Towards a Unified Conceptual Model for Surveillance Theories

"We shall meet in the place where there is no darkness" - 1984, George Orwell

Balbir S. Barn
Dept. of Computer Science
Middlesex University
London, UK
b.barn@mdx.ac.uk

Ravinder Barn
School of Law
Royal Holloway University of London
Egham, UK
r.barn@rhul.ac.uk

ABSTRACT

The erosion of values such as privacy can be a critical factor in preventing the acceptance of new innovative technology especially in challenging environments such as the criminal justice system. Erosion of privacy happens through either deliberate or inadvertent surveillance. Since Bentham's original liberal project in the 1900s, a literature and a whole study area around theories of surveillance has developed. Increasingly this general body of work has focussed on the role of information technology as a vehicle for surveillance activity. Despite an abundance of knowledge, a unified view of key surveillance concepts that is useful to designers of information systems in preventing or reducing unintended surveillance remains elusive. This paper contributes a conceptual model that synthesises the gamut of surveillance theories as a first step to a theory building effort for use by Information Systems professionals. The model is evaluated using a design science research paradigm using data from both examples of surveillance and a recently completed research project that developed technology for the UK youth justice system.

CCS CONCEPTS

• **Security and privacy** → **Social aspects of security and privacy**; • **Social and professional topics** → *Computing and business*; • **Applied computing** → *Law, social and behavioral sciences*;

KEYWORDS

Surveillance, Conceptual Model, Reference Model, Privacy

ACM Reference Format:
Balbir S. Barn and Ravinder Barn. 2018. Towards a Unified Conceptual Model for Surveillance Theories: *"We shall meet in the place where there is no darkness"* - 1984, George Orwell. In *Proceedings of ACM/IEEE International Conference on Software Engineering (ICSE 2018).* ACM, New York, NY, USA, Article 4, 10 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In a recently completed research study [5], the inadvertent erosion of values such as privacy was a critical factor in preventing the acceptance of new innovative technology in the challenging environment of UK youth justice. The work also identified the need to create new methods and approaches for accounting for values such as privacy, autonomy, transparency and trust in the information systems (IS) lifecycle[6]. In an IS context, an individual's privacy is lost through both deliberate design of surveillance technology and through inadvertent surveillance actions. Post-Snowden[27], it is both tempting and relatively common to treat privacy and surveillance as a zero-sum game with trade-offs [4]. IS designers need to acquire a more nuanced understanding of trade-offs before making design decisions that balance privacy and surveillance needs. Therefore we propose that preserving values such as privacy, is predicated by the need for a broader conceptual understanding and representation of surveillance that is expressed in the language of the IS practitioner.

Surveillance has a long history in the social sciences and philosophical disciplines where theories of surveillance have been elaborated through rich discourse. For IS practitioners such discourse remains relatively under utilised in the design of IS systems. One goal of this research is to make such theories accessible and relevant to the IS designer through the development of a conceptual model for surveillance that unifies and represents existing theories of surveillance. This model is also a direct response to the reticence expressed by Haggerty a leading surveillance studies scholar who states that:

"I am wary of the prospect of developing a model of surveillance that can usefully be generalised to all or even a considerable number of surveillance contexts." [22, 39]

Our position is that the formalism available in technologies for developing conceptual models precisely addresses the specific concerns he raises regarding the endless needs to qualify statements about surveillance. A multi-disciplinary approach brings forward new tools for deployment. Conceptual modelling is one such foundational tool from IS practice that offers a viable solution addressing Haggerty's concerns. A conceptual model is an abstraction that stresses the core terms or concepts which describe a domain and are used for fostering better understanding of a domain and communication between stakeholders on a project. A commonly quoted definition describes conceptual models as "*... descriptions of a world enterprise/slice of reality which correspond directly and naturally to our own conceptualisations of the object of these descriptions*" [36]. Three challenges immediately emerge from the idea of developing a conceptual model for a domain.

(1) What should be the sources of information for such a domain?
(2) What should be the form of the conceptual model? and
(3) How can the model be assured of its validity and relevance?

ICSE 2018, May 2018, Sweden

ACM ISBN 123-4567-24-567/08/06…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Our approach to address these challenges is through the use of design science research (DSR) and in particular, the approach is adapted from the DSR methodological innovations proposed by Pries-Heje et al. [39]. Within this approach, this paper makes two key contributions. Firstly, it reviews prevailing theories of surveillance and establishes the key constructs underpinning surveillance conceptualisation research and thus constitutes a contribution to the Hevner et al. notion of a knowledge-base. A knowledge base in Hevner et al's sense is the prior body of IS research and results from reference disciplines that provide foundational theories, instruments and constructs that can be used in further research. We can view this as the *codification* of knowledge. Our proposition is that a model for surveillance does not yet exist and is therefore an addition to a knowledge base that provides "the raw materials from and through which IS research is accomplished. The knowledge base is composed of foundations and methodologies." [1]. This action also addresses challenge (1) head-on, in that we review within, the limitations of space, some of the key resources contributing to this field of study. Secondly, the resulting design artefact is expressed as a conceptual model using a Unified Modelling Language (UML) based modelling approach (challenge (2)) The research outcome, i.e. the conceptual model is evaluated by adopting a Naturalistic and Ex-Post evaluation strategy [39] (challenge 3).

The remainder of the paper is structured as follows: Section 2 outlines the key stages of our approach to developing the conceptual model. Section 3 introduces the background to surveillance research. Component elements of surveillance research are discussed, together with challenges associated with defining surveillance. Section 4 presents the main contribution of our work in the form of a conceptual model for surveillance that has capability for both extension with new concepts and classifying new forms of surveillance. The taxonomy is presented as a UML conceptual model. We recognise that our model is a case of emergent theory building in the sense of Doty & Glick [14] and so we present an evaluation based on DSR that uses a mix of evaluation criteria and a detailed case study scenario in Section 5. Finally, in section 6, we present concluding remarks and further research plans.

2 METHODOLOGY

In section 1, the challenge of a unified, generalised model of surveillance was proposed. This paper proposes a conceptual modelling approach to address the idea of a generalised model of surveillance. Technologies such as conceptual models that can support utility (profit and/or other goals) are appropriately identified and used through design science research methodology (DSR). To that end, we draw upon the DSR process proposed by Peffers et al. [37] and execute three essential activities from their nominal process to realise three design science research cycles of relevance, design and rigour.

- Identify Problem and Motivate: justification of the problem existence
- Design and Development (of the artefact)
- Evaluation of the usage of the artefact (using a naturalistic case study).

The *Identify Problem and Motivate* activity defines the research problem and justifies the value of a solution [37]. Typically, researchers explore theoretical bases that improve the rigour or consider practical relevance that improve the situation on the ground as the basis for identifying the problem.

The *Design* cycle essentially deals with design and development of artefacts. In this cycle, our intended output artefact is a proposal for a unified, generalised conceptual model that attempts to capture the concepts and relationships described in a range of surveillance theories.

The *Evaluation* cycle comprises activities that perform evaluation of the usage of the artefacts. For the evaluation activity, several authors such as Hevner [24] and Prat et al. [38] define a number of criteria for evaluation purposes. The latter, in particular, collate a set of criteria following a review of literature. We select a subset of criteria for evaluation based on our understanding of the problem definition. These criteria are: *efficacy* - the degree to which the artefact achieves its desired effect; *completeness* - akin to and amounts to functionality; *systematic construction* - the approach taken to construct the model including reference to the sources of knowledge. Finally the *Modeling Language* criteria - the choice of language for expressing the model is evaluated. In addition to the criteria, we position our evaluation strategy as one that is Naturalistic and Ex-Post [39]. It is Ex-Post as the evaluation is taking place after the design of the artefact. It is Naturalistic in that we are using a case study to provide our evidence that is based on authentic, primary data. The DSR methodology adopted is shown schematically in Figure 1.

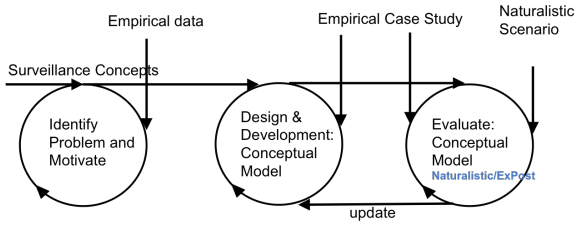

Figure 1: Design Science Research Methodology

Conceptual models as theory representations

A key assumption in this work is the relationship between theory and conceptual models. Here we outline why a conceptual modelling approach is appropriate to define a unified surveillance theory. In their exhaustive review of surveillance theory research, Galič et al. [18] present a number of theories that could broadly be described as mid-range grand theories in that they are relatively well developed and used to describe and explain embedded phenomena [21, 33].

The word "theory" suffers from both an over-use and a reluctance in its use by researchers. Weick comments that most theories that are labeled as theories are actually *approximate* theory in that they go some way to establishing a theory but fall short in some aspect such as: failing to sufficiently articulate relationships between variables/concepts contributing to the theory; or perhaps where ad hoc hypotheses are derived from limited observations [46].

Space does not allow for a detailed philosophical discussion of the nature of a theory, instead, we draw on some representative

descriptions to outline what is generally agreed, constitutes a theory [42, 43, 47]. The main elements of a theory are:

constructs: the basic conceptual elements extracted from the domain of discourse that are generally measurable.
relations: relations describe connections among constructs and their interactions with one another.
boundary: a boundary of a theory describes its scope or the validity of the theory under certain conditions.
propositions: statements that are concerned with making predictions about a theory's constructs.

Others have noted that scientific theories demonstrate analogous properties to conceptual models:

"In system development, the purpose of the conceptual model is to describe the elements of the domain and their relationships (Mylopoulos, 1992). The conceptual model serves as the basis of understanding and problem solving within the domain. It allows analysts to capture and communicate their understanding of a domain and the problems in the domain. In science, it is the role of theories to describe the constituents of the domain, their relationships and their behavior, and to serve as the means by which problems in a domain are specified and solved." [15].

Given this analogy, it appears viable that a conceptual model defining constructs, relations, constraints and possibly behaviours (propositions) could be used as a representative form for describing generalised surveillance theories that can serve to bridge the language gap between social scientists and information systems designers to address concerns such as accidental erosion of privacy when implementing systems.

Theories in general can offer routes to analysis and prediction, explanation, prediction and prescription (a description of a method) [20]. The theories identified from surveillance literature in the next section, all offer routes to analysis and explanation and have implicitly embedded in them, the properties listed above: constructs, relations, boundaries and propositions. The nature of social science theories is such that their primary purpose appears to be as a tool for analysis and it is this paper's task to extract the constructs and relationships. The predictive elements of a theory are out of scope in this paper.

3 SURVEILLANCE LITERATURE AND KEY CONCEPTS

Stage 1 (*Identify Problem and Motivate*) of the adopted DSR approach is achieved through an analysis of existing surveillance theory. The recent review by Galič et al. [18] is a useful starting point as their paper aims to provide an overview of surveillance theories and concepts that can help to understand and debate surveillance in its many forms. In this section, we introduce the concepts in their basic form.

Although mostly chronological, the Galič et al. paper explores a range of theories by clustering them into three phases of surveillance theory building. The first phase is seen as largely architectural and is attributed to Bentham and his octangular shaped designs for prisons and other buildings. Through architectural design, the *prison-Panopticon* enables an illusion of constant surveillance [8]. However, it is Foucault's subsequent analysis of discipline and the use of the Panopticon as an analytical tool for discussing institutions and society [16] that the Panopticon in the form of *panopticism* has become the mostly widely used metaphor for surveillance today. Bentham's liberal political projects extended the panopticon paradigm to other aspects of society including the *Pauper Panopticon*, (industry workhouses), the *Chrestomathic-Panopticon* (the Schoolmaster supervising 600 children without being seen) and the *Constitutional-Panopticon* (where citizens watch those who govern to ensure that there is no misrule) [41]. The latter two examples, exhibit aspects of surveillance without the power relations and disciplining notions so strongly associated with Foucault's focus on the prisonor-Panopticon and also move beyond the margins of society to more mainstream components of society.

The second phase of surveillance theory building offers theories about networked or distributed/remote surveillance that rely primarily on digital technologies. The watched may include data doubles [22] as well as hybrid individuals whose physical and data double goes beyond a straightforward 1-1 correspondence [25]. In this phase, there is also a proliferation of new purposes of surveillance such as deterrence, consumption, entertainment, transparency of accountability, health as well as security [22]. Furthermore, non-human targets of surveillance that were previously neglected are also identified (e.g. satellite imagery on deforestation). The watchers also change. Corporations, in their desire to make effective decisions for their specific needs (rather than national interest), seek to have constant control via continuous monitoring and assessment of markets, workforces, performance, strategies etc. Critically, individuals and their disciplining are of less interest, instead it is the individual's multiple different representations that matter (coined *'dividuals'* by Deleuze) [12]. The focus on digital technologies brings into the foreground, a further complexity: the notion of Surveillance Assemblages - 'multiplicity of heterogenous objects, whose unity comes solely from the fact that these items function together, that they work together as a functional entity' [13] cited in [23]. These assemblages are themselves assemblies: "...any particular assemblage is itself composed of different discrete assemblages which are themselves multiple..." [23, :608]. Such structures are made possible because of technological advances that can integrate disparate systems to function as a single system. In these assemblages, flows the information representations of *dividuals*. The advent of big data, and the emergence of a new economic logic, where the seductive conjunction of prediction and monetization of online transactions of the minutae of both *dividuals*, hybrids and data doubles has created what Zuboff describes as *Surveillance capitalism* [49].

The third cluster of surveillance theories critiqued by Galič et al. focus on more recent contemporary analyses where refinements of ideas originating from the earlier phases are reified and more user-specific perspectives of participative surveillance and resistance are observed. Three specific aspects can be noted.

(1) Increased international terrorism activity post 9/11 led to large-scale mass surveillance of communications of ordinary citizens by nation states often with complicity from commercial providers as revealed by Snowden [27].

(2) The relentless rise of social media has blurred the notion of the watcher and the watched, and added pleasure / entertainment as an intrinsic motivation for surveillance, through for example,

how surveillance is in-built as a design principle in on-line games [3]. The concept of *Participatory surveillance* also seems to be the cornerstone in most social media apps (*sharing*, *liking* and *following* for example)[2, 10] and extends into self-surveillance through for example, the use of mobile health apps, fitness trackers as the ultimate form of nudging where health responsibilities are relocated to the individual [48].

(3) It is also argued that the use of machine learning algorithms offers a more 'objective' judgement of social-sorting or categorisation and can address issues of misjudgement and prejudice of humans but this of course returns us to our opening introduction of how values are embedded in the design of software [17].

3.1 Security and privacy research

Computer science research literature on security and privacy concerns has focussed intensively on technological innovation. Often the emphasis has been reviewing such concerns as non-functional requirements [19]. To place our analysis of surveillance theories in context, it is helpful to review several recent systematic mapping studies on security and privacy research [7, 19, 30, 32]. These studies indicate the role of better human understanding. Hence, Barth and Menno, through their analysis of underpinning theories that reflect human behaviour, suggest that the so-called security paradox, the discrepancy between user attitudes and user behaviour can be explained either by a rational calculation of risk and benefit, or an irrational risk-benefit calculation based on a biased risk assessment. A third explanation offered is that of a failed privacy valuation or information deficit [7]. Indeed, it is this privacy paradox and decision making process by a user that creates opportunities for surveillance capitalism discussed earlier. The most closely related research to our own is the work by Gharib et al. [19]. Here, the researchers propose an ontology of security concepts derived from a systematic literature review of security requirements. Their ontology outlines a range of concepts including Agentive entities (our Actors), Intentional entities (Motivation), Interactions through which actors achieve goals (Actions). Significantly, surveillance research in computer science literature has not reviewed the underpinning sociological perspectives in the manner discussed in this paper.

So the question is how to capture the complexity of these theories?

4 A UNIFIED CONCEPTUAL MODEL OF SURVEILLANCE

In this section we present a conceptual model using the Unified Modelling Language (UML) that is an initial effort at addressing Haggerty's challenge introduced in the first section of this paper. The model is shown in its entirety in Figure 2 and is the principal design science artefact arising from stage 2 of the design science research methodology outlined in section 2. The model concepts were derived from review of surveillance theory literature and aim to provide a compact picture of an integrated view of surveillance theory discussion presented in section 3 earlier.

The literature describes a range of actors who participate in surveillance either as targets of surveillance (*Surveilled*) or as perpetrators of surveillance (*Observer*). Such *SurveillanceParticipants* include the general *Population* at large, individual *Humans*, their *Information Representation* or even *Hybrids* (where there is a close correspondence between the actual human and the information representation). *Surveillance Participants* can also be a *Physical Space* (such as the Amazon rainforest, or the plains of Iraq), a *State Actor* (e.g. NSA) or a *Corporation* such as a Google. *Information Representations* are what Deleuze referred to as 'dvidiuals' (partial representations of an individual) and include for example, say, the monitoring of an email account or a Active Directory account [44]. There are times when a *SurveillanceParticipant* is both the *Observer* and *Surveilled* in an act of *Surveillance*. Such types of *Surveillance*, are categorised as *Self*, *Intrusive* or *Participative*.

The conceptual model attempts to unify the motivations for conducting surveillance from the perspective of the observer (the perpetrator, let us say). Drawing on motivation theory [40], we delineate two categories for *Motivation,* namely, *Intrinsic* (the doing of an activity for its inherent satisfactions rather than for some separable consequence [40] where there is some valency of free choice) and *Extrinsic* (whenever an activity is done in order to attain some separable outcome). The latter can come with some sense of autonomy. We propose that some examples of surveillance (such as participative) are motivated by pleasure.

Any *Surveillance* action is mediated through *Technology* in its broadest sense. Thus Bentham's Panopticon, including the less discussed versions of the *Pauper Panopticon* and *Chrestomatic Panopticon* (School) are technologies based on architectural designs. In Panopticon design, Surveillance is possible because:

> It is obvious that, in all these instances, the more constantly the persons to be inspected are under the eyes of the persons who should inspect them, the more perfectly will the purpose of the establishment have been attained. Ideal perfection, if that were the object, would require that each person should actually be in that predicament, during every instant of time. This being impossible, the next thing to be wished for is, that, at every instant, seeing reason to believe as much, and not being able to satisfy himself to the contrary, he should conceive himself to be so. [9, 69]

Bentham, also introduced the concept of *Constitutional Panopticon* where the concept of continuous visibility is reduced. The surveillant, the government official is only observed in the course of their duties and through technology such as media and the law.

The *Surveillance Assemblage*, the ability for it to be nested and assembled through other assemblages comprising either a *Hardware Platform* or *Software* or a mix of the two is another technology that is essentially the 20th century tooling for surveillance. CCTV for example is an assemblage of cameras, monitor screens, human intervention and machine learning algorithms (for more advanced options). Similarly, the various inter-related systems revealed by Snowden in 2013, describe a state sponsored assemblage where the flow of information includes meta data. *Surveillance*, using mediating *Technology*, is performed through some *Action* which is either: *Human Action* (when accomplished by another human actor); *Hybrid Action* (such as 'following the Twitter handle of some target Surveillance Participant); *Population Action* (when a nation's citizens monitor their government's actions through the

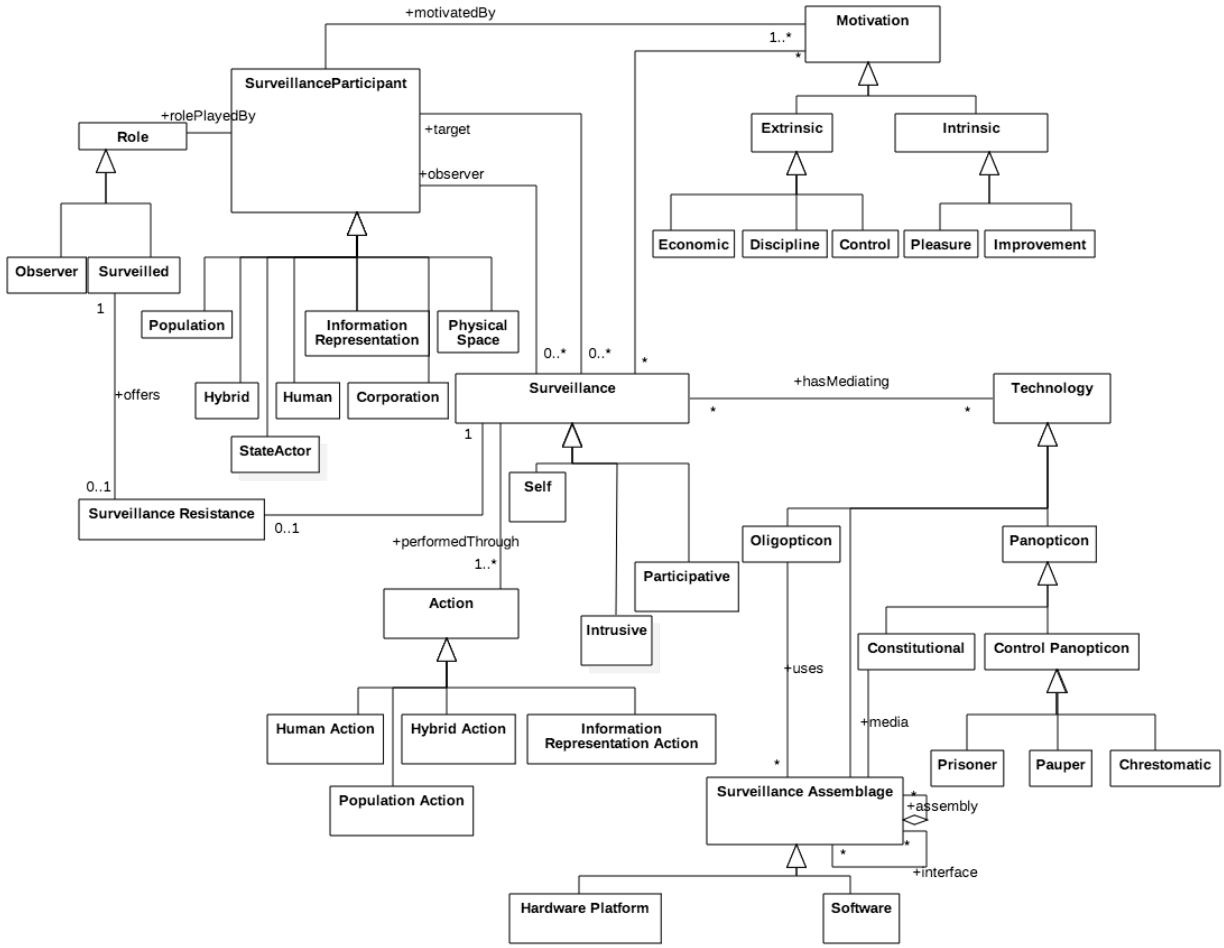

Figure 2: An Integrated Model of Surveillance Theories

independent media); and *Information Representation Action* (when technology is used to monitor the access of file stores using say, Active Directory ID). A benefit of using a type (class) model for is the possibility of ensuring a well-defined semantics. Here, we limit the semantics of this model to be a collection of object models that are instances of the semantic model. The semantic model comprises objects and slots that contain values. Additionally, there are well-formed rules that determine how an instance model is deemed to be correct with respect to the conceptual model.

5 DISCUSSION AND EVALUATION

We discuss and evaluate the surveillance model by two approaches: evaluation of key properties about quality of models and evaluation of two generic use cases that describe functional uses of models. In the introduction, a challenge of how can the model be assured of its validity and relevance was presented. In this section, it is suggested that validity is determined through properties such as representation using a modeling language (with its own semantics) and through notions of completeness. Relevance of the model is demonstrated through the use cases discussed here.

5.1 Model Quality

Conceptual models are widely regarded as pivotal to the design of information systems and technology providing that the models themselves are perceived to be at an appropriate quality threshold. Approaches to evaluation of conceptual models are, therefore, well documented in the IS literature. Much of the research proposes the need to use multiple criteria for assessing notions of quality of a model. Such criteria include: simplicity, understandability (both taken together as efficacy), flexibility, completeness, implementability, correctness, relevance and systematic construction (for example [35]. We propose that such evaluation criteria and their usage depends upon the maturity of the conceptual model under consideration.

Efficacy: The degree to which a model achieves its desired effect can be viewed as an amalgam of two other criteria: simplicity and

Table 1: Sources of Concepts - Part 1

Concept	Specific Literature Source	Notes	Concept	Specific Literature Source	Notes
Surveillance Participant	Derived	General notion of a participant in a surveillance action.	Surveillance	Derived	Generalised notion of Surveillance
Space	Haggerty [22]	Monitoring of a physical space such as via satellite, or a public square.	Self Surveillance	Albrechtslund [2]	Comparing oneself with others
Information Representation	Deleuze [12]	The actual person is of less interest than the individual's representation as an information structure. For example, the consumer and their purcashing power	Participative Surveillance	Boyd and Ellison [10] Marwick [31]	Citizens/users are actively engaged in surveillance themselves as watchers, but they also participate voluntarily and consciously in the role of watched.
Corporation	Zuboff [49], Deleuze [12]	Corporations aim to derive profit from unilateral surveillance and modification of human behaviour.	Motivation	Ryan [40]	Generalised notion of motivation
Hybrid	Haggerty & Ericson [23]	A merged construct representing interactions between an information representation and a physical being	Intrinsic Motivation (Pleasure or Improvement)	Ryan [40] Lyon [29]	Surveillance conducted for pleasure or improvement such as comparing one's performance relative to others on social media

understandability. Reviewing this criteria, we observe that the modelling language emanates from computer science whilst the domain is inter-disciplinary. Hence issues of understandability, the ability of non-computer scientists to interpret the model and benefit from the semantics associated with the notations, used are of concern. At the same time, we note that we are using a very small subset of UML notation (class diagram) and class models can be "read". The relatively small set of concepts focussed on representing things [45] also means that the simplicity criteria is met. Given the technology oriented aspect of modern day surveillance, we anticipate that reading and therefore understanding such models is less of a concern in the future.

Modeling Language: Weber notes that any modelling language requires a number of features for ontological completeness so that the language has sufficiency of completeness to represent ontological phenomena from the real world [45]. Such features include concepts to represent things, properties of things, types states and laws (or constraints). Our choice of UML as noted earlier provides these features and hence is sufficient in its ability to express our desired conceptual model. The availability of software tools, education and widespread acceptance of UML means also that it augments other criteria.

Completeness: Assessing the completeness of a model is challenging and is indeed the problem raised by Haggerty and discussed in the introduction. The response to this completeness property is based on three arguments. Firstly, the conceptual model has been inductively produced from extant literature. Concepts presented in the model have a reference research literature associated with them. Table 1 and Table 2 present the concepts and their cited references. New concepts that are untested or have no ontological reality have not been deployed. Secondly, the use of a semi-formal modelling approach allows generalisations and combinations of valid model elements to be enumerated based on the underlying semantics of UML class models. Thirdly, our chosen semantics for the model is a collection of object models that are instances of the semantic model where the semantics are comprises objects and slots that contain values. Hence we can enumerate scenarios of cases of surveillance using the use cases below as a guide. Table 3 provides examples of such scenarios.

5.2 Use Cases

A second aspect of our evaluation approach is to explore *what* purpose our conceptual model can be used for. We do this by articulating two use cases.

Table 2: Sources of Concepts - Part 2

Concept	Specific Literature Source	Notes	Concept	Specific Literature Source	Notes
Human	Derived	A physical being either conducting surveillance or the target of surveillance actions	Extrinsic Motivation	Ryan [40]	Externalised explanation for conducting of a surveillance action
Population	Haggerty & Ericson [23] Foucault [16]	The disciplining of whole populations through governmentality to optimally regulate social behaviour.	Technology	Derived	Generalised
Role	Derived	The role of a surveillance participant: either the Observer or the Target	Panopticon (Prison, Pauper, Chrestomathic, Consitutional subtypes)	Bentham [8, 9], Foucault [16]	The original technology envisaged by Bentham for control of prisons and then adapted.
Action (Human, Hybrid, Information Representation)	Derived	A form of surveillance action conducted by humans, or information representations	Surveillance Assemblage (Hardware / Software *assembly*)	Deleuze & Guattari [13], Haggerty & Ericson [23]	A networked assembly of hardware and software systems forming a coherent whole with well defined interfaces for conducting surveillance
Surveillance Resistance	Latour [28] Brunton & Nissenbaum [11]	Practical steps to mitigate against surveillance	Oligopticon (*Viewshed* relationship)	Latour [28]	A partial but sturdy view of a target that may be integrated into an assemblage

Table 3: Use Case 1: Review surveillance properties: Case Scenarios

Case	Surveillance Observer	Surveillance Target	Motivation	Surveillance Type	Surveillance Action	Technology	Surveillance Resistance
Snowden Revelations	StateActor: NSA	Population: USPopulation	Control	Surveillance Type: Phone Record Metadata	Surveillance Action: Large Scale Information Action Collection	Surveillance Assemblage: Phone Companies:	None
Performativity	Corporation: University	Human: Academic Staff	Extrinsic: Discipline/Control	Collection of Data Entries to Student Dashboard System	Hybrid Action: Review of Updates of Student Interactions	Surveillance Assemblage: Integrated Student Dashboard	Minimal Data Entry by Academic
Social Network	Hybrid: SNS Member	Hybrid: SNS Member	Pleasure:	Participative	Hybrid Action: Like	Surveillance Assemblage: SNS	None

1. **Review surveillance properties:** This use case is a general action that is used to review a situation to determine the type of surveillance, motivation, actors involved and mediating technology. The use case can also be used in naturalistic/ExPost mode.
2. **Assess surveillance concerns of proposed system:** This use case is used to review data from usage of a system to

provide a qualitative assessment of value concerns arising from the deployment of the system.

5.2.1 Review surveillance properties. Table 3 presents three example reviews of the surveillance aspects of cases where an analysis of surveillance is appropriate. Table 3 can be read as both type and instance data. The first case: *Snowden Revelations* is derived from the 2013 incident of whistleblowing in the security sector [27]. In this scenario, the Surveillance Observer is a StateActor, the National Security Agency (NSA). The Surveillance Target is type Population, the population of the USA. The scenario further details, the type of surveillance and action undertaken, together with the technologies concerned. Given that this is happening in secret, the surveillance targets did not offer any resistance action. The two other cases can be read in a similar fashion. The *Performativity* case describes a notion of accidental surveillance, through the monitoring of academic staff in their daily business of recording their interactions with students on a personalised student dashboard. The academic staff can assert some resistance to this type of surveillance by determining what data is to be entered. The third case, *Social Network* draws out the participative surveillance in social network sites. Here, the motivation for this type of surveillance is pleasure and realised through pressing "Like" buttons on a fellow social network site member's social feed.

5.2.2 Assess surveillance concerns of proposed system. A second use of the surveillance model is to understand the potential impact of surveillance on technology acceptance within a user base, of a proposed system. In particular, how to anticipate accidental or unintended surveillance. To illustrate an evaluation of the model for this, we draw upon primary data from a research study that first exposed us to this central issue of erosion of privacy concerns through accidental surveillance.

The MAYOT (Mobile Applications for Youth Offending Teams) project developed a personalised mobile app for use by young people and their case workers in youth offending teams in the UK Youth Justice system. The app was intended to provide relevant, timely information to a young person to help them manage the requirements of their court order. The project adopted a mixed-methods approach to determine the current and intended/desired use of technology including the use of quantitative survey, co-design workshops, and interviews from three youth offending services, representing inner-city, urban, and rural locations in England. For the co-design workshops, 17 case workers and 10 young people participated to contribute ideas to the design and development of the social technology. One workshop at a hitherto unused site was used to independently assess and evaluate the design ideas. The qualitative comments come from that workshop.

The app offered a range of features, however, in this paper, we are concerned primarily with how values such as privacy and data sensitivity affect the sense of deliberate or accidental surveillance of young people. Thus, we limit our reporting to the impact on the same and use one of the derived functions/features, "Exclusion Zone" for this purpose. The Exclusion Zone feature is a function that is available on the MAYOT app that allows a case worker to define geographic regions from which a young person is prohibited (with a potential risks to violating their youth order with obvious detrimental effects). The feature alerts the young person in possession of the smart phone hosting the app that they are in an exclusion zone.

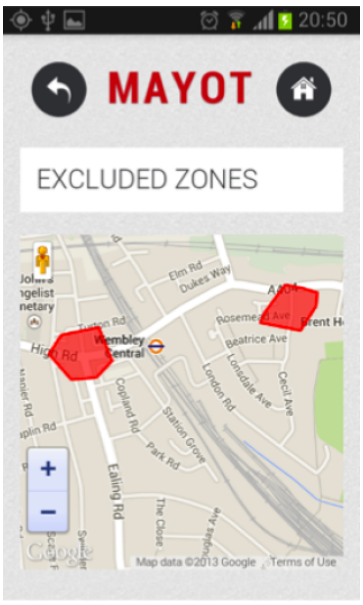

Figure 3: Exclusion Zone Feature of the MAYOT app

The feature is illustrated in Figure 3. The requirement was originally suggested by case workers:

> CASE WORKER (CW): ". . . maybe bespoke. . . some young people are prohibited from going into certain areas so maybe their phone could vibrate if they are getting close to that area".

The *Motivation* was for the *Improvement* of the young person and there was an assumption that the young person would carry the phone with them, such that the phone number became an *Information Representation* of the *Human*. Initial reactions from young people was positive and further support the notion that the Motivation is for Improvement:

> Young person (YP) 1: "So your phone's gonna vibrate when you cross?" "Yeah, that would be alright."
> YP 2: "…yeah, that would be quite useful."

The functionality afforded by the Exclusion Zone feature can be seen as a *Surveillance Assemblage*, in that software, servers, hardware (the mobile phone) act together in a *Hybrid Action* to perform unintended *Surveillance* on the young person. In subsequent workshops, when the design was relayed to a new group of young people, the reaction was diferent:

> J: (interrupts) "so it actually tells the YOT workers and that, that I'm in that area?" (over top - LQ "does it tell anyone?")
> LQ: "like putting yourself on tag by having the app on your phone in the first place so you'd just be like boom"

The young people also identified the *Panopticon* characteristics of the technology where the watched cannot verify they are being watched.

> LQ: "cos you don't know whether they're keeping an eye on you"

Rapidly, it became apparent that there was an implicit trade-off but before any decisions around acceptance could be made, it was important to see the full extent of the *Surveillance Assemblage*.

> LQ "...its got good points to it but to accept the good points you have to feel like you're being watched kind of thing...you should show us their side of the app".

Despite the view that App had potential tracking facilities, there was also recognition, on the part of young people, that the App could be a source of power for them and allow them to offer *Surveillance Resistance* in cases where the police data/perception was inaccurate or out of date or even delete the app.:

> YP 3: "it would be good if like police try to stop you or something and they've got the details and stuff and they put it through the system or you're not allowed in this area and you pull out your phone and be like yeah well I'm not in that area. You know what I mean? To prove them wrong."

Such perceptions of surveillance were not restricted to just the young people, case workers in the independent evaluation workshop identified the surveillance aspects of the app feature:

> CW 2: "And also kind of morally and philosophically, some of it seems slightly Orwellian. Um, in kind of, sort of perspective. I just have issues"
> CW 3: "...well, just, uh I guess in a sort of Orwellian, dystopian... I know obviously there's an essence that's the nature of the work"

This qualitative data extracted from just one of the features of the app, demonstrates the relative efficacy of the conceptual model in that it allows us to develop an annotation language for marking up qualitative data. From a completeness perspective, much of model is exercised. The model also exhibits homomorphism with the realities described in the case example. One issue is that temporal aspects are harder to establish. For example, a perception of an altruistic motivation of surveillance may change during the design of the software or after deployment. Furthermore, the notion of a viewpoint (at a point in time) needs to established.

Both use cases reveal an underlying point that returns to the centrality of the role of technology. Data collection for whatever purpose (performativity, social network, security or design of systems) becomes a proxy for surveillance. That data as Kitchin suggests may be collected through a variety of means [26]:

- targeted, where the gaze of a technology based proxy authority is aimed at individuals or places;
- automated, where data is generated as an inherent function of the ICT component, such as in sensors;
- volunteered, where data is freely given by students or staff, such as through interactions on educational or proxy social media platforms.

Developing the performativity in higher education use case a little further (case 2): Targeted surveillance can be mediated through the use of barcode scanning systems installed in classrooms (ostensibly to measure room usage, but leading to information about an individual's attendance). Automated data collection through the use of under-desk sensors measures presence and hence room usage but provides a deep insight into an individuals' work patterns. Social media platforms such as Yammer (for corporate communications), or Slack for student projects provides quantified data about some notion of engagement.

Integrated data collection and the implicit surveillance suggests that current practices of governance will need adaptation and innovation in parallel to the technological innovation being deployed. Such governance changes will need to consider some key challenges:

Technocratic Governmentality - Leading to 'fabricated' Narratives. Continuous monitoring of activity, analysis and then decision-making presumes an instrumented rationality and then subsequent control. Danger exists where reasons may not be sought, instead, emergent, 'interesting' correlations are acted on that preclude moral reflection (see also value-sensitive algorithms). Moreover, data becomes politicised and is no longer objective. Technocratic governmentality may lead to 'fabricated' narratives which occlude or ignore information that cannot measured easily.

Reification of bio-power. The promise of mobile apps to solve specific perceived issues may create greater autonomy, personalised information and a sense of belonging. The app, however can easily malfunction or be deliberately disabled leading to an asymetrical power relation between those who manage the app and its services and those availing of the services.

Value-sensitive algorithms. In the same manner that data is not apolitical, algorithms underpinning academic analytics and decision making are also not apolitical. Algorithms embed values of their designers and because of the inherent lack of transparency of the computation undertaken in the algorithm nor are the values easy to identify. Algorithms also pose other issues such as: probable results (with inevitable uncertainty), unfair outcomes (inherent biases), and traceability concerns where data combinations are difficult to track back [34].

Corporate data colonialism. The role of large corporations, through the promotion of a neoliberal political economy means that corporations are in a position to colonise data generated in public environments with a view to a monetisation of that data.

5.3 Threats to Validity

This work is an example of a research strategy that is building theory represented as a conceptual model from case studies as case studies are rich empirical descriptions of instances of a phenomenon. We recognise the limitations of case studies in that each case study may represent a discrete experiment. However, through both high level broad descriptions of cases (as in Table 3) and in the more in-depth analysis of our mobile app case study, we have attempted to mitigate against this limitation. Another issue is that of the domain itself. It could be argued that application domain of the MAYOT project is not ideal in terms of the properties of preservation of privacy, security concerns given that surveillance

seems to be a design objective although not one intended by the original conceivers of the app.

The proposed conceptual model is one such extraction of the key properties of theories: concepts and relationships. Limitations associated with the model include: omissions of ethics and compliance (they are left implicit in the model) and a focus on static information rather than dynamic or behavioural information. Our evaluation against the criteria of quality of models, also notes that different criteria are relevant depending upon the stage of acceptance of a conceptual model. Such limitations will form the basis of future work.

We also remark, that inter-disciplinary studies come with some disadvantages, notably, translating from social science to reductive models come with their own challenges of an appropriate shared language.

6 CONCLUSION

In this paper, we have analysed the key literature in surveillance research to extract, unify and relate concepts in surveillance. In doing so, we have contributed what we believe, to be the first such conceptual model of surveillance for use by Information Systems designers. The compactness of the model, enables the domain of surveillance to become more accessible to IS practitioners. The model has an immediate use in that it can be used as part of a process to review design artefacts from information systems development or it can be used to identify technology acceptance concerns. Our next step is to move from the closed world of conceptual modelling to developing an ontology that can be developed through public collaboration.

REFERENCES

[1] R Hevner Alan, Salvatore T March, Jinsoo Park, and Sudha Ram. 2004. Design science in information systems research. *MIS quarterly* 28, 1 (2004), 75–105.

[2] Anders Albrechtslund. 2008. Online social networking as participatory surveillance. *First Monday* 13, 3 (2008).

[3] Anders Albrechtslund and Lynsey Dubbeld. 2002. The plays and arts of surveillance: Studying surveillance as entertainment. *Surveillance & Society* 3, 2/3 (2002).

[4] Balbir S. Barn and Ravinder Barn. 2015. The dilemma of cyber security and privacy: On the role of value sensitive design. In *Working Papers of Sustainability Society Network+, Proceedings of the 1st International Conference Cyber Security for Sustainable Society.*

[5] Balbir S Barn and Ravinder Barn. 2016. An exploration of resilience and values in the co-design of sociotechnical systems. *International Journal of Systems and Society (IJSS)* 3, 1 (2016), 1–17.

[6] Balbir S. Barn, Ravinder Barn, and Franco Raimondi. 2015. On the Role of Value Sensitive Concerns in Software Engineering Practice. In *36th International Conference on Software Engineering, ICSE Companion.*

[7] Susanne Barth and Menno de Jong. 2017. The Privacy Paradox–Investigating Discrepancies between Expressed Privacy Concerns and Actual Online Behavior–A Systematic Literature Review. *Telematics and Informatics* (2017).

[8] Jeremy Bentham. 1995. Panopticon letters. *The Panopticon Writings* (1995), 29–96.

[9] Jeremy Bentham. 2011. The Works of Jeremy Bentham, vol. 4 (Panopticon, Constitution, Colonies, Codification). http://oll.libertyfund.org/titles/bentham-the-works-ofjeremy-bentham-vol-4. (2011), 65–123 pages.

[10] Danah Boyd and Nicole Ellison. 2007. Social network sites: definition, history, and scholarship. *IEEE Engineering Management Review* 3, 38 (2007), 16–31.

[11] Finn Brunton and Helen Nissenbaum. 2015. *Obfuscation: A user's guide for privacy and protest.* Mit Press.

[12] Gilles Deleuze. 1992. Postscript on the Societies of Control. *October* 59 (1992), 3–7.

[13] Gilles Deleuze and Félix Guattari. 1987. *A thousand plateaus: Capitalism and schizophrenia.* Bloomsbury Publishing.

[14] D Harold Doty and William H Glick. 1994. Typologies as a unique form of theory building: Toward improved understanding and modeling. *Academy of Management Review* 19, 2 (1994), 230–251.

[15] Joerg Evermann and Riddhi Mistry. 2008. System Analysis as Scientific Inquiry. *AMCIS 2008 Proceedings* (2008), 154.

[16] Michel Foucault. 1977. *Discipline & punish: The birth of the prison.* Vintage.

[17] Batya Friedman and Helen Nissenbaum. 1996. Bias in computer systems. *ACM Transactions on Information Systems (TOIS)* 14, 3 (1996), 330–347.

[18] Maša Galič, Tjerk Timan, and Bert-Jaap Koops. 2016. Bentham, Deleuze and Beyond: An Overview of Surveillance Theories from the Panopticon to Participation. *Philosophy & Technology* (2016), 1–29.

[19] Mohamad Gharib, Paolo Giorgini, and John Mylopoulos. 2017. Towards an Ontology for Privacy Requirements via a Systematic Literature Review. In *International Conference on Conceptual Modeling.* Springer, 193–208.

[20] Shirley Gregor. 2006. The nature of theory in information systems. *MIS quarterly* (2006), 611–642.

[21] Shirley Gregor and Alan R Hevner. 2013. Positioning and presenting design science research for maximum impact. *MIS quarterly* 37, 2 (2013), 337–355.

[22] Kevin D Haggerty. 2006. Tear down the walls: on demolishing the panopticon. *Theorizing surveillance: The panopticon and beyond* (2006), 23–45.

[23] Kevin D Haggerty and Richard V Ericson. 2000. The surveillant assemblage. *The British journal of sociology* 51, 4 (2000), 605–622.

[24] Alan R Hevner. 2007. A three cycle view of design science research. *Scandinavian journal of information systems* 19, 2 (2007), 4.

[25] Sean P Hier. 2002. Probing the Surveillant Assemblage: on the dialectics of surveillance practices as processes of social control. *Surveillance & Society* 1, 3 (2002), 399–411.

[26] Rob Kitchin. 2014. The real-time city? Big data and smart urbanism. *GeoJournal* 79, 1 (2014), 1–14.

[27] Susan Landau. 2013. Making sense from Snowden: What's significant in the NSA surveillance revelations. *IEEE Security & Privacy* 11, 4 (2013), 54–63.

[28] Bruno Latour. 1999. On recalling ANT. *The Sociological Review* 47, S1 (1999), 15–25.

[29] David Lyon. 2007. *Surveillance studies: An overview.* Polity.

[30] Amjad Mahfuth, Salman Yussof, Asmidar Abu Baker, and Nor'ashikin Ali. 2017. A systematic literature review: Information security culture. In *Research and Innovation in Information Systems (ICRIIS), 2017 International Conference on.* IEEE, 1–6.

[31] Alice E Marwick. 2012. The public domain: Social surveillance in everyday life. *Surveillance & Society* 9, 4 (2012), 378.

[32] Daniel Mellado, Carlos Blanco, Luis E Sánchez, and Eduardo Fernández-Medina. 2010. A systematic review of security requirements engineering. *Computer Standards & Interfaces* 32, 4 (2010), 153–165.

[33] Robert King Merton. 1968. On Sociological Theories of the Middle Range. In *Social theory and social structure.* Simon and Schuster, 448–459.

[34] Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. The ethics of algorithms: Mapping the debate. *Big Data & Society* 3, 2 (2016), 2053951716679679.

[35] Daniel L Moody and Graeme G Shanks. 1994. What makes a good data model? Evaluating the quality of entity relationship models. In *International Conference on Conceptual Modeling.* Springer, 94–111.

[36] John Mylopoulos and Hector J Levesque. 1984. An Overview of Knowledge Representation. In *On Conceptual Modelling.* Springer, 3–17.

[37] Ken Peffers, Tuure Tuunanen, Marcus A Rothenberger, and Samir Chatterjee. 2007. A design science research methodology for information systems research. *Journal of management information systems* 24, 3 (2007), 45–77.

[38] Nicolas Prat, Isabelle Comyn-Wattiau, and Jacky Akoka. 2014. Artifact Evaluation in Information Systems Design-Science Research-a Holistic View.. In *PACIS.* Citeseer, 23.

[39] Jan Pries-Heje, Richard Baskerville, and J Venable. 2008. Strategies for design science research evaluation. *ECIS 2008 proceedings* (2008), 1–12.

[40] Richard M Ryan and Edward L Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology* 25, 1 (2000), 54–67.

[41] Philip Schofield. 2009. *Bentham: A Guide for the perplexed.* Bloomsbury Publishing.

[42] Dag IK Sjøberg, Tore Dybå, Bente CD Anda, and Jo E Hannay. 2008. Building theories in software engineering. In *Guide to advanced empirical software engineering.* Springer, 312–336.

[43] Klaas-Jan Stol and Brian Fitzgerald. 2013. Uncovering theories in software engineering. In *Software Engineering (GTSE), 2013 2nd SEMAT Workshop on a General Theory of.* IEEE, 5–14.

[44] Colin Tankard. 2012. Cultural issues in security and privacy. *Network security* 2012, 11 (2012), 5–8.

[45] Ron Weber et al. 1997. *Ontological foundations of information systems.* Coopers & Lybrand and the Accounting Association of Australia and New Zealand Melbourne.

[46] Karl E Weick. 1995. What theory is not, theorizing is. *Administrative Science Quarterly* (1995), 385–390.

[47] Roel Wieringa, Maya Daneva, and Nelly Condori-Fernandez. 2011. The structure of design theories, and an analysis of their use in software engineering

experiments. In *Empirical Software Engineering and Measurement (ESEM), 2011 International Symposium on*. IEEE, 295–304.

[48] Karen Yeung. 2017. Hypernudge: Big Data as a mode of regulation by design. *Information, Communication & Society* 20, 1 (2017), 118–136.

[49] Shoshana Zuboff. 2015. Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology* 30, 1 (2015), 75–89.