Sensing the City: Designing for Privacy and Trust in the Internet of Things

Dr Caitlin D Cottrill[a], Dr Naomi Jacobs[b1], Dr Milan Markovic[b], Prof Pete Edwards[b]

[a] *Centre for Transport Research, School of Engineering, University of Aberdeen;* [b] *Department of Computing Science, University of Aberdeen*

Abstract: The push for 'smart cities' is seeing increasing interest and investment in public deployments of Internet of Things (IoT) technologies. In many cases, however, these deployments are 'hidden in plain sight', with inadequate attention given to the communication of practices regarding the collection, use and sharing of data. This lack of information may impact negatively upon perceptions of trust by the general public, and lead to diminished engagement and comfort with IoT activities. In this paper, we report the results of a survey undertaken in the United Kingdom in November 2018 designed to gain information on perceptions of trust, risk, and informational desires related to public IoT deployments. Findings are expected to contribute to the development of communication practices regarding IoT deployments in smart city environments.

Keywords — Communication Practices, Internet of Things, Privacy, Smart Cities

[1] Currently Lancaster Institute for the Contemporary Arts, Lancaster University

Sensing the City: Designing for Privacy and Trust in the Internet of Things

Introduction

The British Standards Institute has defined the 'smart city' as one which includes, "effective integration of physical, digital and human systems in the built environment to deliver a sustainable, prosperous and inclusive future for its citizens (British Standards Institution 2014)." Such integration is generally understood to have a significant emphasis on technology to inform physical and human systems, with the use of sensor technologies a key component of this process. Park et al. have stated that, "To prepare the basic infrastructure of a smart city, various sensors, support technologies, and background environments are essential and are being employed in urban areas. Among them, the Internet of Things (IoT) is considered one of the most important aspects for the successful implementation of a smart city (Park, del Pobil and Kwon 2018)." Such integration is, however, not without its detractors, and some key smart city issues have thus far received inadequate attention, including privacy and trust in IoT contexts. These gaps are evident in such recent experiences as the public backlash against Sidewalk Lab's[2] proposed redevelopment of a portion of the Toronto waterfront, which was criticised for being overly opaque regarding practices surrounding data collection, use, and ownership (McLeod 2018).

The IoT is a complex topic to approach from the viewpoint of trust; however, Sicari et al. argue that, "…trust is a fundamental issue since the IoT environment is characterized by different devices which have to process and handle the data in compliance with user needs and rights (Sicari, et al. 2015)." Patel and Patel define IoT systems as "…a concept and a paradigm that considers pervasive presence in the environment of a variety of things/objects that through wireless and wired connections and unique addressing schemes are able to interact with each other and cooperate with other things/objects to

[2] Sidewalk Lab is a subsidiary of Google's parent company, Alphabet.

create new applications/services and reach common goals (Patel and Patel 2016).” Given these characteristics, traditional markers of trust, such as identifiability of involved parties, uncertainty and risk, and faith in honesty and benevolence (Siau and Shen 2003), may be absent. In such cases, the underlying trust that one has towards more generalised actors (such as government or private companies) may serve as a basis for decisions regarding trust under assumed conditions (Gao and Bai 2014).

In the *TrustLens*[3] project, we explore issues of privacy and trust as they relate to the Internet of Things from a viewpoint that considers these more generalised perceptions of trust and familiarity with technology. Through this, we can better our understanding of both the public’s perceptions of risk, and preferences regarding its communication with respect to devices present in public spaces. In such a context, we assume that the public have little actual control over the installation of devices (unlike control they may have in their homes or other private spaces), or knowledge of the ecosystem of data collection, use and exchange. In a 2018 survey conducted in the United Kingdom, we explored these issues as they relate to trust, data privacy, and perceptions of risk with respect to IoT applications, with the expectation of applying our findings to methods that may be used to better encourage trust in the IoT. The survey also asked participants to indicate their informational desires regarding public IoT deployments, to provide a foundational understanding of transparency requirements as a facilitator for trust.

In this paper, we present an overview of the survey findings, with particular attention to the different preferences revealed across categories of participants; namely, those displaying high, medium, or low levels of trust. Through doing so, we aim to introduce more detailed information on the ways in which user preferences regarding trust, willingness to share data, and privacy concerns may be incorporated into the deployment and communication of practices regarding public space IoT deployments, such as those envisaged for smart city development.

[3] https://trustlens.wordpress.com/

Background

Privacy and Trust

The concepts of privacy and trust are generally considered to be multidimensional, with a variety of personal and contextual factors that may influence people's perceptions. In recognition of the variance of privacy preferences demonstrated by consumers, Alan Westin has conducted a series of surveys over time and used these to develop privacy indices, which segment people into clusters of high privacy concern (termed privacy 'Fundamentalists'), moderate privacy concern (or privacy 'Pragmatists') and low privacy concern (or privacy 'Unconcerned') [(Westin et al. 2003) cited in (Kumaraguru and Cranor 2005)]. While this approach has been criticised as being too reductionist (particularly in (Martin and Nissenbaum 2016)), it does introduce a way of taking a clustering approach to assessing people's overall privacy preferences, which may simplify further practices.

In contrast to Westin, Nissenbaum's concept of 'contextual integrity' speaks directly to variations in privacy preferences under changing conditions, particularly in the context of public spaces. She states that, "whether a particular action is determined a violation of privacy is a function of several variables, including the nature of the situation, or context; the nature of the information in relation to that context; the roles of agents receiving information; their relationships to information subjects; on what terms the information is shared by the subject; and the terms of further dissemination (Nissenbaum 2004)." Unfortunately, however, general practices regarding the protection of personal information do not always acknowledge these contexts, driven instead by the static application of privacy policies under a 'notice and consent' approach, which implies "…the individual whose personal data is being processed has—after being informed of the reason, context, and purpose of the processing—given consent (Cate and Mayer-Schönberger 2013)."

A mediating factor that may be introduced is that of trust, as indicated by Joinson et al., who found that "…privacy and trust at a situational level interact such that high trust compensates for low privacy,

and vice versa (Joinson, et al. 2010).” Trust in an actor and their practices may lead to a greater willingness to share information on the part of the consumer (Wirtz and Lwin 2009) (Taddei and Contena 2013); however, such trust may be eroded in the absence of adequate information. In a study regarding the influence of privacy policies on consumer behaviour, for example, (Arcand, et al. 2007) found that the presence of a privacy policy at an online shop may increase consumers perceptions of control. However, for those who read the policy, significantly higher perceptions of control and trust were only seen if an ‘opt-in’ format (in which the user is given more control over access to and management of her data) was used.

Such a finding is concerning, particularly as privacy policies have become the de facto method of communicating data practices with consumers (Kelley, et al. 2010). Such polices, in line with the notice and consent model described above, attempt to provide a standard synthesis of data practices to the consumer; however, such approaches have experienced increasing criticism over time. With respect to online practices, for example, Hetcher argues that, “…privacy policies are not an adequate means to accomplish the task of protecting website visitors from the invasive practices of websites (Hetcher 2000).” Privacy policies have been found to be time consuming to read (McDonald and Cranor 2008); overly complex for the average reader (Jensen and Potts 2004) (Cottrill and Thakuriah 2013); and generally inadequate for conveying data practices to the consumer (Solove 2013) (Pollach 2007). As noted by Pollach, “…these documents are written in a manner that protects companies against privacy lawsuits by integrating privacy legislation that regulates, for example, information gathered from children,… financial data,…and medical records,… or state legislation such as California’s Online Privacy Protection Act. In addition, Internet users have been found not to read online privacy policies because they find them too legalistic and therefore difficult to understand (Pollach 2007).” A compounding concern in an IoT context is that, in addition to their potential inadequacy, access to privacy policies may not be immediately evident given the distributed and public nature of sensors.

Privacy and Trust in the IoT

As noted above, the privacy policy approach described above is likely inadequate for addressing concerns evident in the IoT, particularly when deployed in the public sphere. As described by Peppet, "…four inherent aspects of sensor-based technologies – the compounding effects of what computer scientists call "sensor fusion," the near impossibility of truly de-identifying sensor data, the likelihood that Internet of Things devices will be inherently prone to security flaws, and the difficulty of meaningful consumer consent in this context – create very real discrimination, privacy, security, and consent problems (Peppet 2014)." These considerations are all incompatible with a notice and consent approach to data privacy, as they underscore the implausibility of providing adequate communication of data collection practices (with reference to the collector, the data collected, and the use of this data) to allow for considered consent to be given.

Such concerns are evident in cases such as the Sidewalk Lab example given above, as well as instances such as the Chicago Array of Things (AoT)[4], which faced intensive public scrutiny regarding data collection and privacy practices. A report from an AoT workshop identified some of these considerations, noting the following as key areas of concern:

> "(1) the inability for an individual to easily opt in or out of participation in the AoT; (2) as a general purpose, research oriented, and continuously evolving platform, the AoT presents new challenges for privacy that aren't present in previous sensor or technology deployments that are traditionally more task oriented (i.e. bridge monitoring sensors); and (3) data collected or processed by the AoT could include measurements that are generally considered public (i.e. air quality) and those that are often considered private (i.e. images, albeit AoT is designed without

[4] https://arrayofthings.github.io/

the capacity to store images beyond what is necessary to process them locally on the nodes) (Welch and Catlett 2015)."

These concerns demonstrate the wide variety of privacy considerations evident in IoT systems, and reveal that there is a need for further attention to the subject.

Privacy and the IoT in smart cities

As indicated in the Sidewalk Labs and Array of Things examples described above, smart cities often imply sensed cities, with (Silva, Khan and Han 2018) indicating that, 'The smart city is an application of the IoT (p. 697).' Smart cities worldwide are seeing significant investment, with (Teale 2020) reporting findings from the International Data Corporation (IDC) that spending on smart city initiatives globally will total nearly US$124 billion in 2020. While such investment demonstrates great promise for enhancing smart cities, issues of privacy and trust remain open research questions, as identified in (Braun, et al. 2018), who state that 'Smart cities must ensure individual privacy and security in order to ensure that its citizens will participate. If citizens are reluctant to participate, the core advantages of a smart city will dissolve (p. 499).' With smart city IoT sensors increasingly located in public spaces, developing trust through transparent practices is becoming increasingly complex. Developing appropriate methods of communicating data collection and use processes is key, particularly when collected data may be identifiable or when there is no opportunity to opt-out of data collection in a particular geographic space. However, while extensive research has been undertaken to enhance data privacy preservation and security in IoT systems (see, for example, (Bennati and Pournaras 2018); (Wang, et al. 2014); and (Davies, et al. 2016)), how these practices are communicated to the public has received significantly less attention.

Summary

As the above review demonstrates, privacy and trust are interrelated concepts that may demonstrate a wide variety of preferences depending on both the individual and the context. Current notice and consent practices, however, demonstrated most clearly through the use of privacy policies, do

not adequately address consumer needs with respect to their privacy and trust preferences. IoT systems deployed in public spaces as part of 'smart city' initiatives present a particularly complex case for privacy and trust issues, as they introduce uncertainties aligned with spatial deployment, integration into existing systems, and a lack of transparency regarding data collection and use practices. In the following sections, we describe and analyse a 2018 survey undertaken in the UK that was designed to evaluate consumer perceptions of risk and trust in IoT deployments, as well as determine what informational communications elements are required to engender trust in the IoT.

Methodology

The 20-question survey instrument was developed following an extensive literature review and a number of participatory research activities with the public, undertaken as part of the *TrustLens* project and reported in (Jacobs, et al. 2019). Initial IoT-related privacy and trust concerns identified by citizens during engagement activities were collected and checked against existing research (discussed above) to ensure that the survey addressed a comprehensive set of questions that could beneficially contribute to the design of effective communication practices in public IoT deployments.

Survey Deployment

The TrustLens survey was launched online on 21st November 2018 and concluded on 22nd November. To ensure a representative sample, recruitment activities were undertaken through a specialty survey company through a sub-contract. The target population consisted of United Kingdom residents aged 16 or over, or roughly 53,240,570 persons according to census records. 227 completed responses were received, giving a 95% confidence level, with a confidence interval of +/- 6.5. The geographic distribution of responses was fairly representative of the UK population (as determined by those who voluntarily provided the first four characters of their postcodes – 216/227) as seen in Fig. 1. Many respondents were clustered around larger cities such as London, Liverpool, Leeds, and Edinburgh;

however, a reasonable sample was also obtained from more rural and exurban areas. Summary statistics of the sample are included in Table 1 and demonstrate a good mix of respondents from a socio-demographic perspective.

Figure 1: Map of survey respondents (4-character postcode limit)

The survey was designed, first, to obtain an overall indication of people's familiarity with 'smart' technologies, such as smartphones[5] and other smart devices. Prior research has indicated that willingness to use technology is a complex construct, with trust, usefulness, and familiarity all contributing to overall attitudes (Pavlou 2003) (Gao and Bai 2014); thus, we utilised questions regarding adoption of various types of smart devices to serve as an indication of familiarity. The overall profile of awareness, ownership, and use is seen in Fig. 2.

[5] Defined as: "A mobile phone that performs many of the functions of a computer, typically having a touchscreen interface, Internet access, and an operating system capable of running downloaded apps (Lexico.com, n.d.)."

Table 1: Summary Survey Demographics

Age Category	Frequency	Percent
16-24	7	3.1%
25-34	27	11.9%
35-44	42	18.5%
45-54	45	19.8%
55-64	62	27.3%
65-74	35	15.4%
75+	7	3.1%
Prefer not to say	2	0.9%
Gender	**Frequency**	**Percent**
Male	120	52.9%
Female	104	45.8%
Prefer not to say	3	1.3%
Yearly Income Category	**Frequency**	**Percent**
<£10,000	36	15.9%
£10,000 - £24,999	70	30.8%
£25,000 - £49,999	74	32.6%
£50,000 - £74,000	12	5.3%
£75,000+	7	3.1%
Prefer not to say	28	12.3%
Highest Education Level Completed	**Frequency**	**Percent**
Lower secondary school qualification (e.g. Standard Grade, O Grade, GCSE)	48	21.1%
Upper secondary school qualification (e.g. Highers or A levels)	64	28.2%
University or FE college qualification below a degree	28	12.3%
University undergraduate degree	46	20.3%
Graduate or professional degree	32	14.1%
None of these	4	1.8%
Prefer not to say	5	2.2%
Housing Tenure	**Frequency**	**Percent**
Own (with or without a mortgage)	145	63.9%
Rented from private landlord	32	14.1%
Rented from housing association or Council	40	17.6%
Other	1	0.4%
Prefer not to say	9	4.0%

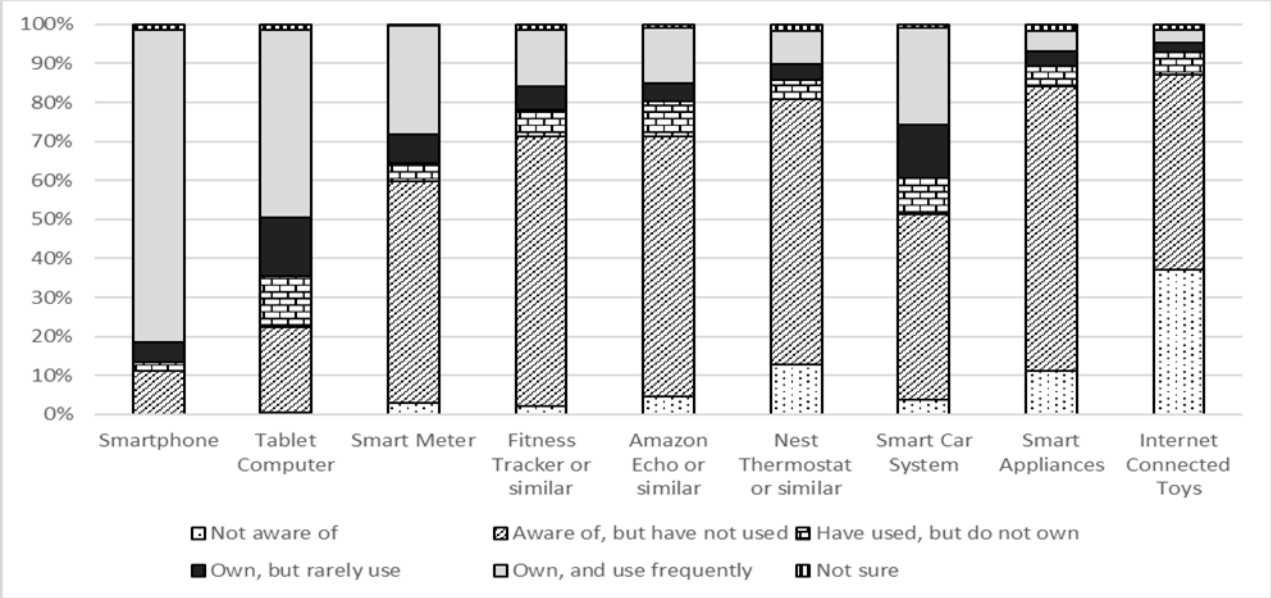

Figure 2: Awareness, ownership, and use of technology

It is evident from these results that there is good awareness of smart devices, ranging from smartphones, which all respondents indicated familiarity with, with 80% reporting that they 'own, and use frequently', to internet connected toys, which 62% of respondents reported at least being familiar with. In terms of considerations leading to the decision to purchase (or not to purchase) such devices, as seen in Figure 3, most people reported a wide variety of factors related to the device itself (such as cost and convenience), the impacts of device use (such as safety and health impacts), and issues related to the providers of the device (such as data security and trust). Of note is that, while all factors were rated quite highly (all factors had over 50% of respondents indicating that they were "most" or "very" important), those related to Security, Privacy and Data Security were ranked most highly. This may be reflective of the importance assigned to these factors; however, it may also be that these considerations apply to all smart devices.

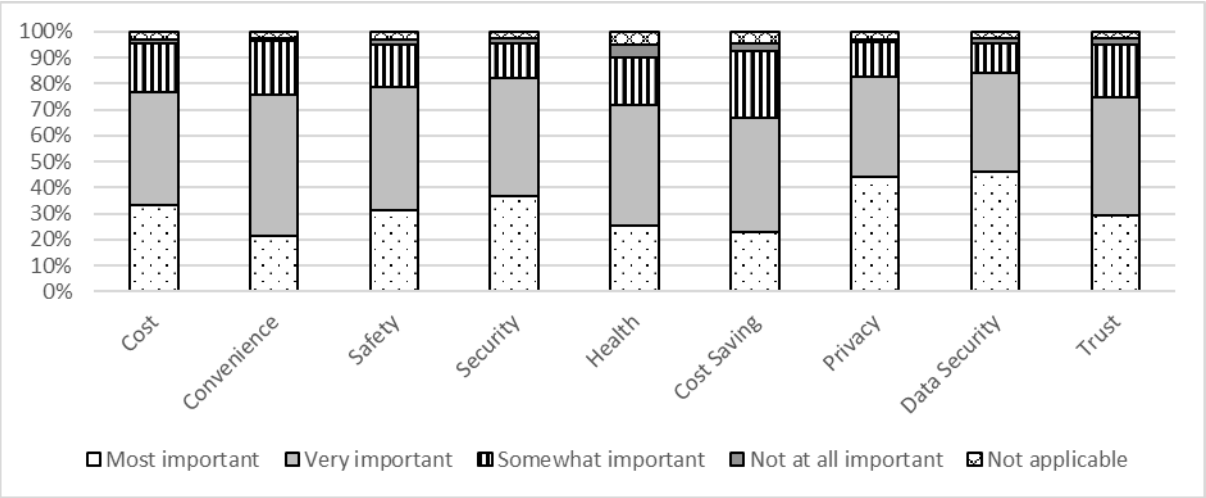

Figure 3: Factors contributing to decisions regarding to purchase smart devices

These initial findings provide an overall indication that the survey population is quite familiar with smart devices and may have various reasons for adopting or not adopting them. In the following section, we look more closely at issues related to trust and privacy to gain a better understanding of the preferences that may influence attitudes to IoT devices in public spaces.

Survey Analysis

Perceptions of Trust

When considering the use of smart devices that may have the ability to collect detailed data on the user, a core concern is that of trust. In line with work undertaken by Westin (A. F. Westin 2003), survey participants were asked to indicate the extent to which they agreed with the following three statements in order to gain a generalised baseline understanding of trust tendencies:

- I think that most people can be trusted.
- I believe most government agencies can be trusted.
- I believe most private companies can be trusted.

The results shown in Figure 4 provide a useful indication of people's beliefs overall with respect to different actors, which in turn may impact upon their degree of comfort with sharing data with these actors. It is notable that the patterns for each actor differ, with median values (where 'Disagree completely' = 1 and 'Agree completely' = 5) and standard deviations as follows:

- Government agencies: 2.8062; 1.14728
- Private companies: 2.7137; 1.07750
- People: 3.0881; 1.06480

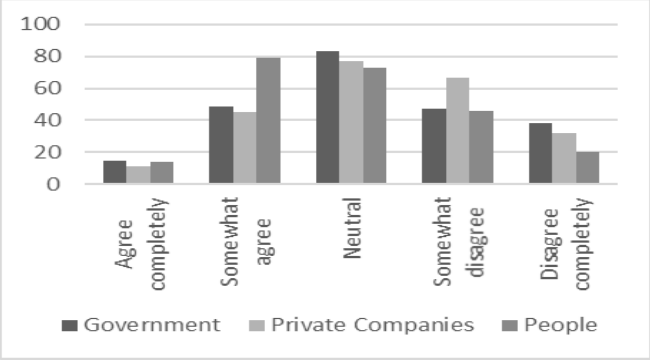

Figure 4: Extent to which survey respondents agree that specific actors can be trusted

Overall, respondents reported slightly higher trust in government agencies than private companies, though the larger standard deviation is indicative of the greater range of responses for these organisations.

Following on from Westin's three 'levels' of privacy (Unconcerned, Pragmatist, and Fundamentalist), a cluster analysis was performed to classify survey respondents by their responses to the above-mentioned questions, with results shown in Table 2.

Table 2: Median values regarding trust

	High Trust	Medium Trust	Low Trust
I think most people can be trusted.	4.03	3.19	2.19
I believe most government agencies can be trusted.	4.10	3.01	1.49
I believe most private companies can be trusted.	4.02	2.74	1.63
Total N	*58*	*97*	*72*
% N	*25.6%*	*42.7%*	*31.7%*

As may be expected, the “Medium Trust” group shows the highest number of responses overall, though the percentage of respondents in this category is lower than may have initially been expected. Of note is that the “High trust” group demonstrates a higher overall trust in government than both people in general and private companies, while the “Low trust” group demonstrates a lower trust in government than in the other groups, perhaps accounting for the higher standard deviation indicated above.

We were further interested in exploring the characteristics of users in each of the three groups, to assess if socio-demographics contribute to perceptions of trust. Variables tested included age, income, housing tenure, education, and number of smart devices owned. A multinomial regression was run, with the ‘Low trust’ group serving as the reference variable. Resulting goodness of fit characteristics, seen in Table 3, indicate a good fit for the model. A Nagelkerke pseudo R-square value of .406 further indicates that the explored characteristics explain the trust categories reasonably well (approximately 40% of observed differences), though it would be useful to further explore additional variables. Table 4 shows the influence of the considered variables.

Table 3: Model Fit Information

Model Fitting Information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	Df	Sig.
Intercept Only	461.432			
Final	360.52	100.912	60	0.001

Table 4: Model Fit Information II

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	Df	Sig.
Intercept	360.52	0	0	
OwnTech	**396.154**	**35.634**	**16**	***0.003***
Age category	380.832	20.312	14	0.121
Income level	371.023	10.503	10	0.398
Education	379.772	19.252	12	0.083
Housing Tenure	**376.279**	**15.759**	**8**	***0.046***

Characteristics with significant influence regarding category membership include the amount of technology owned and housing tenure (seen in bold). As might be expected, persons displaying high levels of trust are more likely to own more smart devices (36% of respondents in the 'high trust' category own 4 or more devices, compared to 17% in the 'low trust' category). For housing, persons who rent from a housing association or council tend to report lower levels of trust than those who own their home or rent from a private landlord.

Perceptions of Risk

While understanding the characteristics that may correspond to trust categories is useful for ascertaining how a population may demonstrate trust characteristics, for our purposes it is perhaps more relevant to understand how these populations may consider risk in public IoT deployments. Respondents were asked to provide their ranking of concern regarding a series of fictional scenarios related to IoT deployments by government agencies and private companies. Significant association is seen between category membership and degree of concern over the posed scenarios, with a chi-square of 157.052 and a significance of .000 (88 degrees of freedom), and a Nagelkerke pseudo R-square of .565. Likelihood ratio fit is shown in Table 5, while average levels of concern displayed across all categories is shown in Table 6.

Table 5: Concern with different scenarios of public IoT deployments

Effect	Labels	Model Fitting Criteria	Likelihood Ratio Tests		
		-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept		288.891[a]	0.000	0	
The City Council has access to information on your home energy use via your smart meter	***Concern_Council_SmartMeter***	***311.230***	***22.340***	***8***	***0.004***
You are required to swipe a personal ID card in order to deposit rubbish in a bin	Concern_Council_RubbishID	302.652	13.761	8	0.088
You are required to register your pet with the Council, and microchip it to allow it to be located instantly	Concern_Council_PetRegister	301.313	12.422	8	0.133
Sensors are installed at key traffic crossings, allowing drivers committing traffic offences to be identified and ticketed by local police	***Concern_Council_TrafficOffences***	***308.345***	***19.455***	***8***	***0.013***
Sensors are installed along key traffic corridors allowing vehicle emissions to be monitored and non-compliant vehicles to be identified	Concern_Council_TrafficEmissions	295.186	6.296	8	0.614
Sensors are installed along local bodies of water to monitor for potential flooding	***Concern_Flooding***	***320.101***	***31.210***	***8***	***0.000***
A private company is able to track your home energy use via your smart meter	Concern_PC_SmartMeter	301.486	12.595	8	0.127
A private company offers you insurance incentives if you track and report your physical activity levels via a personal fitness device	Concern_PC_FitnessInsurance	298.697	9.807	8	0.279
A private company provides you with sensors that will automatically detect when you are running low on staple items (such as milk, toilet roll, or laundry detergent) and place an order for delivery	Concern_PC_AutoOrder	303.304	14.413	8	0.072
You or your child is given a toy that is able to record conversations and send them to the manufacturer	Concern_ChildToy	299.853	10.963	8	0.204
A smart kettle is installed in your place of work or education that records when the kettle is boiled and by whom	Concern_SmartKettle	***313.736***	***24.845***	***8***	***0.002***

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Table 6: Average concern with scenarios across all trust levels (*0=No concern; 4=High concern*)

Scenario Label (from Table V)	Mean	Std. Deviation
Concern_Flooding	1.6608	1.12274
Concern_Council_PetRegister	1.8062	1.25415
Concern_Council_TrafficEmissions	2.0661	1.26563
Concern_Council_TrafficOffences	2.1278	1.26787
Concern_Council_SmartMeter	2.3128	1.31498
Concern_PC_SmartMeter	2.4802	1.27025
Concern_PC_FitnessInsurance	2.5066	1.30476
Concern_SmartKettle	2.5286	1.2631
Concern_PC_AutoOrder	2.5507	1.32065

Concern_Council_RubbishID	2.8634	1.23869
Concern_ChildToy	3.2952	1.18476

Of interest is that while there are significant differences in levels of concern regarding the above scenarios by trust category, these do not always follow expected patterns. In general, those persons falling in the 'Medium' trust category evidenced the lowest levels of concern for the scenarios presented, as seen in Table 7. Such figures may be placed into context when considering willingness to share information, as seen in Figure 5.

Table 7: Average level of concern by trust category *(0=No concern, 4=High concern*)

Sample Scenarios of Concern	High Trust	Medium Trust	Low Trust
Concern_SmartKettle	3.98	2.38	3.21
Concern_Flooding	2.67	1.6	2.15
Concern_Council_ TrafficOffences	3.41	2.04	2.75

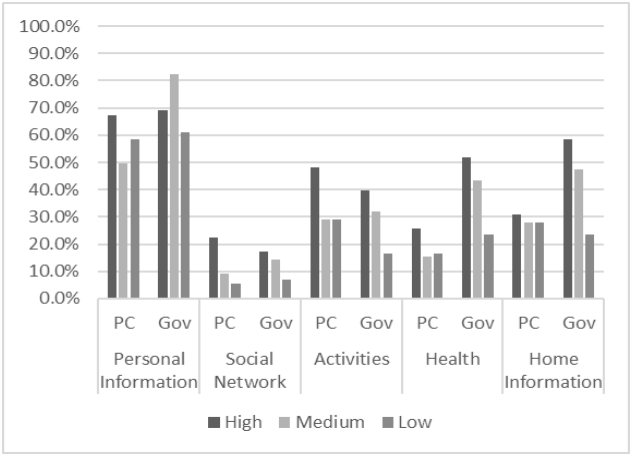

Figure 5: Willingness to share information with private companies and government by trust category

Here, we see that while respondents in the 'Medium' (or pragmatist) trust category have generally moderate options of the trustworthiness of different actors, when asked to consider sharing particular

data with those actors, they may reflect less positive responses. This is particularly evident in cases where they are asked to share data with private companies, where they are, in the instances of personal and health information, less likely to demonstrate willingness to share data than those who demonstrate lower overall trust. The generally lower lack of concern shown in the scenarios presented above is therefore somewhat surprising.

These results, however, may indicate behaviours in line with Nissenbaum's concept of 'contextual integrity', which argues that privacy is transgressed when norms of either appropriateness or norms of flow/distribution are violated. From the survey responses, it is evident that respondents are generally in agreement regarding the 'bounds' of information sharing, with most reporting that they are most willing to share personal, housing, or health information with government agencies, which may be reflective of norms of appropriateness regarding the need to provide information to receive services in return. Social network information, however, is least likely to be shared, perhaps reflecting a lack of nexus – i.e. a lack of clear connection between the data collected and the purpose for which it is being collected.

In general, the findings reflected in Tables 5, 6 and 7 indicate that all respondent clusters are least concerned about scenarios where there is a clear nexus between the data to be collected through the IoT sensors and the use of that data (for example, in the case of flood detection), or where a clear benefit to the consumer can be seen (for instance, in the pet sensor scenario). Those where the most concern is evidenced tend to have a less clear connection, or could be read to imply that data will be collected that may track personal behaviour either knowingly or unknowingly (as in the scenarios reflecting the child's toy or rubbish ID, seen in Table 5).

Preferences Regarding Information Provision

The concerns identified above indicate that communication of the data use and handling practices that sit behind such deployments are not insignificant when considering trust. To further examine this, we

asked participants to indicate what information they would want available to them when presented with the following scenario: “A bus stop you use frequently has been instrumented with technologies that will turn on lights when a person is present. It does this by periodically turning on an audio sensor to determine if a conversation is taking place and by searching for Bluetooth- or WiFi-enabled devices such as smartphones. If a person or device is sensed, the light will turn on until such time as they leave the vicinity.” Participants were asked to select information that they would want to be posted at the stop, as well as what they would want to be available online, with the responses obtained shown in Table 8.

As seen here, there is a clear preference for information to be available online, with all groups indicating higher preferences for this mode of communication in general, though with less differentiation across the low trust cluster. Obvious differences are also seen in the amount of information desired by each of the three clusters, with the low trust cluster demonstrating the highest preferences for information availability. To determine if these differences were statistically significant, an analysis of variance (ANOVA) was run to test the differences in information desired between clusters. Those attributes that showed significant differences between clusters are shown in Table 9.

Table 8: Information desired by trust cluster (*attributes receiving ≥50% response are highlighted*)

	Information Desired Online			Information Desired at Stop		
Information Attribute	**High Trust**	**Medium Trust**	**Low Trust**	**High Trust**	**Medium Trust**	**Low Trust**
What data are collected	45%	57%	64%	53%	54%	65%
Whether data are identifiable	45%	54%	68%	50%	46%	58%
Duration of retention	43%	44%	58%	36%	36%	50%
Purpose	55%	60%	71%	40%	48%	61%
Sharing policies	50%	55%	69%	47%	42%	64%
Access	33%	38%	60%	21%	28%	53%
Contact for more information	41%	40%	53%	33%	39%	54%
Privacy policy	45%	52%	64%	36%	43%	57%
Manufacturer	21%	24%	36%	24%	25%	39%
How data are protected	52%	62%	72%	53%	52%	57%
Risk indicator	31%	37%	60%	28%	28%	57%
None	12%	18%	18%	7%	19%	17%

It is notable that key differences are seen in the information desired at the location of the deployment, particularly as this reflects the wide divergence between the requirements of the low trust cluster and those of the high and medium clusters. With the exception of information on the Manufacturer, a high proportion of persons in the low trust cluster indicated that they wanted all applicable information available both at the stop and online, though a slightly higher preference for online information was revealed. It is notable that 'Privacy Policy' did not rank particularly highly as a preference across the three categories. This may be indicative of the criticisms levied against privacy policies outlined above, as it is clear that there is information that is desired; however, this may be presented more effectively in a simplified format that responds directly to user preferences.

Table 9: ANOVA test of differences between trust clusters, significant values only

		Sum of Squares	df	Mean Square	F	Sig.
Whether collected data will be identifiable (Online)	Between Groups	1.826	2	0.913	3.778	0.024
	Within Groups	54.121	224	0.242		
	Total	55.947	226			
Whether you may view collected data (Online)	Between Groups	2.842	2	1.421	6.008	0.003
	Within Groups	52.982	224	0.237		
	Total	55.824	226			
An indication of the risk the system poses to your privacy (i.e. a safety rating) (Online)	Between Groups	3.178	2	1.589	6.797	0.001
	Within Groups	52.372	224	0.234		
	Total	55.551	226			
For what purposes collected data will be used (At Stop)	Between Groups	1.532	2	0.766	3.107	0.047
	Within Groups	55.217	224	0.247		
	Total	56.749	226			
With whom collected data will be shared (At Stop)	Between Groups	2.037	2	1.018	4.169	0.017
	Within Groups	54.712	224	0.244		
	Total	56.749	226			
Whether you may view collected data (At Stop)	Between Groups	3.935	2	1.967	9.387	0.000

	Within Groups	46.946	224	0.210		
	Total	50.881	226			
Who to contact to find out more about the sensors (At Stop)	Between Groups	1.637	2	0.818	3.409	0.035
	Within Groups	53.764	224	0.240		
	Total	55.401	226			
A privacy policy concerning use of your data (At Stop)	Between Groups	1.489	2	0.744	3.039	0.050
	Within Groups	54.864	224	0.245		
	Total	56.352	226			
An indication of the risk the system poses to your privacy (i.e. a safety rating) (At Stop)	Between Groups	4.193	2	2.096	9.638	0.000
	Within Groups	48.724	224	0.218		
	Total	52.916	226			

Discussion

It is evident from the survey results analysed above that trust must be considered carefully with regards to public IoT deployments. Given the varying levels of trust identified across different actors and by the three clusters of respondents, it is clear that there is no 'one size fits all' approach to facilitating perceptions of trust in IoT systems. Considerations of risk identified indicate that the public are most concerned about deployments that have the potential to either track their personal behaviours and habits, or where no clear nexus is seen between the collection of data and the benefits obtained. This is also relevant when considering that 'Purpose of collection' was identified by all three groups as an informational element that would be desirable for a public IoT deployment. This is in line with previous work, which found that the intent and purpose behind public IoT deployments is information viewed as important by members of the public (Jacobs, et al. 2019).

This finding is also aligned with reported willingness to share data. Clear differences are seen in responses regarding the sharing of different types of information with government agencies and private companies, both supporting the clustering of individuals by trust levels and Nissenbaum's contextual

integrity argument. In reviewing the reported personal information sharing preferences of respondents, it is clear that all data are not viewed as equally sensitive. In general, people's preferences reflected the utility of sharing information – in cases where private companies or government agencies may not be viewed as 'needing' to have access to certain data, the preference would be not to share. Distinctions seen between willingness to share with government agencies and private companies reflected this differentiation, with the 'low trust' cluster showing a distinct lack of trust in government agencies with respect to their personal data.

Interestingly, the responses regarding concern with specific IoT scenarios reflected these findings, but also introduced an element of disaggregation. Again, issues of utility and nexus emerged fairly consistently, with those scenarios demonstrating a clear alignment between data collection and use ranking fairly low on the scale of concern. Also, as indicated in the findings regarding sensitivity of data as revealed through willingness to share, those scenarios that included both the potential tracking of personal behaviours and a lack of nexus were viewed as more concerning by respondents. The medium trust cluster, however, demonstrated the lowest level of concern with all scenarios presented, which was unexpected in the context of other findings. It is possible, however, that this is aligned with the pragmatism of this group – in general, their levels of trust are moderate; however, when presented with specific scenarios, the trade-off between privacy and utility may sway in favour of the latter. This finding warrants further exploration.

Finally, with respect to information provision, the findings indicate, first, that there is a clear preference for hosting information on an available online source rather than providing all information at the location of the sensor deployment. While most participants did indicate that some information should be available at the scenario location (most notably, what data are collected and how they are protected), with the exception of the 'low privacy' cluster, respondents showed an overwhelming preference for having information available online.

There were, however, clear distinctions regarding what information should be provided across the three clusters. This finding would reflect a need to design different models of information provision depending on the trust characteristics of the user. It should be possible, for example, to design three separate information interfaces for users depending upon their trust levels that would provide sufficient information to ensure their comfort, without requiring them to read through information they feel is unnecessary. Such an approach might encourage users to engage with the information provided and become informed members of the IoT ecosystem. This, in turn, would allow users to have more feelings of control, as they would be provided with a clearer and more direct understanding of the data collection purpose and practices, tailored to their preferences. While a 'plain language' approach is consistent with requirements put forward under the General Data Protection Regulation (GDPR), discussed below, the general lack of desire for access to a privacy policy indicates that the conveyance of information to the consumer in such scenarios should be presented in a manner that addresses their actual concerns.

Implications of Findings for Regulatory Purposes

According to (Greenleaf 2019), the number of countries that had enacted privacy laws from 2017-2018 had risen by 10%, from 120 to 132. This increase is demonstrative of the heightened recognition that privacy is receiving, particularly in the context of increased data collection across technology applications. While the GDPR, enacted in May 2018, has received perhaps the greatest recognition (due to a number of factors including geographic scope, severity of penalties for non-compliance, and requirements for data protection by design (Albrecht 2016), additional regulations have been enacted or updated in places including California, Maine, and Nevada in the United States; Brazil; and India (the later two of which have been modelled after the GDPR) (Determann 2020). Taking the GDPR as an exemplar, the following requirements, summarized by the UK's Information Commissioner's Office (UK ICO N.D.) are particularly pertinent to IoT deployments given the preceding discussion:

- "Individuals have the right to be informed about the collection and use of their personal data;
- You must provide individuals with information including: your purposes for processing their personal data, your retention periods for that personal data, and who it will be shared with. We call this 'privacy information';
- The information you provide to people must be concise, transparent, intelligible, easily accessible, and it must use clear and plain language (UK ICO N.D.)."

Of note is that these requirements pertain directly to 'personal data', which is defined under the GDPR as:

> … any information which are related to an identified or identifiable natural person. The data subjects are identifiable if they can be directly or indirectly identified, especially by reference to an identifier such as a name, an identification number, location data, an online identifier or one of several special characteristics, which expresses the physical, physiological, genetic, mental, commercial, cultural or social identity of these natural persons (GDPR 2016).

Determining whether data collected by IoT sensors in smart city environments is 'personal' is thus pertinent to discussions of requirements for information conveyance; however, the potential for collected data to be linked across platforms in a way that makes it personally identifying is not insignificant (European Commission 2012). Privacy-by-design approaches may experience trade-offs between the anonymization of data and the utility of data, and be susceptible to exploitation by malicious actors (Veale et al, 2019). It may therefore be advisable to respond to these requirements, particularly in public spaces, as identification of individuals may be possible based on information collected.

The findings of the survey discussed above are reflective of emerging regulatory requirements, though we also found that additional information about systems and their operation is also often desired. The requirements that such information be 'concise' and 'easily accessible' are also pertinent here, perhaps most closely aligned with attitudes towards privacy policies, which tend to be neither. A difficulty for IoT deployments in public spaces, however, is how to provide this information in a manner that is

informative, visible and accessible. As indicated from the survey responses analysed above, respondents desire access to relevant information, including who to contact to find out more about the sensors. This contact, however, may not be immediately evident, particularly given likely underlying complexities within the deployment of IoT devices in public spaces – though many cities have deployed such sensing devices, they have often partnered or contracted with third-party technology developers, who may have ownership rights over or control access to collected data (as described in (Jacobs, et al. 2019) and (Perera, et al. 2014)). This is germane not only in the context of GDPR requirements for notification and communication, but also with reference to degrees of trust and willingness to share data indicated across the high, medium, and low privacy categories discussed above. The complexities of data ownership and access rights seen in public-space deployments of IoT devices may make clear demarcations of use and processing unclear, thus the communication of practices at the outset may be useful in engendering trust across all privacy preference categories.

Conclusions and Future Research

In this paper, we have used a UK survey to explore issues of trust, privacy of personal information, and risks associated with public deployments of Internet of Things sensors. Findings indicate that, as the 'smart city' evolves, attention must be paid to consumer preferences regarding trust; information collection, sharing and use; and the methods of communication used to convey data practices. Public IoT networks introduce levels of contextual complexity that may not be addressed through conventional methods of communication; thus, further attention must be paid to this area in order to develop sufficiently robust and acceptable communication practices. Introducing such considerations at the system design stage, and aligning the method of communication with the technology deployed and location of deployment, will allow for more effective decisions to be made and will hopefully support

public acceptance of IoT devices. In addition, taking such an approach will provide more effective and robust responses to evolving regulatory requirements.

Due to the complexity of IoT networks indicated above, it could also be argued that leveraging the use of intelligent infrastructures and Artificial Intelligence (AI) will be an essential requirement to support consumer trust in future smart city development projects. Such an approach would create the need for vocabularies to allow developers to describe their IoT systems and deliver machine-understandable transparency. To address this, the *TrustLens* project has begun exploring semantic models for characterising IoT infrastructures in terms of their plans (i.e. expected behaviours), execution traces (i.e. actual behaviours) and additional contextual information such as data protection policies and physical deployment context (Markovic, et al. 2019). The results of the survey reported in this paper highlight the types of information that need to be captured by such models (e.g. what data are being recorded, what are the data being used for, who manages the IoT system, etc.) in order to allow effective translation to the public sphere. In addition, the survey suggests that consumers may be receptive to more aggregated information availability on IoT systems, particularly if these are presented in a tailored fashion that responds to the user preferences. The potential for this approach to be taken represents an area for further exploration within smart city design and development.

Acknowledgments

The work described here was funded by the award made by the RCUK Digital Economy programme to the University of Aberdeen (EP/N028074/1).

References

Albrecht, J. P. 2016. "(2016). How the GDPR will change the world." *European Data Protection Law Review* 2: 287-289.

Arcand, M., J. Nantel, M. Arles-Dufour, and A. Vincent. 2007. "The impact of reading a web site's privacy statement on perceived control over privacy and perceived trust." *Online Information Review* 31 (5): 661-681.

Bennati, S., and E. Pournaras. 2018. "Privacy-enhancing aggregation of Internet of Things data via sensors grouping." *Sustainable cities and society* 39: 387-400.

Braun, T., B. C. Fung, F. Iqbal, and B. Shah. 2018. "Security and privacy challenges in smart cities." *Sustainable cities and society* 39: 499-507.

British Standards Institution. 2014. *Smart city framework—guide to establishing strategies for smart cities and communities.* PAS 181:2014, BSI Standards Publication.

Cate, F. H., and V. Mayer-Schönberger. 2013. "Notice and consent in a world of Big Data." *International Data Privacy Law* 3 (2): 67-73.

Cottrill, C. D., and P.V. Thakuriah. 2013. "Privacy in context: an evaluation of policy-based approaches to location privacy protection." *International Journal of Law and Information Technology* 22 (2): 178-207.

Davies, N., N. Taft, M. Satyanarayanan, S. Clinch, and B. Amos. 2016. "Privacy mediators: Helping IoT cross the chasm." *Proceedings of the 17th International Workshop on Mobile Computing Systems and Applications.* 39-44.

Determann, L. 2020. *Determann's Field Guide To Data Privacy Law: International Corporate Compliance.* Cheltenham, UK: Edward Elgar Publishing.

European Commission. 2012. *IoT Privacy, Data Protection, Information Security: Fact Sheet of the European Commission.* European Commission. http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=1753.

Gao, L., and X. Bai. 2014. "A unified perspective on the factors influencing consumer acceptance of internet of things technology." *Asia Pacific Journal of Marketing and Logistics* 26 (2): 211-231.

GDPR. 2016. "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC." *Official Journal of the European Union* L 119.

Greenleaf, G. 2019. "Global data privacy laws 2019: 132 national laws & many bills." *157 Privacy Laws & Business International Report* 14-18.

Hetcher, S. 2000. "FTC As Internet Privacy Norm Entrepreneur." *The. Vand. L. Rev., 53, 2041.* 53: 2041-2062.

Jacobs, N., M. Markovic, C. D. Cottrill, P. Edwards, and K. Salt. 2019. "Public Sector Internet of Things Deployments: Value, Transparency, Risks and Challenges." *Data for Policy.* London.

Jensen, C., and C. Potts. 2004. "2004, April). Privacy policies as decision-making tools: an evaluation of online privacy notices." *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* . ACM. 471-478.

Joinson, A. N., U. D. Reips, T. Buchanan, and C. B. P. Schofield. 2010. "Privacy, trust, and self-disclosure online." *Human–Computer Interaction* 25 (1): 1-24.

Kelley, P. G., L. Cesca, J. Bresee, and L. F. Cranor. 2010. "Standardizing privacy notices: an online study of the nutrition label approach." *Proceedings of the SIGCHI Conference on Human factors in Computing Systems .* ACM. 1573-1582.

Kumaraguru, P., and L. F. Cranor. 2005. *Privacy indexes: a survey of Westin's studies.* Carnegie Mellon University, School of Computer Science, Institute for Software Research International., 368-394.

Markovic, M., Garijo, D., Edwards, P., & Vasconcelos, W. (2019, October). Semantic modelling of plans and execution traces for enhancing transparency of iot systems. In *2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS)* (pp. 110-115). IEEE.

Martin, K., and H. Nissenbaum. 2016. "Measuring privacy: an empirical test using context to expose confounding variables." *Colum. Sci. & Tech. L. Rev., 18, 176.* 18: 176-218.

McDonald, A. M., and L. F. Cranor. 2008. "The cost of reading privacy policies. ISJLP, 4, 543." *ISJLP* 4: 543-568.

McLeod, J. 2018. *Facing privacy backlash, Sidewalk Labs proposes giving data to a public trust.* October 15. Accessed June 30, 2019. https://business.financialpost.com/technology/facing-privacy-backlash.

Nissenbaum, H. 2004. " Privacy as contextual integrity. ." *Wash. L. Rev., 79, 119.* 79: 119-158.

Park, E., A. del Pobil, and S. Kwon. 2018. "The role of internet of things (IoT) in smart cities: Technology roadmap-oriented approaches." *Sustainability* 10 (5): 1388.

Patel, K. K., and S. M. Patel. 2016. "Internet of things-IOT: definition, characteristics, architecture, enabling technologies, application & future challenges." *International journal of engineering science and computing, 6(5).* 6 (5): 6122-6131.

Pavlou, P. A. 2003. "Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model." *International journal of electronic commerce* 7 (3): 101-134.

Peppet, S. R. 2014. "Regulating the internet of things: first steps toward managing discrimination, privacy, security and consent. ." *Tex. L. Rev.* 93: 85-178.

Perera, C., A. Zaslavsky, P. Christen, and D. Georgakopoulos. 2014. "Sensing as a service model for smart cities supported by internet of things." *Transactions on emerging telecommunications technologies* 25 (1): 81-93.

Pollach, I. 2007. "What's wrong with online privacy policies?" *Communications of the ACM* 50 (9): 103-108.

Siau, K., and Z. Shen. 2003. "Building customer trust in mobile commerce." *Communications of the ACM* 46 (4): 91-94.

Sicari, S., A. Rizzardi, L. A. Grieco, and A. Coen-Porisini. 2015. "Security, privacy and trust in Internet of Things: The road ahead." *Computer networks* 76: 146-164.

Silva, B. N., M. Khan, and K. Han. 2018. "Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities." *Sustainable Cities and Society* 38: 697-713.

Solove, D. J. 2013. "Privacy self-management and the consent paradox." *Harvard Law Review, 126(7), 1-880* 126 (7): 1880-1903.

Taddei, S., and B. Contena. 2013. "Privacy, trust and control: Which relationships with online self-disclosure?" *Computers in Human Behavior* 29 (3): 821-826.

Teale, C. 2020. *IDC: Global smart city spending to total $124B.* February 14. Accessed May 3, 2020. https://www.smartcitiesdive.com/news/idc-worldwide-smart-city-spending-124B-2020/572352/.

UK ICO. N.D. *The Right to be Informed.* Edited by Information Commissioner's Office. Accessed May 2020. https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-law-enforcement-processing/individual-rights/the-right-to-be-informed/.

Veale, M., Binns, R., and Ausloos, J. 2018. "When data protection by design and data subject rights clash". *International Data Privacy Law*, 8(2), 105-123.

Wang, X., J. Zhang, E. M. Schooler, and M. Ion. 2014. "Performance evaluation of attribute-based encryption: Toward data privacy in the IoT." *2014 IEEE International Conference on Communications (ICC) .* IEEE. 725-730.

Welch, V., and C. Catlett. 2015. "Urban Sensor Data Privacy Issues: Findings of the Array of Things (AoT) Privacy Breakout Group." *Position paper for STREAM2015.* Indianapolis, IN.

Westin, A. F. 2003. "Social and political dimensions of privacy." *Journal of social issues* 59 (2): 431-153.

Westin, A., and and the Staff of the Center for Social & Legal Research. 2003. *Bibliography of Surveys of the U.S. Public, 1970-2003.* *http://www.privacyexchange.org/iss/surveys/surveybibliography603.pdf.* http://www.privacyexchange.org/iss/surveys/surveybibliography603.pdf.

Wirtz, J., and M. O. Lwin. 2009. "Regulatory focus theory, trust, and privacy concern." *Journal of Service Research* 12 (2): 190-207.