to ordinary citizens and the time scale involved meant that politicians (who want to get re-elected) were not suitable to lead on Smart City projects. 7 commenters stated specifically they were anti-business due to the need for the corrupting influence of a profit margin, while 6 commenters thought that Smart Cities would arrive only when businesses could make a profit.

The views expressed by our learners reflect the unanswered question regarding who should lead Smart Cities with politicians, city administrators, businesses and community leaders all suitable candidates [25].

V. CONCLUSION AND FURTHER WORK

The main goal of this paper was to explore whether the comments left in discussion threads on MOOCs could be mined to gather meaningful opinions from the public on unsolved issues around Smart Cities. The analysis we have performed has demonstrated this with each examined topic – open data, privacy and leadership – providing corroborative data to support existing views expressed in the research literature. This highlights that the views expressed are not unique to single locales but that there are commonalities internationally. This is particularly useful when considering the transferability and standardization of best practice.

It is important to highlight how analysing MOOC comments is distinguishable from other methodologies, particularly how it differs from simply sending out a questionnaire. The main difference is the nature of the respondents. MOOCs attract an international audience who are typically highly educated. They are also, through the process of learning through the MOOC, knowledgeable about the topic being discussed. While these features are not unobtainable through other methodologies, they are inherent features of the population who study MOOCs, making them much more obtainable. However, this is also the key limitation of our work. While MOOCs provide access to a certain demographic, this also leads to a skewed sample.

This work very much remains a work in progress with a number of avenues for further investigation. While our analysis indicates that MOOCs could be used by researchers as a source of information on public views, we haven't been able to present a more detailed breakdown of differences across international contexts, nor have we investigated how our learners may have been influenced by the material presented within the MOOC.

A more in-depth analysis of this data is needed to provide more comprehensive evidence of the potential of the approach we have taken. Such an analysis would also assist us in determining whether MOOCs can be used in this fashion as

they are currently designed or whether there need to be adjustments, either to the MOOCs or the way learners use MOOCs to study, before these investigations bear fruit.

Our insights lead us to conclude that a MOOC could provide a valuable lens for assessing people's views on different facets of the Smart City agenda or other topics where people have a wide range of views.

REFERENCES

[1] Tuba Bakıcı, Esteve Almirall, and Jonathan Wareham, 2013. A smart city initiative: the case of Barcelona. *Journal of the Knowledge Economy*, *4*(2), pp. 135-148.

[2] Louise Barkhuus. 2012. The mismeasurement of privacy: using contextual integrity to reconsider privacy in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '12). ACM, New York, NY, USA, 367-376. DOI=http://dx.doi.org/10.1145/2207676.2207727

[3] S. Bayne, and J. Ross, "The pedagogy of the massive open online course: the UK view". The Higher Education Academy, 2014.

[4] Karim Benouaret, Raman Valliyur-Ramalingam, and François Charoy, 2013. CrowdSC: Building Smart Cities with Large-Scale Citizen Participation. *Internet Computing, IEEE*, *17*(6), 57-63.

[5] Kirsten Boehner and Carl DiSalvo. 2016. Data, Design and Civics: An Exploratory Study of Civic Tech. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16). ACM, New York, NY, USA, 2970-2981. DOI=http://dx.doi.org/10.1145/2858036.2858326

[6] Erik Braun, Leo van den Berg and Jan van der Meer, eds. National Policy Responses to Urban Challenges in Europe. Ashgate Publishing, Ltd., 2012. Harvard.

[7] British Standards Institution (BSI) (2014) *PAS 181:2014 Smart Cities Framework: Guide to Establishing Strategies for Smart Cities and Communities*. http://www.bsigroup.com/en-GB/smart-cities/Smart-Cities-Standards-and-Publication/PAS-181-smart-cities-framework/ (accessed 7 September 2015).

[8] Andrea Caragliu, Chiara Del Bo, and Peter Nijkamp, 2011. Smart cities in Europe. *Journal of urban technology*, *18*(2), 65-82.

[9] Hafedh Chourabi, Taewoo Nam, Shawn Walker, J. Ramon Gil-Garcia, Sehl Mellouli, Karine Nahon, Theresa A. Pardo, Hans Jochen Scholl, 2012. Understanding smart cities: An integrative framework. In *System Science (HICSS), 2012 45th Hawaii International Conference on* (pp. 2289-2297). IEEE.

[10] Juliet Corbin and Anselm Straus. 2008. *Basics of qualitative research*. Sage, London, UK.

[11] ENoLL (2015) ENoLL Website http://www.openlivinglabs.eu/aboutus

[12] European Commission and Smart Cities and Communities (2013) *Smart Cities Stakeholder Platform: Financing Models for Smart Cities*. https://eu-smartcities.eu/sites/all/files/Guideline-%20Financing%20Models%20for%20smart%20cities-january.pdf (accessed 22 September 2015).

[13] Bent Flyvbjerg, 2006. Five misunderstandings about case-study research. *Qualitative inquiry*, *12*(2), 219-245.

[14] Hendrik Hielkema and Patrizia Hongisto, 2013. Developing the Helsinki smart city: the role of competitions for open data applications. *Journal of the Knowledge Economy*, *4*(2), pp. 190-204.

[15] Lorraine Hudson, Gerd Kortuem, Annika Wolff, and Patrina Law, 2016. Smart Cities MOOC: Teaching citizens how to co-create smart cities. In *Proceedings of ICT4S*.

[16] Marijn Janssen, Yannis Charalabidis and Anneke Zuiderwijk, 2012. Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, *29*(4), pp. 258-268.

[17] Thierry Karsenti, 2013. "What the research says." *International Journal of Technologies in Higher Education* 10(2), pp. 23-37.

[18] J. Richard Landis and Gary G. Koch, 1977. The measurement of observer agreement for categorical data. *Biometrics* 33:159-174.

[19] Melissa Lee, Esteve Almirall, and Jonathan Wareham. 2015. Open data and civic apps: first-generation failures, second-generation improvements. *Communications of the ACM*, 59, 1 (December 2015), 82-89. DOI=http://dx.doi.org/10.1145/2756542

[20] Yibin Li, Wenyun Dai, Zhong Ming, and Meikang Qiu, 2016. Privacy protection for preventing data over-collection in smart city. *IEEE Transactions on Computers* 65(5), pp 1339-1350.

[21] Matthew Lombard, Jennifer Snyder-Duch, and Cheryl Campanella Bracken, 2002. Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human communication research*, *28*(4), pp. 587-604.

[22] Antoni Martínez-Ballesté, Pablo A. Pérez-Martínez, and Agusti Solanas, 2013. The pursuit of citizens' privacy: a privacy-aware smart city is possible. *IEEE Communications Magazine* 51(6), pp. 136-141.

[23] Donald McMillan, Arvid Engström, Airi Lampinen, and Barry Brown. 2016. Data and the City. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16). ACM, New York, NY, USA, 2933-2944. DOI=http://dx.doi.org/10.1145/2858036.2858434

[24] Taewoo Nam and Theresa A. Pardo. 2011. Conceptualizing smart city with dimensions of technology, people, and institutions. In *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times* (dg.o '11). ACM, New York, NY, USA, 282-291. DOI=http://dx.doi.org/10.1145/2037556.2037602

[25] Manuel Pedro Rodríguez-Bolívar, 2015. *Transforming city governments for successful smart cities*. Vol. 8. Springer.

[26] Dimitri Schuurman, Bastiaan Baccarne, Lieven De Marez and Peter Mechant, 2012. Smart ideas for smart cities: Investigating crowdsourcing for generating and selecting ideas for ICT innovation in a city context. Journal of theoretical and applied electronic commerce research, 7(3), pp. 49-62.

[27] Stephen Snow, Kenneth Radke, Dhaval Vyas, and Margot Brereton. 2014. Privacy in the new era of visible and sharable energy-use information. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design* (OzCHI '14). ACM, New York, NY, USA, 342-351. DOI=http://dx.doi.org/10.1145/2686612.2686664

[28] Howard EA Tinsley and David J. Weiss, Interrater Reliability and Agreement. *In:* Tinsley, H. E., & Brown, S. D. (Eds.). (2000). *Handbook of applied multivariate statistics and mathematical modeling*. Academic Press.

[29] Andrea Wiggins and Kevin Crowston, 2011. From conservation to crowdsourcing: A typology of citizen science. In *System Sciences (HICSS), 2011 44th Hawaii International Conference on* (pp. 1-10). IEEE